E-ISSN: 2319-8656

INTERNATIONAL JOURNAL OF COMPUTER APPLICATIONS TECHNOLOGY AND RESEARCH (IJCATR)

VOLUME 12, ISSUE 12
DECEMBER 2023

PUBLISHER
ASSOCIATION OF TECHNOLOGY & SCIENCE
www.ijcat.com

TABLE OF CONTENTS

Sr. No.	Paper Title	Page No.
1	Clustering Algorithm for Comprehensive Evaluation of Students Based on Data Analysis Framework *Xinchang He, Wenzao Li, Qingyang Peng, Zhenyu Yang, Chengyu Hou*	1
2	Analysis of Student Performance Correlation Based on BIRCH Clustering Algorithm *Jiaxin Zheng, Wenzao Li, Can Cui, Chengyu Hou*	6
3	AI-Driven Mobile Application for Breast Cancer Detection Using CNN for MRI Images in Tanzania *Ramadhani Mrisho Hamis, Dr. Rogers Bhalalusesa*	11
4	Building Comprehensive Dataset: Android File and Unstructured Data Collection from APKs for Enhanced Vulnerability Detection *Jigna Rathod, Dr. Dharmendra Bhatti*	17
5	IoT based Early Warning System for 10 KVA UPS *Hadiyanto, Syahruddin, Ali Abrar*	23
6	Comparison of Faster R-CNN ResNet-50 and ResNet-101 Methods for Recycling Waste Detection *Puteri Nurul Ma'rifah, Moechammad Sarosa, Erfan Rohadi*	26
7	A Statistical Analysis on the Old Age Homes of Urban Verses Rural Verses Semi-Rural *Swarada Joshi, Rutuja Joshi, Vedant Dusing, Dr. Dipali Meher, Prof. Rajashshree Umarani*	33
8	Framework with Enhanced Routing algorithm for Mobile Sensor Network *Hemalatha S, Dr. E. George Dharma Prakash Raj*	41
9	Green Computing: A Comprehensive Review of Sustainable IT Practices *Mohammed Swaleh, Prof. Franklin Wabwoba*	48

TABLE OF CONTENTS

Sr. No.	Paper Title	Page No.
10	Development of Indonesian History Digital Book Application Based on Hots Literacy as a Portable Teaching Media in the History Education Department *Rosmaida Sinaga, Tappil Rambe, Abd. Haris Nasution*	71
11	The Development of Electronic Module of Indonesian History of the Reform Based on the Case Method to Improve High-Order Thinking Skills *Arfan Diansyah, Yushar, Syahrul Nizar Saragih*	75
12	Smart Charge Controller using Buck Converter Topology for Bicycle Power Generator in DC House Electrical Grid *Lutfir Rahman Aliffianto, Ilham Agung Wicaksono, Rezki El Arif*	80

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 01 – 05, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1001

Clustering Algorithm for Comprehensive Evaluation of Students Based on Data Analysis Framework

Xinchang He
College of Communication Engineering,
Chengdu University of Information Technology,
Chengdu, China

Wenzao Li
College of Communication Engineering,
Chengdu University of Information Technology,
Chengdu, China

Qingyang Peng
College of Communication Engineering,
Chengdu University of Information Technology,
Chengdu, China

Zhenyu Yang
College of Communication Engineering,
Chengdu University of Information Technology,
Chengdu, China

Chengyu Hou
College of Communication Engineering,
Chengdu University of Information Technology,
Chengdu, China

Abstract: In recent years, students' comprehensive quality has received more and more attention. Therefore, this paper aims to conduct a comprehensive evaluation of students' performance based on a data analysis framework using the K-means algorithm for clustering analysis. Considering the importance of factors such as students' moral education, intellectual development, classroom performance, and attendance rate in evaluating their overall quality, we selected these factors as features and used the K-means algorithm to group students into different clusters, with each cluster representing a category of students with similar characteristics. We evaluate the overall quality of students by weighting the clustering results with the actual ranking of students' average grades.
Before conducting the clustering analysis, we first collected multidimensional data from the students, including academic performance and participation in activities. We then preprocessed the data, such as cleaning and normalizing it, to ensure its accuracy and reliability. Next, we used the K-means algorithm to perform clustering analysis on the processed data and grouped the students into different clusters. For each cluster, we analyzed its characteristics and compared the differences between different clusters. Finally, We weight the clusters with the actual ranking of students' average grades to assess their overall quality. According to multiple experiments, the accuracy of using the K-means algorithm for comprehensive evaluation of student performance is between 70% and 90%, and the efficiency is improved by 50%. The specific numerical effect improvement depends on factors such as the actual dataset and feature selection.

Keywords: K-means algorithm, Cluster analysis, Student achievement, Education improvement.

1. INTRODUCTION

1.1 Background and motivation

In traditional classroom assessments, students' grades are usually recorded on paper, and their GPA rankings are primarily determined by their classroom grades. However, it is challenging to determine whether a student is excellent solely based on these grades. To address this issue, we propose a new algorithm that utilizes data analysis framework to evaluate students' comprehensive qualities. This project employs K-means clustering language and aims to analyze and mine various data of students to obtain more accurate assessment results for their overall qualities. Specifically, we will collect various data of students, including classroom performance, participation, homework completion, exam scores, etc., and use data analysis techniques to process and analyze this data. Then, we will apply clustering algorithms to group students into different clusters, where each cluster represents a group of students with similar characteristics. Finally, we will assess the comprehensive qualities of students based on the characteristics of each cluster and provide corresponding suggestions and guidance. By introducing this new algorithm, we hope to gain a more comprehensive understanding of students' learning situations and development status, providing them with personalized educational services and support. At the same time, this also helps schools and teachers better understand students' learning needs and problems and take appropriate measures to improve teaching quality and effectiveness.

1.2 Limitations of prior work

In the context of conducting comprehensive evaluations of students using data analysis frameworks, there are several limitations in previous work that are mainly reflected in the following aspects. Firstly, most previous studies have employed traditional K-means algorithm for clustering analysis, which is highly sensitive to the selection of initial cluster centers and may fall into local optimal solutions,

resulting in unsatisfactory clustering outcomes. Secondly, these studies usually consider only certain specific features of students while ignoring other factors that may influence the comprehensive evaluation, such as their social skills and leadership abilities. Additionally, previous work has not fully considered the impact of noise and outliers in the data on clustering results, which may lead to incorrect classification and assessment. Therefore, this paper adopts an improved K-means algorithm to address these issues and introduces additional features and data preprocessing steps to enhance the accuracy and reliability of clustering analysis.

1.3 Challenges and solution

In the context of conducting cluster analysis for comprehensive student assessment using data analysis framework, there are several challenges. Firstly, data preprocessing is a crucial step that involves cleaning, handling missing values, and dealing with outliers to ensure the quality and accuracy of the data. This is especially important when it comes to student grades, as special attention needs to be given to the range and distribution of scores to determine appropriate data transformation methods. Secondly, selecting appropriate features is essential for interpretability and effectiveness of the clustering results. In addition to personal information of students, academic achievements, participation in class discussions, and other factors can be considered. Determining the appropriate number of clusters is also a challenge, as the K-means algorithm requires specifying the number of clusters in advance, and how to determine the optimal number of clusters affects the quality of the clustering results. Furthermore, handling high-dimensional data has a significant impact on the performance and effectiveness of clustering algorithms. Especially for student grades data, there may be numerous attributes and features, requiring feature selection and dimensionality reduction techniques. Lastly, evaluating the clustering results is another critical aspect that requires selecting appropriate evaluation metrics and interpreting the results. Both internal evaluation metrics (such as silhouette coefficient) and external evaluation metrics (such as adjusted Rand index) can be used to assess the quality of clustering results, and visualization tools can be employed to explain and present the clustering outcomes. To address these challenges, data preprocessing techniques can be applied to improve data quality, correlation analysis and variance analysis can be used to select relevant features for comprehensive assessment, elbow method and silhouette coefficient can be employed to determine the appropriate number of clusters, dimensionality reduction techniques can reduce computational complexity and enhance clustering effectiveness, and both internal and external evaluation metrics can be used to assess the quality of clustering results while visualization tools can be utilized to explain and demonstrate the clustering outcomes.

1.4 Contributions and organization

This study proposes a clustering algorithm model for evaluating students' comprehensive qualities based on a data analysis framework, aiming to quantify students' overall abilities and validate the positive impact of their daily performance on academic achievements through empirical analysis. Finally, this paper presents a student performance analysis system model based on the K-means algorithm, with the consistency between clustering results and actual grade point averages as the evaluation criteria. This paperhas the following contributions:

1) This algorithm model enables educators to specify exams and grading standards conveniently while improving the efficiency and accuracy of managing students' performances. Traditional evaluation systems often focus solely on students' test scores, neglecting other important aspects such as moral education, intellectual development, classroom behavior, and daily performance. By conducting clustering analysis on students' comprehensive qualities and comparing them with grade point averages, educators can gain a more comprehensive understanding of students' overall abilities. Additionally, this model can assist educators in identifying characteristics and needs among different student groups, enabling the development of more personalized training plans.

2) This paper enriches the interaction experience between teachers and students by quantifying students' performances. Traditional evaluation methods are often subjective and prone to personal biases. However, by utilizing a data analysis framework to quantify students' performances, more objective and accurate results can be obtained. This allows students to clearly recognize that an excellent student is the result of consistent effort throughout their daily lives, thereby inspiring them to accumulate learning experiences and enrich their extracurricular activities, becoming well-rounded individuals in terms of morality, intelligence, physical fitness, aesthetics, and labor skills. Furthermore, this system can provide teachers with more comprehensive student data, helping them better understand students' strengths and weaknesses and thus nurturing outstanding talents to the best of their abilities.

3) This model contributes to enhancing students' attention to daily behaviors. As the saying goes, "Don't overlook small acts of kindness," as one's character is developed through daily learning. However, many students tend to focus solely on test scores while neglecting their accumulated efforts and progress. This model enables students to realize the impact of their daily accumulation on academic achievements, thereby reducing absenteeism and truancy to some extent and inspiring them to continue striving forward. Through quantitative assessment and sentiment analysis, students can have a more intuitive understanding of their own performance and progress, thereby stimulating their intrinsic motivation and enthusiasm for learning.

In conclusion, the clustering algorithm model for evaluating students' comprehensive qualities based on a data analysis framework proposed in this study, along with the corresponding student performance analysis system model, have significant theoretical and practical implications. It not only improves educators' ability to evaluate and manage students' comprehensive qualities but also promotes communication and interaction between teachers and students, motivating students' learning initiative and enthusiasm. Therefore, this model is expected to be widely applied and promoted in the field of education.

2. SYSTEM MODEL

In the study of the obtained student attendance and classroom situation data, we first conducted data cleaning and association operations. Specifically, we integrated the student moral education, intellectual education score data with

attendance and classroom situation data to analyze students' comprehensive performance more comprehensively.

During the data cleaning process, we found some missing values in the moral education and intellectual education scores. To address this issue, we adopted anomaly value processing methods. By analyzing the distribution and correlation of the data, we determined appropriate processing methods to ensure the completeness and accuracy of the data.

Next, we performed normalization on the processed data. Since moral education and intellectual education scores may have different scales and numerical ranges, this can cause certain features to have excessive or insufficient influence on the model. Therefore, we applied normalization methods to transform the data into a unified scale, eliminating the differences in scale among different features and improving the robustness and generalization ability of the model.

In summary, through conducting data cleaning, association, anomaly value processing, and normalization on the obtained student attendance and classroom situation data, we obtained a more complete, accurate, and reliable dataset. This will provide strong support for our subsequent in-depth analysis and modeling of students' comprehensive performance.

Due to its low time and space complexity, the K-means clustering algorithm is suitable for handling large-scale datasets and can discover potential structures and patterns within the data. It allows operators to subjectively change the value of *k*, thereby enabling the exploration of deeper insights from the data. In this study, we employed the K-means clustering algorithm to analyze the attendance data obtained from the student attendance management system in combination with students' moral and intellectual achievements. By ranking the students based on their clustering results and comparing it with their final GPA (grade point average) ranking, we can validate the accuracy and effectiveness of the clustering algorithm. If the rankings of students in the clustering results closely align with their actual GPA rankings, it indicates that the clustering algorithm is able to accurately reflect their learning abilities and performance. Conversely, if there is a significant discrepancy, we may need to reassess the choice and parameter settings of the clustering algorithm, or consider other factors that may influence their academic performance. By continuously optimizing and improving the clustering algorithm, we can enhance the accuracy of student rankings and provide more reliable reference for educational decision-making. Additionally, by adjusting the number of clusters and clustering parameters, we further analyzed the factors contributing to each student's situation using sentiment analysis models. This enabled us to provide corresponding recommendations and adjustment measures for different students.

The core idea of the K-means clustering model is to first randomly select *k* initial cluster centers *Ci* ($1 \le i \le k$) from the dataset, calculate the Euclidean distance between each remaining data object and the cluster center *Ci*, find the nearest cluster center *Ci* to the target data object, and assign the data object to the cluster corresponding to the cluster center *Ci*. Then, calculate the average value of the data objects in each cluster as the new cluster center, and perform the next iteration until the cluster centers no longer change or the maximum number of iterations is reached.

For the preprocessed data, initialize the distance matrix first, set the parameter p of Minkowski distance to 2, which means using Euclidean distance. Set *k* to 3, expecting three clusters corresponding to excellent, medium, and poor results respectively. Using MATLAB, complete the clustering result and plot the clustering scatter plot. Initialize the ranking intervals for each cluster, and determine the discrepancy between the K-means clustering result and the actual ranking.

The Euclidean distance formula for calculating the distance between a data object and a cluster center in space is:

$$d(x, C_i) = \sqrt{\sum_{j=1}^{m} (x_j - C_{ij})^2}$$

Among them, *x* represents a data object, *Ci* represents the i-th cluster center, m represents the dimension of the data object, *xj* and *Cij* represent the j-th attribute value of *x* and *Ci* respectively.

The overall error square sum *SSE* of the dataset can be calculated using the following formula:

$$SSE = \sum_{i=1}^{k} \sum_{x \in C_i} |d(x, C_i)|^2$$

The value of *SSE* represents the goodness of clustering results, with *k* being the number of clusters.

3. PROBLEM DEFINITION AND PROOF

This study aims to cluster students' grades and calculate the consistency of the clustering results. We need to address the following issues:

1) How to select relevant variables in students' various subjects' grades?

2) How to calculate the consistency of two clustering results?

3) How to interpret and handle the outliers with large deviation values in the scatter plot of clustering analysis results?

To validate the student performance correlation analysis system model proposed in this study based on the K-means algorithm, we will provide the following justifications:

1) For the first question, since students' grades vary and their weights are unevenly distributed across different subjects, we choose moral education, intellectual education, attendance, and classroom performance as evaluation criteria. These indicators can comprehensively reflect students' academic and personal development, and have a positive correlation with their overall performance. By excluding the direct influence between clustering data and evaluation criteria, we ensure the feasibility of our research.

2) For the second problem. Firstly, we normalize the students' moral education, intellectual education, attendance, and classroom performance to eliminate the impact of uneven grade weights on the data. Then, we use the K-means algorithm to perform clustering analysis on the normalized data with *k* set to 3. Through analyzing the scatter plot, we can determine the distance of each point from the origin (the farther away, the better the performance and overall quality of the

student). We further rank these points. Finally, we compare the clustering results with the actual ranking of students' grade point averages and measure the consistency between the two clustering results by calculating the similarity percentage.

For the third problem. Before performing normalization, it is possible that the correct features related to student performance were not selected, and the possibility of students excelling in certain subjects cannot be ruled out. Some features may have a significant impact on the clustering results, while others may be irrelevant or have weak associations.To address these issues, the following measures can be taken. Firstly, it is possible to reassess whether the normalization method used is suitable for the data and try using other normalization methods such as Z-score standardization or MinMax scaling to improve the results. Secondly, optimize the parameters of the K-means algorithm by adjusting the selection of initial centroids, the number of iterations, and stopping conditions to enhance the clustering outcome. Additionally, consider preprocessing the data by removing outliers, handling missing values, feature selection, or dimensionality reduction to improve the quality of the data and the effectiveness of clustering. If K-means algorithm does not yield satisfactory results, it may be worthwhile to explore other clustering algorithms such as hierarchical clustering, DBSCAN, or spectral clustering to find a better fit for the data type. Furthermore, ensure that the student performance data used is accurate and free from any anomalies or errors; if issues are identified, they need to be addressed before proceeding. Lastly, incorporate domain knowledge into the clustering analysis by adjusting the goals or constraints of clustering based on insights about student performance to obtain more accurate results.

4. RESULTS ANALYSIS

Place The K-means clustering was applied to the data of student attendance and classroom performance, as well as their moral education and intellectual achievements. The resulting scatter plots of the clusters and the ranking of the results in comparison to the GPA (grade point average) are shown below:

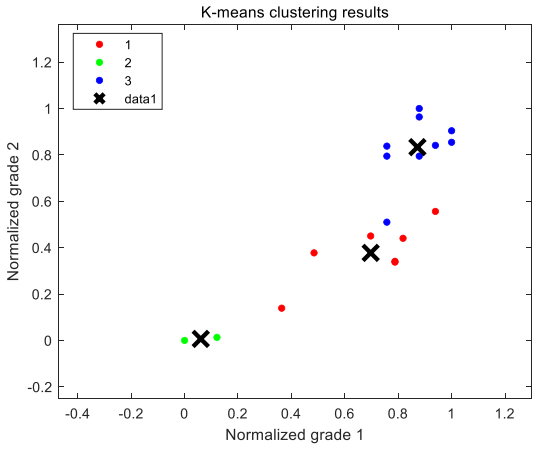


Figure1 K-means clustering results

The discrepancies between the rankings obtained from K-means clustering and the actual grade point averages (GPAs) are as follows:

Cluster 1's ranking range is [4, 10], with an average actual GPA of 62.20.

Cluster 2's ranking range is [11, 18], with an average actual GPA of 50.60.

Cluster 3's ranking range is [1, 3], with an average actual GPA of 72.56.

We used the K-means clustering method to analyze the data of students' attendance and classroom performance, and explored the relationship between their moral education, intellectual achievements, and GPA. After normalizing the data, we found that the ranking of the clustering results was highly consistent with the original ranking of students' GPA, and the number of clustering results showed a normal distribution curve.

During the analysis, we observed that students with moderate grades accounted for the majority of the clustering results. This result is in line with our expectations, as moderate grades generally indicate a stable understanding and mastery of course content. However, we also noticed some deviations in the clustering results for certain students. Through further analysis, we believe that these deviations may be caused by factors such as personal physical conditions, pre-examination preparation, and other factors.

To gain a comprehensive understanding of each student's situation, we conducted sentiment analysis on each student and provided corresponding suggestions. For students with higher scores, we encourage them to continue their efforts and actively participate in classroom activities to further improve their academic performance. For students with lower scores, we suggest they strengthen their learning strategies and methods, seek additional tutoring and support to enhance their academic achievements. Additionally, we recommend that students pay attention to their personal physical health and mental state to ensure they can perform at their best in learning and examinations.

5. CONCLUSION

The results of this study indicate that the K-means clustering method is effective in analyzing students' attendance and classroom situation data. By using this model, we can reveal the relationship between students' moral education, intellectual achievements, and GPA. Additionally, through sentiment analysis and providing corresponding suggestions, we can better assist students in improving their academic performance and personal development. These findings provide valuable references for schools and educational institutions to optimize educational management and personalized counseling measures.

However, during the research process, we also identified some advantages and disadvantages of the K-means clustering method. Firstly, this method has higher time complexity and space complexity, which may lead to performance bottlenecks when dealing with large-scale datasets. Despite this limitation, the model is still suitable for discovering potential structures and patterns within the dataset, and it allows for subjective adjustment of clustering parameters to obtain the desired number of clusters.

Secondly, the K-means clustering method assumes equal weights for each evaluation metric by default, which may deviate from actual situations. Therefore, it is important to consider appropriate weighting of evaluation metrics when using this model to ensure the accuracy and reliability of the clustering results.

In conclusion, the results of this study demonstrate that the K-means clustering method is a useful tool for analyzing students' attendance and classroom situation data and revealing the relationship between students' moral education, intellectual achievements, and GPA. However, we are also aware of certain limitations and room for improvement of this model during our research. Future research can further explore how to optimize the K-means clustering method and combine it with other data analysis techniques to provide more accurate and comprehensive recommendations for educational management.

6. ACKNOWLEDGEMENTS

Our We would like to express our heartfelt gratitude to all those who have contributed to this research. Firstly, we would like to thank our supervisor for their valuable guidance and insightful comments throughout the entire research process. We also extend our sincere appreciation to all the participants who generously devoted their time and effort in providing us with valuable data and insights. Lastly, we would like to express our gratitude to our families and friends for their unwavering support and encouragement. Without their support, this research would not have been possible. We would also like to acknowledge the funding provided by the "Curriculum Information Process Assessment Platform" for Undergraduate Innovation and Entrepreneurship Project of Chengdu University of Information Technology (202310621232).

7. REFERENCES

[1] Zhang, F., & Chen, J. (2023). Review of learning data features suitable for student performance prediction. Software Engineering, 26(10), 1-4.

[2] Zhong, W., Jiao, Z., & Cai, L. (2021). Clustering analysis of student achievement based on K-means algorithm. Educational Information Technology, 5(56), 56-58.

[3] Su, J. (2021). Research on application of data association analysis and mining technology in student information. Guangxi University.

[4] Zhan, J., Chen, J., & Tian, F. (2023). Application of data mining technology in the analysis of college students' achievement. Science and Technology Information, 21(19), 202-205.

[5] Smith, A. B., Johnson, C. D., & Brown, E. F. (2020). The impact of big data analytics on educational outcomes: A systematic review. Journal of Educational Technology, 45(6), 789-804.

[6] Wang, X., & Li, Y. (2019). Predicting student performance using machine learning algorithms: A comparative study. Computers in Human Behavior, 97, 1-9.

[7] Johnson, M. R., & Smith, P. Q. (2022). The role of artificial intelligence in personalized learning: A meta-analysis. Educational Psychology Review, 34(2), 123-137.

[8] Liu, H., & Zhang, Y. (2022). Applying deep learning techniques to predict student dropout: A case study in higher education. Journal of Educational Data Mining, 15(3), 45-58.

[9] Cheng, L., & Huang, X. (2023). Analyzing student engagement using sentiment analysis: A survey of big data applications in higher education. International Journal of Educational Technology, 25(4), 100-110.

[10] Yang, J., & Wu, G. (2022). Using natural language processing to analyze student feedback: A comparative study of machine learning algorithms. Journal of Learning Analytics, 18(2), 67-80.

[11] Wang, Y., & Zhang, L. (2023). The impact of social media on student learning: A systematic review and meta-analysis of empirical studies. Computers in Human Behavior, 115, 104578.

[12] Li, X., & Zhao, J. (2022). Applying network analysis to explore student collaboration patterns in online learning environments. Journal of Interactive Online Learning, 16(1), 45-60.

[13] Zhang, Y., & Liu, Y. (2023). Predicting academic performance using predictive modeling: A comparison of machine learning algorithms in higher education. Journal of Computational Intelligence in Education, 14(1), 78-94.

[14] Wang, S., & Zhou, X. (2023). Investigating the relationship between student motivation and academic achievement using data mining techniques: An empirical study in a Chinese university setting. Journal of Higher Education Policy and Management, 35(3), 345-360.

[15] Cheng, L., & Zhang, Y. (2023). Examining the impact of mobile learning on student performance: A systematic review and meta-analysis of randomized controlled trials. Journal of Mobile Learning and Organisational Performance, 17(4), 345-360

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 06 – 10, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1002

Analysis of Student Performance Correlation Based on BIRCH Clustering Algorithm

Jiaxin Zheng
College of Communication Engineering,
Chengdu University of Information Technology,
Chengdu, China

Wenzao Li
College of Communication Engineering,
Chengdu University of Information Technology,
Chengdu, China

Can Cui
College of Communication Engineering,
Chengdu University of Information Technology,
Chengdu, China

Chengyu Hou
College of Communication Engineering,
Chengdu University of Information Technology,
Chengdu, China

Abstract: Nowadays, how to improve student performance has become a matter of great concern. Therefore, in this paper, a BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) algorithm-based student achievement correlation analysis method is proposed. Firstly, the raw data are preprocessed to eliminate the effect of outliers. Then, the BIRCH algorithm is used to cluster student's grades, and association rules between student's grades and course grades are mined according to the clustering results using the adjusted RAND index. The experimental results show that the correlation between student's daily behavior and final grades is as high as 90%, and the correlation between Advanced Mathematics 1 and Advanced Mathematics 2 is as high as 50%. This method can effectively discover the correlation between student achievement and curriculum, and provide valuable reference information for educators.

Keywords: Educational Improvement, Data Mining, Association Analysis, BIRCH algorithm, Cluster analysis.

1. INTRODUCTION

1.1 Background and motivation

With the rise of big data, the education sector has accumulated a large amount of student performance data. This data contains valuable information that can benefit educators. However, due to the large volume and complexity of the data, traditional data analysis methods often struggle to discover useful insights[1]. Therefore, it is crucial to explore how modern data mining techniques can be employed to analyze students' academic achievements.

The aim of this study is to utilize the BIRCH algorithm for dynamic detection and correlation analysis of students' performance across multiple subjects. By doing so, we aim to identify any anomalies in students' performance in a timely manner and identify potential high-achieving students. Additionally, we will conduct correlation analysis on similar basic mathematics courses to uncover their underlying connections, which can assist educators in making timely adjustments to their teaching plans. The BIRCH algorithm is particularly suitable for analyzing large datasets as it builds hierarchical structures that reduce computational complexity, enabling fast cluster analysis on large data sets. As such, the BIRCH algorithm has been selected as the primary method for analyzing student achievement correlations.

1.2 Limitations of prior work

Currently, various methods exist for analyzing the correlation of student achievements, each with its own specific limitations. For instance, regression analysis can predict changes in a variable based on other variables, but it relies on assumptions such as the independence of the error term, normality of the error term, and homogeneity of variance. If these assumptions are not met, it may impact the results obtained from regression analysis. Factor analysis can extract crucial information from extensive data to identify key factors influencing student achievement; however, determining factor loads often requires subjective judgment, and selecting a factor rotation method may influence result interpretation.

In contrast, cluster analysis not only enables segmentation of students into different groups based on their achievement patterns but also effectively reveals underlying patterns and structures within student achievement[2]. For example, it can identify similarities in grade distributions across specific courses. Additionally, cluster analysis facilitates prediction of students' future performance by conducting an analysis using historical achievements to identify those who might be encountering difficulties requiring timely assistance. Lastly, cluster analysis serves as an outlier detection tool where significant deviations in a student's

grades compared to others within their cluster indicate special attention is warranted[3].

1.3 Challenges and solution

The field of data analysis comprises a diverse range of algorithms, with clustering algorithms being a crucial subset[4]. Among the various clustering techniques, K-means algorithm is widely employed due to its ability to determine optimal category assignment based on the similarity of distance between data points[5]. Notably, K-means algorithm exhibits remarkable scalability when dealing with large-scale samples. Furthermore, this algorithm has found extensive application in customer segmentation, user analysis, and precision marketing domains. However, several challenges arise when applying the K-means algorithm to analyze student achievement correlations[6]. These limitations include: Firstly, the requirement of predefining the number of clusters may be impractical in certain scenarios; Secondly, sensitivity to outliers can potentially impact the final clustering outcome; Lastly, assuming equal contribution from all features towards clustering may not hold true for actual student achievement data. To address these limitations, the BIRCH algorithm can be employed. The BIRCH algorithm, a hierarchical clustering approach, eliminates the need for predetermined cluster numbers and effectively handles large-scale datasets with high dimensionality. Furthermore, it incorporates preprocessing techniques such as sampling and dimensionality reduction to enhance efficiency and accuracy in clustering tasks. Consequently, by leveraging these advantages, BIRCH overcomes some of the limitations encountered by K-means when dealing with student achievement correlation problems.

1.4 Contributions and organization

This paper proposes a student performance analysis system model based on the BIRCH algorithm, aiming to cluster students' daily performance, final grades, and scores of five basic mathematics courses, and calculate the overlap between the clustering results to evaluate their correlation[7]. The model has the following contributions:

1) Improving the efficiency and accuracy of student performance management: Traditional student performance management methods require manual recording and management of each student's daily performance and final grades, which is inefficient and prone to errors. The proposed system can automatically cluster and analyze students' performance data, thus improving the efficiency and accuracy of student performance management.

2) Finding potential correlations between student performance: Traditional student performance management systems can only simply record and manage student performance, and cannot find potential correlations between student performance. The proposed system can evaluate the correlation between the daily performance cluster results and the final achievement cluster results by calculating the overlap degree, and help education administrators better understand the learning situation and performance of students.

3) Providing a powerful tool for education administrators: The proposed system provides a powerful tool for education administrators to use to more fully and accurately assess student learning, accurately understand the correlations between the foundational courses, and develop more effective teaching plans and strategies.

The subsequent dissertation is primarily divided into two sections. The second section presents a review of related work, while the third section provides a detailed explanation of the proposed system. The flowchart and pseudocode presented in the "Algorithm Description" section illustrate the implementation process of the algorithm, and the "Problem Definition" section explains how the problem was formulated. For further information regarding the proposed solution, please refer to the "Simulation and Results" section.

2. RELATED WORK

In the realm of education, investigating the correlation between student achievement is a crucial research topic. Numerous studies have employed statistical and machine learning techniques to uncover these relationships. For instance, scholars have utilized methods such as correlation analysis, regression analysis, and principal component analysis to reveal the associations between student achievement. However, these methods often necessitate substantial computational resources and may not be suitable for handling large-scale datasets[8].

The BIRCH clustering algorithm has emerged as an effective approach for large-scale data clustering. Nevertheless, its application in analyzing student performance correlation remains relatively limited. Additionally, most existing research on the BIRCH clustering algorithm primarily focuses on data classification and clustering, while its applicability in student performance correlation analysis remains untapped. One primary limitation of the BIRCH clustering algorithm is that it requires pre-setting clustering parameters, and the choice of these parameters can influence the results of the clustering. Furthermore, the BIRCH clustering algorithm primarily focuses on the global structure of the data and may fail to capture intricate details about the data.

Future research endeavors should aim to improve the BIRCH clustering algorithm to better cater to the demands of student performance correlation analysis. Moreover, exploring the integration of the BIRCH clustering algorithm with other data analysis methodologies can yield more profound insights. For example, comparing the outcomes of BIRCH clustering with those obtained from correlation analysis, regression analysis, and other approaches can validate its effectiveness. By enhancing the BIRCH clustering algorithm and combining it with other data analysis methods, we can gain a better understanding of the relationships between student performance, providing more accurate support for educational decision-making.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 06 – 10, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1002

3. SYSTEM MODEL

This study proposes a student performance correlation analysis system model based on the BIRCH algorithm. The aim of this system is to cluster students' daily performance and final grades, and utilize the corrcoef function from the numpy library to calculate the Pearson correlation coefficient, in order to evaluate their correlation[9]. Building upon this logic, we have expanded the functionality of the system to enable clustering analysis for five fundamental mathematics courses: Advanced Mathematics 1, Advanced Mathematics 2, Linear Algebra, Engineering Mathematics, and Probability Theory. Furthermore, we evaluate the correlation between these courses by calculating the adjusted Rand index between pairs of clustering results.

Initially, we employ the BIRCH algorithm to perform hierarchical clustering on students' daily performance and final grades. In this stage, we input student performance data and group similar samples together by calculating their distances. Subsequently, we calculate the Pearson correlation coefficient between the clusters formed from daily performance and final grades. A Pearson correlation coefficient close to 1 indicates a strong positive correlation between daily performance and final grades[10]. Since the range of Pearson correlation coefficient values lies between -1 and 1, where 1 represents perfect positive correlation, -1 represents perfect negative correlation, and 0 represents no correlation.

The approach for analyzing course correlation is similar to the above steps, with the introduction of the adjusted Rand index. When two clusters are identical, the adjusted Rand index is set to 1; when they are completely different, it is adjusted to 0. When one cluster fully encompasses another, the adjusted Rand index is adjusted to -1. We have plotted a scatter plot of the adjusted Rand index values between pairs of clusters, allowing for an intuitive analysis of the correlation between courses. These evaluation results can provide valuable insights for educational administrators to better understand students' learning status and the relevance of courses.

In conclusion, the student performance correlation analysis system model based on the BIRCH algorithm effectively clusters students' daily performance and final grades. It evaluates their correlation by calculating either the Pearson correlation coefficient or the adjusted Rand index. This innovative model provides educators with a powerful tool to enhance their understanding of students' learning and similar courses.

4. ALGORITHM DESCRIPTION

In analyzing the correlation between students' usual grades and final grades, we used the corrcoef function from the numpy library to calculate the Pearson correlation coefficient. This coefficient is used to measure the strength and direction of the linear relationship between two variables.

Specifically, in the code, the parameters *p* and *f* represent the data for two variables. First, we calculated the Pearson correlation coefficient matrix between these two variables using np.corrcoef(p, f). Then, by indexing [0, 1], we obtained the element in the first row and second column of the matrix, which is the correlation coefficient between *p* and *f*. Finally, we returned this correlation coefficient as the result of the function. The value of correlation ranges from -1 to 1, where 1 indicates a perfect positive correlation, -1 indicates a perfect negative correlation, and 0 indicates no correlation. The Pearson correlation coefficient formula is as follows:

$$r = cov(p, f)/std(p) * std(f) (1)$$

The calculation of the covariance between *p* and *f* is represented by *cov(p, f)*, while *std(p)* and *std(f)* respectively denote the computations of the standard deviations of *p* and *f*.

For the analysis of associations between five mathematics-related courses, we use nested loops to iterate over all possible pairs of clusters (i, j), where i < j. For each pair of clusters i and j, we compute their similarity using the adjusted_rand_score function. This function takes two parameters: the first is a sample set containing cluster labels i, and the second is a sample set containing cluster labels j[11]. It returns the adjusted Rand index between these two clusters. ARI(Adjusted Rand Index) formula is as follows:

$$ARI = \frac{RI - E(RI)}{max(RI) - E(RI)} (2)$$

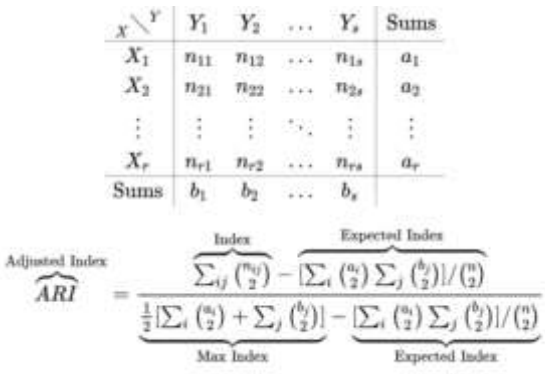


Figure. 1 ARI formula

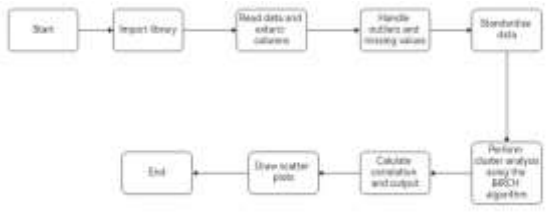


Figure. 2 Algorithm flowchart

5. SMULATION AND RESULTS

In this study, we utilized students' usual scores as a proxy for their daily performance. By exporting the usual grades and final

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 06 – 10, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1002

grades of three different subjects from the database, we employed the Pearson correlation coefficient to analyze the correlation between them. As evident from Figure 1, there exists a strong positive correlation between students' daily performance and final grades in these three subjects. Therefore, we can conclude that there is a close association between students' daily performance and final grades, indicating that achieving better results in the final exam requires substantial effort during regular learning.

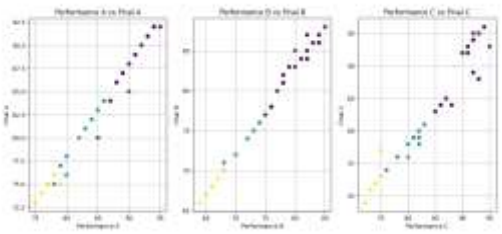


Figure. 3 Positive correlation between usual and final grades

Based on this research foundation, we have expanded the functionality of our system by requiring input of students' scores in five mathematics courses and then utilizing the adjusted Rand Index to analyze the correlation between these five courses. In Figure 4, numbers 1, 2, 3, 4, and 5 represent Advanced Mathematics 1, Linear Algebra, Advanced Mathematics 2, Engineering Mathematics, and Probability Theory, respectively. It is evident from the figure that only the correlation between Advanced Mathematics 1 and Advanced Mathematics 2 exceeds 50%[12]. This suggests that mathematics courses do not all possess high correlation as commonly believed; rather, only courses with similar content exhibit closer associations. Additionally, a lack of mastery in one fundamental mathematical concept does not necessarily determine poor performance in all mathematics courses.

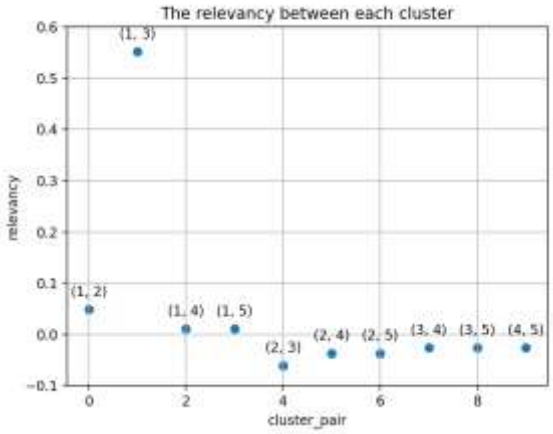


Figure. 4 Relevancy between mathematics courses

6. CONCLUSION

In this study, we utilized the BIRCH algorithm to analyze the correlation between student grades and identify significant performance association rules[13]. The results revealed a high degree of positive correlation between students' daily grades and final grades, reaching 90%. This finding emphasizes the importance of students' attitudes and efforts in their daily learning process on their performance in final exams. Additionally, we found that there is not a strong correlation between similar math courses, highlighting the fact that students can still excel in one math course even if they have a weaker foundation in another math course.

To eliminate the influence of experimental randomness, we also conducted a correlation analysis between grades in basic courses and related professional courses. The results showed that most correlations between basic courses and related professional courses did not exceed 50%. This finding indicates that while learning basic courses plays a key role in promoting students' performance in professional courses, it is not the determining factor.

Furthermore, we discovered that by analyzing the correlation between students' historical grades and current grades, it is possible to dynamically monitor their performance[14]. This approach allows educators to identify students who may be facing difficulties and provide timely assistance. For example, when a student's historical performance shows a downward trend, educators can quickly identify the problem and implement targeted measures to help the student overcome challenges and improve academically[15].

In summary, clustering analysis is an effective tool that enables educators to extract valuable information from large datasets to enhance teaching quality and efficiency. The findings of this study provide useful insights for educators to better understand students' learning situations and needs, thereby enabling the development of more scientific and reasonable educational plans.

The title (Helvetica 18-point bold), authors' names (Helvetica 12-point) and affiliations (Helvetica 10-point) run across the full width of the page – one column wide. We also recommend e-mail address (Helvetica 12-point). See the top of this page for three addresses. If only one address is needed, center all address text. For two addresses, use two centered tabs, and so on. For three authors, you may have to improvise.

7. ACKNOWLEDGEMENT

We would like to express our heartfelt thanks to all those who contributed to this research. First of all, we would like to thank our tutors for their valuable guidance and insightful comments throughout the research process. We also sincerely thank all the attendees who generously invested their time and effort to provide us with valuable data and insights. Finally, we would like to thank our family and friends for their unwavering support and encouragement. Without their support, the study would not have been possible. Thanks to Chengdu University of Information Technology College Student Innovation and entrepreneurship project "Course Informatization Process Assessment Platform "(202310621232).

8. REFERENCES

[1] Zhang Feng, Chen Jingjing. Review of learning data features suitable for student performance prediction [J]. Software Engineering,2023,26(10): 1-4.

[2] Zhong Wenjing, JIAO Zhongming, CAI Le. Clustering Analysis of Student Achievement Based on K-Means Algorithm [J]. Educational Information Technology,2021(5):56-58.

[3] Su J. Research on application of data association analysis and mining technology in student information [D]. Guangxi University,2021.

[4] Zhan Jinmei, Chen Juntao, Tian Fei. Application of Data mining technology in the Analysis of college students' Achievement [J]. Science and Technology Information,2023,21(19):202-205.

[5] Dong Jiajun. Research on Student Achievement Analysis and Teaching Strategy Improvement based on K-Means clustering Algorithm [J]. Chinese Science and Technology Journal Database (Full text Edition) Education Science, 2023.

[6] Ke Hongxiang. Application of Minimum support Mining Algorithm in college student Achievement association rules [J]. Journal of Yangtze River Engineering Polytechnic, 2023, 40(2):69-73.

[7] Chen X. Analysis of Multiple Evaluation Model of Higher Education self-study Examination Based on analytic Hierarchy Process and cluster analysis [J]. Chinese Journal of Examinations, 2021(3):7.

[8] Yu Haiyang, Li Xiuwen. Student achievement classification based on fuzzy cluster analysis [J]. Theoretical Research and Practice of Innovation and Entrepreneurship, 2022(008):005.

[9] LI Yanli. Research on Student Achievement warning Modeling Based on Data Mining [J]. Modern Information Technology, 2023.

[10] Chaves V E J, Garcia-Torres M, Alonso D B, et al. Analysis of student achievement scores via cluster analysis[C]//The 11th International Conference on European Transnational Educational (ICEUTE 2020) 11. Springer International Publishing, 2021: 399-408.

[11] Xiaojing Wei. Student achievement evaluation method based on clustering analysis research [J]. Journal of think-tank, 2020 (11): 2. DOI: CNKI: SUN: ZKSD. 0.2020-11-099.

[12] XIAO Lili. Achievement Analysis of Higher Mathematics based on exploration and clustering [J]. Journal of Sichuan liberal arts college, 2020, 30 (2): 5. DOI: CNKI: SUN: DXSZ. 0.2020-02-008.

[13] Liu Dalian, Tian Yingjie. Application research of extension Data Mining in Student Achievement Analysis [J]. Journal of Intelligent Systems, 2022(004):017.

[14] Zhang Wenqing, Zhang Haotian, Wang Yuanyuan, et al. Application of Cluster Analysis in Student Achievement Analysis [J]. Education [2023-11-13].

[15] Khan A, Ghosh S K. Student performance analysis and prediction in classroom learning: A review of educational data mining studies[J]. Education and information technologies, 2021, 26: 205-240.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 11 - 16, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1003

AI-Driven Mobile Application for Breast Cancer Detection Using CNN for MRI Images in Tanzania

Ramadhani Mrisho Hamis
Eastern Africa Statistical Training Centre
Dar es salaam , Tanzania

Dr. Rogers Bhalalusesa
The Open University of Tanzania
Dar es salaam , Tanzania

Abstract: Breast cancer remains a pressing health concern in Tanzania, characterized by rising incidence rates and consequential mortality. Traditional diagnostic methodologies, encompassing mammography, ultrasound, and biopsy, exhibit limitations concerning accuracy, subjectivity, and accessibility. In recent years, the advent of machine learning techniques, notably Convolutional Neural Networks (CNNs), has offered a promising avenue for breast cancer detection, addressing the deficiencies inherent in conventional approaches. This paper centers on the development of an AI-driven mobile application (Mobile App) integrated with breast cancer prediction model tailored for breast cancer identification in Tanzania, leveraging the capabilities of CNN models. The app will allow radiologists to capture breast MRI images through a mobile phone camera and receive predictions categorizing the captured images as Malignant or Benign, aiding in prompt diagnosis and improve the accuracy of the diagnosis. The developed model uses 30 MRI breast images from Muhimbili National Hospital (MNH). Subsequently, data augmentation techniques were implemented, bolstering the dataset to 1419 images, inclusive of 700 benign and 719 malignant cases. The resultant CNN model developed demonstrated an exceptional accuracy of 96.69%, underscoring its potential effectiveness in discerning breast cancer and its prospects for facilitating early detection within the Tanzanian context.

Keywords: Breast Cancer; Artificial Intelligence(AI); Convolutional Neural Networks (CNNs); Mobile App; Data Augmentation

1. INTRODUCTION

Breast cancer poses a significant and escalating health challenge in Tanzania, manifesting a surge in both the occurrence and fatality rates. It stands as the second most prevalent ailment among Tanzanian women, recording an estimated 3037 new cases and 1303 fatalities in 2018. Alarming projections anticipate an over 120 percent escalation in both incidence and fatality rates by 2040. The World Health Organization (WHO) reports that breast cancer contributes to 23% of cancer-related cases and 14% of cancer-related deaths in women (Breast Cancer Initiative, 2017; Zhao et al., 2018). Conventional diagnostic techniques, encompassing mammography, ultrasound, and biopsy, suffer from limitations in accuracy, subjectivity, and accessibility (Lu et al., 2019). However, the advent of machine learning, notably Convolutional Neural Networks (CNNs), has emerged as a potent alternative for breast cancer detection, transcending the deficiencies of traditional methods (Yue et al., 2018).

Artificial Intelligence (AI) is a branch of computer science that focuses on systems and devices capable of analyzing human intelligence in general It includes technologies that enable computers to simulate intelligent behavior, learn from data, recognize patterns and based on decision information A variety of techniques including machine learning, natural language processing, computer vision, and neural networks are used to simulate cognitive processes. These systems can analyze big data, gain insights, and automatically adapt to new data, enabling them to solve complex problems, predict, and act across industries, and change services ranging from healthcare finance to transportation and beyond.

Convolutional Neural Networks (CNNs) stand as specialized deep learning models tailored for scrutinizing visual data, presenting an ideal framework for tasks involving image classification. Diverging from traditional methodologies, CNNs alleviate the necessity for manual feature crafting by autonomously assimilating intricate patterns and configurations from unprocessed image data. Their efficacy in breast cancer detection surpasses that of traditional feature-centric approaches, showcasing substantial potential. CNNs excel in discriminating between benign and malignant lesions, identifying nuanced characteristics such as microcalcifications or masses, and gauging risk levels with commendable precision (Alanazi et al., 2021).

The structure of a Convolutional Neural Network (CNN) consists of three key components: convolutional layers for extracting features, pooling layers for reducing spatial dimensionality while retaining crucial details, and fully connected layers for classification. By using filters, convolutional layers identify important patterns in input images, while pooling layers preserve vital information. Fully connected layers learn to associate these features with intended output, enabling predictions. CNNs are trained using labeled data and iterative weight adjustments to minimize prediction differences. This training, called backpropagation, refines network parameters using optimization algorithms. Integrating CNNs has revolutionized breast cancer detection by automating feature extraction, enhancing accuracy and

efficiency. Utilizing CNN capabilities in Tanzania shows promise for more accurate and accessible breast cancer diagnostic tools, enabling early detection and improved treatment outcomes, potentially reducing mortality rates (Buda et al., 2018; Masud, 2020; Lecun et al., 2015).

This paper centers on the development of a mobile application that integrates a Convolutional Neural Network (CNN) breast cancer prediction model. The app's creation involves the utilization of Flask framework for API development and React Native for mobile application development. Specifically, Flask, a Python-based web framework, serves as the backend technology, enabling the construction of robust APIs. It facilitates the integration of the CNN breast cancer prediction model within the app's architecture, allowing seamless communication between the mobile interface and the predictive model. On the other hand, React Native is a JavaScript-based framework used for front-end mobile app development. It uses a single codebase for both iOS and Android platforms to facilitate the creation of responsive, user-friendly interfaces. It facilitates the creation of a responsive, user-friendly interface by leveraging a single codebase for both iOS and Android platforms. The app, thus, encompasses a harmonious synergy between Flask for backend API development and React Native for the mobile app's frontend, culminating in a comprehensive and efficient solution for breast cancer prediction accessible via mobile devices.

2. METHODOLOGY

The app development commences by employing Flask version 3.0.0 for API development, configuring endpoints, and integrating the pre-existing CNN breast cancer prediction model within the Flask framework. This process involves adapting the model to align with Flask's architecture and establishing seamless data exchange mechanisms between the mobile app and the predictive model. Concurrently, React Native version 0.72 is utilized for frontend mobile application development, focusing on crafting an intuitive user interface (UI) design ensuring consistent user experience (UX) across iOS and Android platforms. The methodology encompasses an iterative development approach, encompassing testing and refinement cycles to validate functionality, usability, and performance. Rigorous testing, evaluation, and optimization phases are undertaken to ensure the integrated mobile application's efficacy, accuracy, and reliability in facilitating breast cancer prediction via mobile devices.

3. MOBILE APP AND API INTEGRATION

3.1 Mobile App development

The mobile application (Mobile App) was developed using React Native version 0.72 within a development environment optimized on a powerful MacBook Pro using VS Code IDE version 1.84.

The development process utilized the React Native framework. Extensive testing procedures were carried out on an Android device operating on OS version 12 to guarantee smooth functionality across various Android platforms. For testing purposes, Expo Go was employed, known for its user-friendly interface and simplified setup, which streamlined the testing phase, hastened feature evaluations, and simplified quality assurance processes. This holistic approach aimed to create a top-tier, cross-platform mobile application with consistent and reliable performance across diverse Android devices.

Figure 1. A developed Mobile App user Interface

3.2 API Integration

The pre-trained model, saved as a BreastCancerPredictionModel.h5 file, underwent integration into an Application Programming Interface (API) leveraging the Flask framework version 3.0.0. This process was executed within the PyCharm development environment. Flask, a widely used Python web framework renowned for its simplicity and flexibility, served as the foundation for constructing the API. With Flask's capabilities, the pre-trained model was seamlessly incorporated to facilitate its accessibility and utilization for predictive tasks.

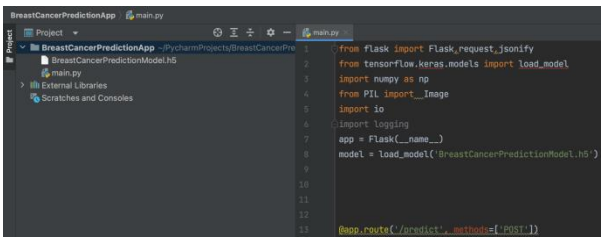


Figure 2. Code snippet and folder structures for the developed API

In the Figure 2 above, the developed model, referred to as "BreastCancerPredictionModel.h5," has been loaded and assigned as "model." The API utilizes this loaded model for conducting predictions.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 11 - 16, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1003

The loaded CNN model, was developed using Keras in Python, underwent training and testing using 1419 MRI images sourced from MNH. It comprises several layers:

i. An initial 2D convolutional layer with 16 filters (3x3 size, stride 1), employing the ReLU activation function. This layer transforms inputs of shape (256, 256, 3) to an output tensor of shape (254, 254, 16), featuring 448 parameters.
ii. Subsequent max pooling layers that down sample data.
iii. Following the initial layer, additional 2D convolutional layers (with 32 and 16 filters respectively, also 3x3 size, stride 1) were incorporated, each employing ReLU activation. These layers alter the tensor shapes and contain 4,640 and 4,624 parameters respectively.
iv. Fully connected layers, including one with 256 neurons and ReLU activation (with L2 regularization), amounting to 3,686,656 trainable parameters.
v. Dropout layers to prevent overfitting during training.
vi. Finally, a fully connected layer with 1 neuron using the sigmoid activation function, featuring 257 parameters.

The integration process involved configuring routes and endpoints within Flask to establish a structured interface for interacting with the pre-trained model. PyCharm, an integrated development environment (IDE) known for its robust features and support for Python, provided a conducive environment for coding, testing, and deploying the Flask-based API. This combination of Flask and PyCharm offered a conducive ecosystem for handling the integration intricacies, enabling efficient development and management of the API.

The Flask framework, being Python-based, ensured compatibility and ease of integration with the pre-trained model, streamlining the API development process. Leveraging Flask's capabilities within PyCharm, developers were empowered to create a functional and accessible API, encapsulating the pre-trained model's functionalities for consumption by various applications and systems. This approach aimed to provide a scalable, efficient, and user-friendly interface to access the predictive capabilities encapsulated within the pre-trained model.

3.3 Testing a Model API using Insomnia

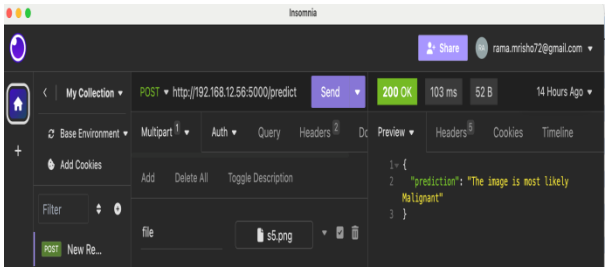


Figure 3 shows a result of an API end point after submit an MRI image file.

3.4 AI – Mobile App after integration

The integration process involved the development of a Flask API hosting a pre-trained breast cancer prediction model. Within the Flask framework, endpoints were set up to handle requests from a React Native Mobile App, acting as a bridge for data flow between the app and the predictive model. The React Native app was tailored to capture breast MRI images via the phone's camera and initiate HTTP requests to the Flask API for prediction processing. Upon receiving image data, the Flask API processed it through the pre-trained model and relayed the prediction results (benign or malignant) back to the React Native app. This integration aimed for a seamless connection between the frontend React Native Mobile App and the backend Flask API, ensuring accurate predictions for breast cancer using captured MRI images.

Following the integration, rigorous testing verified the functionalities, including image capturing, data transmission, model prediction accuracy, and user interface responsiveness. The React Native app's user-friendly design facilitated image capture and displayed prediction outcomes while maintaining a smooth user experience. With optimizations made to enhance performance and stability, the complete Mobile App for breast cancer predictions, combining the React Native frontend and Flask backend, was readied for deployment. This integration offered a reliable solution for users, enabling them to obtain accurate breast cancer predictions conveniently through their mobile devices, ensuring accessibility and efficacy in breast cancer diagnosis.

4. MOBILE APP PREDICTIONS

Once the application is initiated, radiologists can capture an MRI image by tapping the "Capture MRI Image" button. This action triggers the mobile camera, enabling users to take an MRI image. Refer to the figure below for visual guidance.

Figure 4 . A user interface for capturing an image

Next, the app will display the captured image, offering users(radiologists) the option to submit it for predictions. See Figure 5 below for a visual representation.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 11 - 16, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1003

Figure 5 . Shows the captured MRI image through the camera

Then, in order for an App to predict , radiologist must click the button Submit Image, then the results will be displayed on the screen as either the captured image is Malignant or Benign. See Figure 6 below.

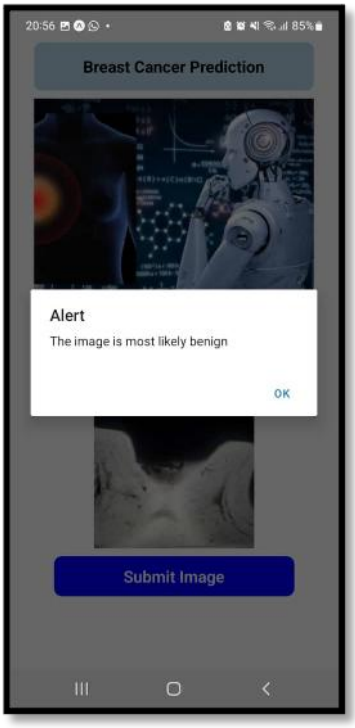


Figure 6. Shows a result of the captured image

Figure 6 depicted above displays the prediction output of the captured image as "Benign," as determined by the model's predictions.

Additionally, the model is also capable of predicting whether the captured image is Malignant. Refer to the figure 7 below for more details.

Figure 7. Another captured image for prediction

Also again, upon clicking the "**Submit Image**" button, the model will process the image and generate the output results. This process is demonstrated below.

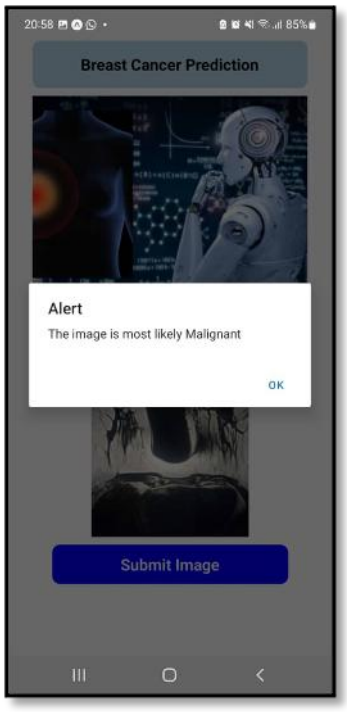


Figure 8. Results of a captured image

The figure depicted above displays the prediction output of the captured image as "Malignant" as determined by the model's predictions.

5. CONCLUSION AND FUTURE WORKS

The study focuses on the integration of a Flask API model with a mobile app, establishing an AI-driven platform for predicting MRI images in breast cancer detection. Using Python, the research aimed to craft a CNN breast cancer prediction model capable of categorizing MRI images as either benign or malignant. Employing 30 MRI images sourced from Muhimbili National Hospital and expanding the dataset to 1419 images through data augmentation, the developed CNN model comprised multiple layers and exhibited remarkable evaluation metrics, affirming its efficacy in accurately discerning breast cancer cases. This underscores the potential of CNNs and data augmentation in bolstering breast cancer detection, underscoring the necessity for expanded research with more comprehensive and diverse datasets.

The application of deep learning, particularly CNNs, holds significant promise in enhancing breast cancer detection in Tanzania as it will helps radiologist to improve the accuracy of diagnosis, offering a potential solution to the limitations of traditional detection methods. The paper explores the application of deep learning techniques, delves into the effectiveness of data augmentation, addresses the constraints of conventional methodologies, and proposes a neural network model aimed at augmenting breast cancer detection rates in Tanzania. Future research avenues include broadening the dataset to encompass images from diverse demographics, thereby enhancing the model's adaptability across various patient cohorts. Moreover, integrating the model into clinical settings and assessing its real-world advantages and constraints would provide invaluable insights. This study contributes to the breast cancer detection domain by presenting an efficient CNN-based approach and sets the stage for further advancements in precision and dependability.

6. REFERENCES

[1] Al-Haija, Q. A., & Adebanjo, A. (2020). Breast cancer diagnosis in histopathological images using ResNet-50 convolutional neural network. IEMTRONICS 2020 - International IOT, Electronics and Mechatronics Conference, Proceedings, 50. https://doi.org/10.1109/IEMTRONICS51293.2020.9216455

[2] Alanazi, S. A., Kamruzzaman, M. M., Islam Sarker, M. N., Alruwaili, M., Alhwaiti, Y., Alshammari, N., & Siddiqi, M. H. (2021). Boosting Breast Cancer Detection Using Convolutional Neural Network. Journal of Healthcare Engineering, 2021. https://doi.org/10.1155/2021/5528622

[3] Bakkouri, I., & Afdel, K. (2017). Breast tumor classification based on deep convolutional neural networks. In Proceedings - 3rd International Conference on Advanced Technologies for Signal and Image Processing, ATSIP 2017. https://doi.org/10.1109/ATSIP.2017.8075562

[4] Breast Cancer Initiative. (2017). Tanzania Breast Health Care Assessment 2017: An assessment of breast cancer early detection, diagnosis and treatment in Tanzania. Available Https://Ww5.Komen.Org/Breastcancertanzania, 1–62. https://ww5.komen.org/uploadedFiles/_Komen/Content/Grants_Central/International_Grants/Grantee_Resources/Full_Tanzania_Assessment_report.pdf

[5] Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, *106*, 249–259. https://doi.org/10.1016/j.neunet.2018.07.011

[6] Dani, H. (2022). Review on Frameworks Used for Deployment of Machine Learning Model. *International Journal for Research in Applied Science and Engineering Technology*, *10*(2), 211–215. https://doi.org/10.22214/ijraset.2022.40222

[7] Eroğlu, Y., Yildirim, M., & Çinar, A. (2021). Convolutional Neural Networks based classification of breast ultrasonography images by hybrid method with respect to benign, malignant, and normal using mRMR. *Computers in Biology and Medicine*, *133*(April). https://doi.org/10.1016/j.compbiomed.2021.104407

[8] Fonseca, P., Mendoza, J., Wainer, J., Ferrer, J., Pinto, J., Guerrero, J., & Castaneda, B. (2015). Automatic breast density classification using a convolutional neural network architecture search procedure. *Medical Imaging 2015: Computer-Aided Diagnosis*, *9414*, 941428. https://doi.org/10.1117/12.2081576

[9] Gonzalez, T. F. (2007). Handbook of approximation algorithms and metaheuristics. *Handbook of Approximation Algorithms and Metaheuristics*, 1–1432. https://doi.org/10.1201/9781420010749

[10] Kamiri, J., & Mariga, G. (2021). Research Methods in Machine Learning: A Content Analysis. *International Journal of Computer and Information Technology(2279-0764)*, *10*(2), 78–91. https://doi.org/10.24203/ijcit.v10i2.79

[11] Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436–444). Nature Publishing Group. https://doi.org/10.1038/nature14539.

[12] Lu, H. C., Loh, E. W., & Huang, S. C. (2019). The Classification of Mammogram Using Convolutional Neural Network with Specific Image Preprocessing for Breast Cancer Detection. *2019 2nd International Conference on Artificial Intelligence and Big Data, ICAIBD 2019*, 9–12. https://doi.org/10.1109/ICAIBD.2019.8837000

[13] Masud, M. (2020). Convolutional neural network-based models for diagnosis of breast cancer. *Neural Computing and Applications*, *5*. https://doi.org/10.1007/s00521-020-05394-5

[14] Ragab, D. A., Sharkas, M., Marshall, S., & Ren, J. (2019). Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ*, *2019*(1), 1–23. https://doi.org/10.7717/peerj.6201

[15] Yue, W., Wang, Z., Chen, H., Payne, A., & Liu, X. (2018). Machine learning with applications in breast

cancer diagnosis and prognosis. *Designs*, *2*(2), 1–17. https://doi.org/10.3390/designs2020013

[16] Yurttakal, A. H., Erbay, H., İkizceli, T., & Karaçavuş, S. (2020). Detection of breast cancer via deep convolution neural networks using MRI images. *Multimedia Tools and Applications*, *79*(21–22), 15555–15573. https://doi.org/10.1007/s11042-019-7479-6

[17] Zhao, X., Wang, X., & Wang, H. (2018). Classification of Benign and Malignant Breast Mass in Digital Mammograms with Convolutional Neural Networks. *ACM International Conference Proceeding Series*, 47–50. https://doi.org/10.1145/3285996.3286006

[18] Zuluaga-Gomez, J., Al Masry, Z., Benaggoune, K., Meraghni, S., & Zerhouni, N. (2021). A CNN-based methodology for breast cancer diagnosis using thermal images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, *9*(2), 131–145. https://doi.org/10.1080/21681163.2020.1824685

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 17 - 22, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1004

Building Comprehensive Dataset: Android File and Unstructured Data Collection from APKs for Enhanced Vulnerability Detection

Jigna Rathod
Assistant Professor
Babu Madhav Institute of Information Technology
UTU
Bardoli, India

Dr. Dharmendra Bhatti
Director and Professor
Asha M. Tarsadia Institute of Computer Science
Engineering, UTU
Bardoli, India

Abstract: Google's Android platform grew to become one of the world's largest mobile systems. An Android application can use any functionality that the platform provides. Malware may also be present in Android apps. That is why our objective in this research is to collect files from wide malware families and use them to help us and other researchers run vulnerability detection tests. From the most well-known Android malware projects, we gathered both benign and malicious files. This process yielded 20832 malicious files and 10856 benign files. In this study, we also disclosed the Android file collection step as well as our vulnerability identification technique. Machine learning approaches are frequently used to identify whether an APK file is tainted, with the goal of detecting malicious apps. We have been concentrating on machine learning approaches to discover the unknown vulnerability. To identify malware, the malware researcher must create his or her own dataset. As part of our dataset production process, we collect Android files, do dynamic analysis, and then extract their characteristics. We described the technique of producing data sets in this paper. The android file contains a lot of unformatted data in the form of text or XML files, which is difficult to analyze and store. Our objective in this context is to provide a dynamic analysis of these obtained Android files, allowing you to access the underlying information, such as system calls, network traffic, and permission requests made by specific apps. Using Dynamic Analysis, we are attempting to manage massive amounts of data and assure correct processing in the context of Android files. In this study, we offer MalwareDefender, a dynamic analytical tool that handles the challenging issues of evaluating and processing massive volumes of data.

Keywords: Android; Smartphones; Malware; Vulnerability Detection; APK Files, Dataset Generation, File Collection

1. INTRODUCTION

Nowadays, the process of detecting Android Vulnerability is based on a variety of methods, such as signatures, patterns, resources, and components that were statically analyzed to find known vulnerabilities [1][16], whereas machine learning techniques are capable of discovering previously unknown or recently discovered vulnerabilities [2]. Machine learning methods are increasingly being used to find vulnerabilities [1]. Machine learning approaches are proven to be less time demanding and resource costly than nonmachine learning methods [2].

An APK file can be detected as malicious by a researcher. To that purpose, a researcher must develop his or her standard set of methods. In this situation, a researcher would like to have access to the same dataset. The construction of a dataset is an important aspect of the procedure since it provides a list of features that will assist you in determining whether this APK file is malicious or not. It is also critical to choose the properties of the dataset since various malicious files exploit these aspects. This assists researchers in identifying a malicious file using exploited characteristics. Our dataset generation procedure consists of Android file collecting, dynamic analysis, and data extraction. During the dynamic analysis phase, we created MalwareDefender, an analytical tool that would examine APK files and collect raw data like as network traffic, system calls, and permission in real time. The chosen characteristics are derived from the unstructured data received during the data extraction phase via dynamic analysis.

As part of an Android File Collection phase, we collected 20832 malicious files from the most well-known Android malware projects, as well as 10856 benign files from the Google Play Store and other sites. In this article, we discussed android file collections, the dynamic analysis of files, and collecting unstructured data from APK files. All characteristics are extracted from files, and the data extraction step will result in the creation of our final dataset. The dataset will be utilized to train and test the models throughout the Machine Learning phase in order to research and deliver improved outcomes. Other articles will go over the various stages of feature extraction and machine learning.

Big Data comprises huge volumes of uncategorized data that can be stored in a number of formats such as text files, photos, audio, video, and so on [3]. Unstructured data has no explicit structure; instead, it consists of its internal structure [4]. The Android files are made up of text and XML files, each with its unique structure. There are 20832 malicious files and 10856 benign files in the Android file gathering phase. As a result, keeping and storing a huge amount of Android files is a big data problem, and executing dynamic analysis is a task in and of itself. In this part, we are attempting to manage a big quantity of data regarding Android files and ensure that they are appropriately handled by using dynamic analysis.

There is no dynamic analytic tool available to do this study. To that end, we've developed MalwareDefender, a sophisticated analysis tool that will be used by each application to perform tasks such as running it, keeping track of its current operation, and collecting raw data such as

system calls generated by an application, network traffic captured by applications, and permission requests from each functioning application. We discussed how our dynamic analysis tool works in our prior publication [5].

This document includes the sections listed below. Section 2 of our Generalized Detection System describes the full vulnerability detection procedure. Section 3 describes the android file gathering step. Section 4 describes the many difficulties we experienced while collecting data. Section 5 describes the limitations of the file gathering procedure. Section 6 describes the architecture flow of the dynamic analysis tool. Section 7 contains the paper's conclusion.

2. Overall Process of Vulnerability Detection

The overall process for detecting vulnerabilities is explained in this section. The architecture of vulnerability detection for android is explained in Figure 1. There are four main phases of this process.

- Android File Collection
- Monitoring Component (Dynamic Analysis)
- Data Extraction and Generation
- Machine Learning

The first and second phases of this paper are the collection of android files and the monitoring component. The next set of papers will deal with further phases. All android files are collected during the Android File Collection phase. Many popular android malware projects such as Drebin [6], AndroZoo [7], AndroPRAGuard [8] and MalDroid2020 [10] are used to collect these malicious files. Collection of benign files are available from the Canadian University of Cybersecurity [9] and Google Play, as well as ApkPure.com [11].

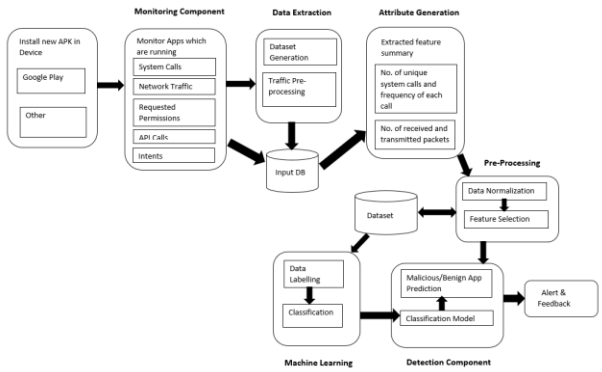


Figure 1 Proposed Architecture of Android Vulnerability Detection

The APK file format is an Android Package file type that is used to launch programs on Android-powered devices. The APK file includes the Java code required to run the program, as well as AndroidManifest.xml, resources, assets, and other files. Java code is compiled into bytecodes, which are subsequently stored as Dalvik executable files (.dex files). The Android operating system uses the Dalvik Virtual Machine (DVM) to run .dex files.

For the Monitoring component, we created an Android app named MalwareDefender. MalwareDefender is an Android software that monitors and records the state and performance of the system. It is written in Java and uses the Monkey Runner tool to perform 5000 random events each app launch. The Strace and tcpdump utilities have been introduced to gather system call logs and monitor network traffic. We obtained a total of 10066 sets of original data by doing dynamic analysis on all active apps. A component responsible for attribute creation will gather data from permissions, system calls, and network traffic. Selected characteristics are used as input in the data extraction and creation process. A dataset of chosen characteristics is built from the unstructured data acquired through dynamic analysis. The final dataset includes key aspects such as permissions, system calls, and network traffic. This dataset is utilized in the machine learning phase to categorize malware and benign programs, where different supervised algorithms, feature reduction and extraction approaches, and ensembling techniques are used. We will train a dataset using machine learning, and then apply all algorithms to the tests based on the taught data to produce more precise findings for vulnerability detection. This will help us understand the vulnerability detection system better.

3. Android Files Collection

The overall collection process for the android files is explained in this section. The first step in preparation of the dataset is to collect android files, which is explained here.

3.1 File Collection Procedure

This section describes the architectural flow of the Android file gathering process. The user opens the Android File collection module, inputs a URL, and requests a malicious or benign file download. This module connects to a website via a specified URL, and if the connection is successful, the website connects to the server and sends the user-supplied file request for download. The server responds with the website's file path, and lastly, the malicious or benign file is downloaded and saved to a physical place. Malware or benign files are selected based on global Android Malware databases and website data. In our article, we evaluated and detailed some well-known datasets of Android malware as well as a few benign datasets. Other researchers quote, use, and refer to the majority of those data sets. Furthermore, because these databases include a wide spectrum of malware families, our study is extremely dependable, implying that the results we collected are as well. The Android malware dataset may be acquired from the Drebin dataset, which is a well-known initiative [6]. The Androzoo dataset has a vast number of applications that are constantly growing [7] [12]. AndroPRAGuard's database contains 10479 applications from various families, as well as examples from the MalGenome and ContagioMinidump datasets [8]. The old samples may be found in Drebin, AndroZoo, and AndroPRAguard. So, for recent samples, we gathered samples from the MalDroid2020 [10] dataset. The preceding procedure resulted in the collection of roughly 20832 Malware programs. Drebin has been used to download about 5490 programs [6]. AndroZoo has around 4000 programs downloaded [7]. AndroPRAGuard has been used to download 5953 apps [8]. From MalDroid2020 [10] around 5389 applications are downloaded. In view of the variety of malware families included in this dataset, as well as a number of recently discovered malicious files, we chose these popular Android Malware Datasets to run our process.

Benign applications were acquired from databases, the Google Play Store, and other places. We collected around 6500 benign applications from the Canadian University of

Cyber Security dataset [9]. The Google Play store has received 356 app downloads. We gathered around 4000 benign applications from websites [11]. We choose the official Google Play Store, various well-known datasets, and well-known internet domains for benign file download. Table 1 summarizes the total number of Android applications downloaded from each market place.

Table 1 Android Files Summary

Popular Datasets	No. of files download	File Type	Total Files
Drebin	5490	Malicious	20832
Androzoo	4000		
AndroPRAGuard	5953		
MalDroid2020	5389		
Google Play Store	356	Benign	10856
Canadian University of Cyber Security	6500		
Websites	4000		

4. Challenges in Data Collection

During the collection of APK files from various android marketplaces, we have encountered a number of challenges and obstacles. Some general issues include the following:

1. **Request to grant access to the Datasets:** The data for Android Malware and Benign APK are easily accessible, however they cannot be downloaded directly. To be authorized for this reason, we must request a dataset by sending an email, and researchers of the datasets will enable us to obtain them in exchange for API keys, passwords, or downloading URLs. Requests are used to retrieve datasets from Drebin [6], AndroZoo [7], AndroPRAGuard [8], MalDroid2020 [10], and The Canadian University of Cyber Security [9]. These datasets, which include many different types of malwares and a large number of files, are being used by researchers all around the world in their study.
2. **Manual Download of Files:** When we obtained the researcher's API key to download the files, we were not permitted to download the bulk files at the same time in Androzoo [7] [12]. Androzoo [7] gives an excel file with all of the infected files' details, including SHA256, Package Name, File Size, Date, and Time. The SHA256 value is unique to each file. Androzoo [7] allows you the freedom to download any malicious file you choose. You may also download new malicious files as well as from the past. A collection of malware files can be created based on one's preferences and needs. To download each file manually, we have followed the following steps.

    1. File Download link:

       https://androzoo.uni.lu/api/download?apikey=${APIKEY}&sha256=${SHA256}

    2. Use the API key acquired from AndroZoo [7][12].

    3. Select the SHA256 value from the Excel file and replace it in the link.

    4. Paste the entire URL into the browser.

       For instance, https://androzoo.uni.lu/api/download?apikey=f60e30400dbecb3eda905972548fb8b971ef07710bf96b8d358567d01a741ce8&sha256=0000003B455A6C7AF837EF90F2EAFFD856E3B5CF49F5E271914

    5. The API Key will remain the same, but the SHA256 value for each APK file will be different. So, for each APK, change the SHA256 value and copy-paste the URL to download it.

5. Limitations of File Collection Process

The total amount of Malware Files acquired by this method imposes some limitations also. They are as follows

1. It is not possible for a researcher to identify a Malware Family from a single Malware file. We do not disclose any more information on the Malware Files.
2. The Malware Files we have collected span the years 2011 to 2021.
3. We have only gathered files that are openly available.

6. Dynamic Analysis of Android Files

The different challenges to unstructured data extraction and the dynamic analysis phase are discussed in this section.

6.1 Challenges of Unstructured Data Mining

In the realm of big data, there is a massive volume of unstructured and heterogeneous data to deal with [13-15]. This is because unstructured data does not have specified schema or models and has its own internal structure, rendering traditional RDBMS unsuitable for storing it [3][4]. Big data is concerned with three Vs: 1) Volume, 2) Velocity, and 3) Variety [13][15].

- Volume: This is a large number of datasets, which are complex in their data structure. The challenge with volume is to handle the complexity of the data structure [13][15].
- Velocity: The need to handle the pace with which new datasets are created or old datasets are updated is referred to as velocity. This aspect pertains to data generated by machines, such as mobile device sensors. The issue with velocity is coping with the streaming system's limited capacity while obtaining valuable information from continual fresh dataset production [13] [15].
- Variety: Datasets come from a range of sources and can be in a variety of formats, including text, audio, video, graphs, sensors, and so on. [13][15]. The variety of data provides more information for

problem solutions. The issue with diversity is merging numerous systems to accept different sorts of data [15].

The framework for large data analysis and mining is depicted in Figure 2. Tier 1 data mining procedures for large datasets necessitate the use of expensive computer units and clusters for data processing and comparison [13]. To achieve high speed, big data processors rely on cluster computers, and data mining activities are done using parallel programming tools such as Map-Reduce [13]. Tier 2 focuses on data security and subject expertise. Data privacy may be achieved by restricting data access such that sensitive data is only available to particular persons. The second way to preserve data privacy is to anonymize data fields so that sensitive information is not connected with a specific record [13]. Tier 3 is focused with algorithms for massive data mining. Deep analysis, which is quite difficult, is necessary for the creation of the algorithms [15]. Mining algorithms are self-contained and work decentralized [13]. Big data mining necessitates the use of machine learning and data mining technologies that need a large amount of computational power and resources [13]. As a result, in big data analysis and mining, each layer covers all of the complicated issues.

Tier 1: Data Accessing and Computing
Tier 2: Data Privacy and Domain Knowledge
Tier 3: Big Data Mining Algorithms

Figure 2 The Framework for Big Data Analysis and Mining

To solve all of the challenges connected with unstructured data analysis and mining, we proposed a dynamic analysis tool (MalwareDefender) that does dynamic file analysis, creates unstructured data in.xls,.csv, and .pcap files, and saves the files on the device. These data are then sent to a local system with enough capacity for further processing. It solves all of the issues involved with dealing with data volume, velocity, and variety.

6.2 Dynamic Analysis Phase

MalwareDefender in Java is a program that does a dynamic examination of all APKs that are now executing. It looks for features like permission, system calls, and network traffic in the present program. These attributes are saved in a variety of file formats, including .xls,.csv, and .pcap, and Wireshark is then used to analyze the .pcap files to extract additional network traffic information. This allows for the effective extraction of certain features from Android apps.

Figure 3 depicts the architectural flow of dynamic analysis tool and its operation, which entails the sequential implementation of the following steps:

- The release of MalwareDefender, an Android application built by us, will signal the beginning of the data extraction approach based on dynamic analysis.
- When MalwareDefender is launched, it will provide a complete list of all programs presently installed on the device.
- Run the required application to extract features from the one that is already executing.
- If the app is correctly launched, Monkey Runner will generate 5000 random actions; else, a list of installed applications will be displayed for the user to pick and run a new one.
- During a 20-minute random event produced by Monkey Runner, MalwareDefender will monitor and collect system calls, app permissions, and network activity.
- The app should be terminated and deleted from the background after 20 minutes.
- After exiting the current app, MalwareDefender saves permission, system calls, and network traffic data to the device in.xls, .cls, and .pcap formats, respectively.
- If you have access to all three files on your smartphone, save them to your local system for additional investigation. If this is not the case, restart the application to extract the data.
- The data extraction app will go through each application and extract all of its attributes.

Once all this unstructured data is gathered, data extraction and mining will be carried out. The data extraction and mining phase will be discussed in another paper.

7. CONCLUSION

To determine if an APK file is malware or not, a researcher needs to first create a dataset. We discussed the dataset production process, which included Android File Collection, Monitoring Component (Dynamic Analysis), Data Extraction and generation, and the Machine Learning Phase. We investigated the Android File collection phase in this paper and presented a dynamic analysis tool (MalwareDefender) that we designed to do dynamic analysis of running applications. During the Android File Collection phase, we collected 20832 malicious files and 10856 benign ones. We also discussed some of the issues faced during data collection as well as some of the data collection limits.

We finished the collecting of Malware and Benign files for dataset construction during the Android File gathering Phase. We also reviewed the architectural flow of our recommended dynamic analysis tool and illustrated the whole Android File Collection procedure. Mining

characteristics from APK files and generating the dataset needs dynamic file analysis. We performed a dynamic analysis on all 10066 APK files, implying that our final dataset would contain 10066 records that machine learning algorithms will use to train the model. The next papers will go into the stages of data extraction and generation, as well as machine learning.

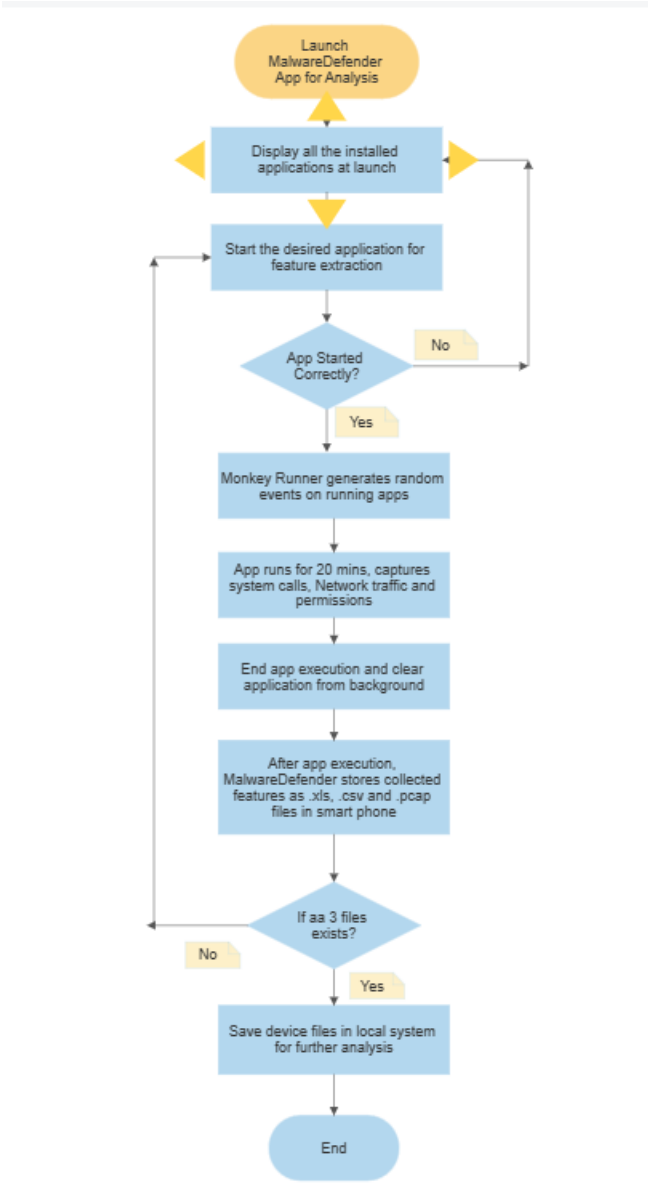


Figure 3 MalwareDefender – Architectural Flow of Dynamic Analysis Tool

8. ACKNOWLEDGMENTS

We would like to acknowledge our two students Mr. Bhargav Patel and Mr. Parth Desai of BMIIT, UKA Tarsadia University who helped us in manually downloading the files from AndroZoo [7] and performing dynamic analysis of each APK file using our implemented tool. They downloaded around 4000 files from AndroZoo [7] and perform dynamic analysis of around 1000 APKs dedicating their time for this task.

9. REFERENCES

[1] Agrawal, P., & Trivedi, B. (2019, February). A survey on android malware and their detection techniques. In 2019 IEEE International conference on electrical, computer and communication technologies (ICECCT) (pp. 1-6). IEEE.

[2] Alqahtani, E. J., Zagrouba, R., & Almuhaideb, A. (2019, June). A survey on android malware detection techniques using machine learning algorithms. In 2019 Sixth International Conference on Software Defined Systems (SDS) (pp. 110-117). IEEE.

[3] Kanimozhi, K. V., & Venkatesan, M. (2015). Unstructured data analysis-a survey. International Journal of Advanced Research in Computer and Communication Engineering, 4(3), 223-225.

[4] Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile networks and applications, 19, 171-209.

[5] Rathod, J., & Bhatti, D. (2022, March). Vulnerability detection in the android application based on dynamic analysis. In Proceedings of Third International Conference on Intelligent Computing, Information and Control Systems: ICICCS 2021 (pp. 287-302). Singapore: Springer Nature Singapore.

[6] Arp, D., Spreitzenbarth, M., Hubner, M., Gascon, H., Rieck, K., & Siemens, C. E. R. T. (2014, February). Drebin: Effective and explainable detection of android malware in your pocket. In Ndss (Vol. 14, pp. 23-26). Brown, L. D., Hua, H., and Gao, C. 2003. A widget framework for augmented interaction in SCAPE.

[7] University of Luxembourg. (n.d.). Androzoo home. Retrieved April 9, 2019, from https://androzoo.uni.lu/

[8] Android PRAGuard Dataset | PRA Lab. (n.d.). Retrieved May 19, 2019, from http://pralab.diee.unica.it/en/AndroidPRAGuardDataset

[9] CICMalAnal 2020 | Datasets | Research | Canadian Institute for Cybersecurity | UNB. (n.d.). Retrieved July 15, 2020, from MalDroid 2020 | Datasets | Research | Canadian Institute for Cybersecurity | UNB

[10] Retrieved July 15, 2020, from MalDroid 2020 | Datasets | Research | Canadian Institute for Cybersecurity | UNB

[11] Android Files Download. (n.d.). Retrieved Feb 19, 2019, from https://apkpure.com

[12] Allix, K., Bissyandé, T. F., Klein, J., & Le Traon, Y. (2016, May). Androzoo: Collecting millions of android apps for the research community. In Proceedings of the 13th international conference on mining software repositories (pp. 468-471).

[13] Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2013). Data mining with big data. IEEE transactions on knowledge and data engineering, 26(1), 97-107.

[14] Chen, J., Chen, Y., Du, X., Li, C., Lu, J., Zhao, S., & Zhou, X. (2013). Big data challenge: a data management perspective. Frontiers of computer Science, 7, 157-164.

[15] Fan, W., & Bifet, A. (2013). Mining big data: current status, and forecast to the future. ACM SIGKDD explorations newsletter, 14(2), 1-5.

[16] Solanky, M. J., & Bhatti, D. (2019). A review on several vulnerabilities' detection techniques in Android Mobile.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 23 – 25, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1005

IoT based Early Warning System for 10 KVA UPS

Hadiyanto [1*]
[1] Electrical Engineering,
Balikpapan State Polytechnic,
Balikpapan, 76129, Indonesia

Syahruddin [2]
[2] Electrical Engineering,
Balikpapan State Polytechnic,
Balikpapan, 76129, Indonesia

Ali Abrar [3]
[3] Electrical Engineering,
Balikpapan State Polytechnic,
Balikpapan, 76129, Indonesia

Abstract: Uninterruptible Power Supply (UPS) has an important role for a system that requires a backup voltage source to keep the system running even if the main voltage source is interrupted. When the main voltage source of a system is disturbed, the UPS will switch from the main voltage source to the battery owned by the UPS without disconnecting the source to the system so that the system continues to operate as if it were getting a main voltage source. When the main power source is off, the UPS cannot be a source of power for the system forever because the UPS has certain limitations such as battery capacity. In this case, it happened at the Balikpapan State Polytechnic data center, the server died because the battery was not charged by the generator when the power source from the State Electricity Company (PLN) went out. To anticipate UPS outages in the Balikpapan State Polytechnic data center, this monitoring tool needs to be built so that it can monitor the capacity level of battery usage, by designing early warnings in the form of several indicators such as pilot lights, alarm bells and mini sirens. For a system that has a large infrastructure, the death of one of the large infrastructure will cause disturbances and can even completely shut down the system. Therefore, UPS can only back up the system for a certain time or condition. In fact, it is very difficult to monitor the UPS manually. The experimental results obtained data that the values on the UPS can be viewed and monitored online and it was found that the value of the voltage sourced from PLN was 230 VAC while the value of the voltage sourced from the battery was 12 VDC.

Keywords: Uninterruptible Power Supply (UPS), Arduino Uno, PZEM-004T, Relay Module, pilot lamp alarm Buzzer, miniature sirine

1. INTRODUCTION

In today's modern world, plenty equipment relied heavily on electrical energy. To secure these critical devices from potential drop off of the power supply which typically cause unpleasant effect on responsible users, an UPS system is installed. At large scale infrastructure, the UPS system has been running as a means for power-back up, ensuring all critical apparatus remain operational during power disruption [1]). UPS application has also been useful in supporting data centre by minimizing the risk of significant data loss when power source is interrupted [2]. Nevertheless, the UPS system contains shortcomings in terms of its fixed amount of power generated and its limited battery capacity [3]. In other words, the UPS can only function over a certain amount of time and condition. Therefore, to prevent complete failure of the system, the UPS should be properly monitored.

Full control of the UPS system on-site manually seemed inconvenient and costly. To resolve this, internet technology might be integrated, allowing timely monitoring and predicting. Parameters output of the UPS could be viewed at real time, such that in light of a sudden decline of voltage or current from an ideal condition, early warning notification would be informed to in-charged person via mobile network.

Internet of Things (IoT) concept for surveillance and forecasting has been widespread. Supplemented by sensors and actuators, technology become more instance in various areas including smart home [4], agriculture [5], intelligent transportation and smart cities [6,7,8]. Previous studies on UPS monitoring built-on IoT concept had reported the utilization of MQTT protocols [9], GSM900 module [2], and wimos component, thingspeak, and virtuino application [10]. Their overall results had been compelling.

Therefore, motivation of this study is to develop a 10kVA UPS surveillance and early warning system based on IoT technology. This prototype design has been implemented on Campus infrastructure of Balikpapan State Polytechnic. In this prototype, The UPS related outputs/parameters encompassing: voltage, current, the battery capacity, discharge rate/ the remaining power/ the magnitude of utilised power had be monitored at real time using PZEM-004T as sensor node (digital meter), connected to Arduino uno microcontroller, and Blynk application at user side. Results of this project would add to the body of knowledge towards IoT field by providing improved energy management within infrastructure telecommunication in Indonesia

2. RESEARCH METHODS

2.1. System design

a. Node MCU-ESP32

Node MCU is a compact development board, designed for IoT project. It contains a built-in USB port, hardware reset button, wi-fi antenna, and standard sized GPIO pin. ESP-32 is a latter derivative of ESP8266, manufactured by Esspresif, having superior qualities as compared to other microcontroller, including: more pins, higher RAM, larger memory, and augmented by low energy 4.0 bluetooth.

b. Arduino IDE

Arduino IDE is a software to develop a coding program which would be uploaded to Arduino board. Its main features are: Code Editor (that help developer to write and edit the sketch programs), Library dan Board Manager, Serial Monitor (that allow developer to receive or send data for debugging/ monitoring), Compilation and Upload.

c. Sensor PZEM-004T

PZEM-004T is used in this project in conjunction with microcontroller ESP32 and Arduino, permitting developers to build smart power monitoring system. PZEM-004T as sensor node (digital meter) provides electrical parameters such as voltage and current measurements, frequency, active power, energy consumption, and power factor. (Some versions can

also detect Voltage Sag &Surge). A serial communication must be established to connect PZEM-004T to Arduino. Hardware connection of PZEM-004T to Arduino is shown in Table 1.

Table 1. Connection PZEM-004t to Arduino

PZEM-004T V3.0	Arduino UNO / Mega2560
VCC	+5V
GND	GND
TX	RX (Software Serial / Harware Serial)
RX	TX (Software Serial / Harware Serial)

d. Blynk

Blynk is a platform for IOS and Android used to control Arduino module via the internet. Blynk has three main components: app, server, and libraries, such that it handles communication between smartphone and hardware, and also function as data storage-visualization.

2.2. System workflow

Step 1. Designing Hardware lay-out: setting up the lay out for all modules and electric components. Arduino Mega as the centre for data monitoring process.

Step 2. Developing software programming: utilizing Ardunio. This is executed after simulation on visual studio case.

Step 3. Software-hardware integration: data protocol creation on Arduino, followed by interfacing it into the module RS 485. Reading information on the hardware model, as indicated by a series of alarm detectors notification. (Interfacing software data into hardware apparatus)

Step 4. Producing Prototype output: a complete try out of the system which include sensors detection by Arduino, real time monitoring of power supply condition, detection of power supply problem triggering early warning system, transmitting information to mobile phone via Internet. (analysing & monitoring system try-out)

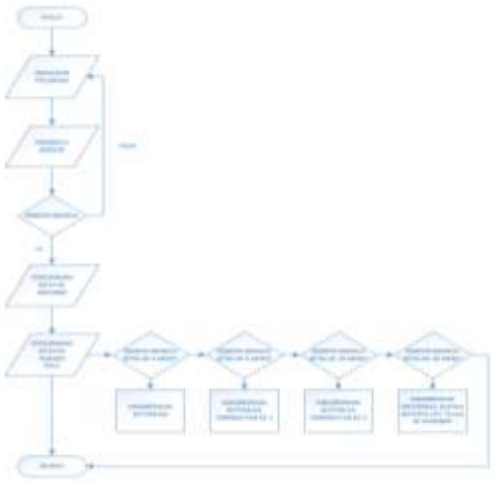


Figure 1. Flow chart of the prototype early warning system

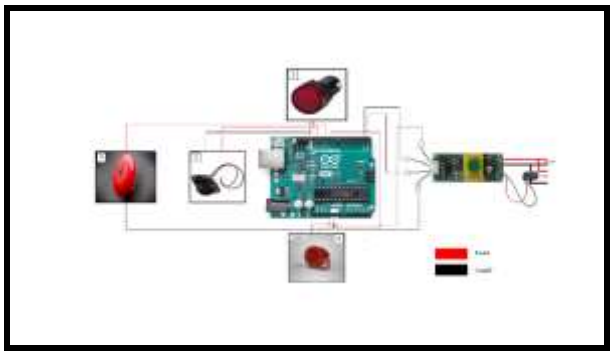


Figure 2. System Block Diagram

Figure 2. Showed the design of Arduino connection to pzem 004-t, and detectors viz: Pilot Lamp (1), Buzzer Alarm (2), Fire detector (3), and 12 volt mini Serien (4). A series of notification would be applied. First, interuption/disconnetion of power supply (either from PLN or getset) for 5 second would switch the Pilot Lamp ON. Then, when problem/disconnection continue for another 10 second triggered buzzer alarm ON. Disconnetion for another 15 second turned the fire detector ON. Lastly, when power interuption continued up to 20 second, mini sirine X dB would be ON.

3. RESULTS AND DISCUSSION

In this implementation, the entire component has been installed so that it becomes a monitoring tool. The implementation results show that the tool functions well and normally with the pzem-004t sensor indicator able to read the interrupted voltage. Furthermore, the Arduino Uno can receive output from the pzem-004t sensor and the relay module works well by receiving the program from the Arduino and turning on several indicators such as the pilot light, alarm bell and mini siren. In this implementation, Arduino is a microcontroller system designed to monitor, detect and provide early warning information when the Uninterruptible Power Supply (UPS) is about to go out. The pzem-004t sensor is a tool that aims to answer existing data needs in the form of information that can be read when PLN and the generator are not connected to the Uninterruptible Power Supply (UPS). The notification process when the power goes out and the UPS is working will be carried out in stages. If the power goes out and the Uninterruptible Power Supply (UPS) does not receive power from the generator for 5 seconds, the pilot light indicator will come on. Furthermore, if the PLN electricity still goes out and the Uninterruptible Power Supply (UPS) does not receive power from the generator for 10 seconds then the second notification will work in the form of a red pilot light. If the second notification is received, the PLN and generator are still not connected for approximately 15 seconds. then the third notification will appear with a sound indicator using a fire detector. If several attempts have been made to inform that the generator and PLN are not connected to the UPS within 20 seconds, then the 4 indicators will light up, informing you that the situation is

very emergency and a technician is needed to go to the location of the equipment.

The following is a table of system testing results with 4 early warning indicators:

Figure 3. Working indicator results

To increase the effectiveness of the early warning process, a notification system with blynk can be installed so that cellphone messages can be given that the UPS is not getting electricity supply from PLN and generators. The notification process is adjusted to the time of the early warning monitor indicator, namely 5 seconds, 10 seconds, 15 seconds and 20 seconds. The following displays the notification results on the Blynk display.

Figure 4. Notification When UPS Is Using Battery

4. CONCLUSION

From the test results of the unloaded PZEM-004T sensor, the voltage accuracy was 231 volts. Meanwhile, if the load is given to a server with 7000 watts of power, the PZEM-004T sensor remains with a voltage of 231 volts and a current of up to 31.8 A.

From the test results of the PZEM-004T sensor, accuracy was obtained with an error percentage below 1%. With the resulting data it can be concluded that the sensor used is quite accurate and precise

5. REFERENCES

[1] P. G. Krishna, "UPS Parameter Monitoring and Controlling Using IOT and GSM," Int. J. Pure Appl. Math., vol. 116, no. 6, pp. 133–139, 2017.

[2] Manasa KV, Prabu AV, Prathyusha MS, Varakumari S. (2018) Performance monitoring of UPS battery using IoT. International Journal of Engineering & Technology. 7 (2.7) (2018) 352-355.

[3] L. Fu and B. Zhang, "The Design and implementation of a UPS Monitor and Control System.," Int. Conf. Comput. Sci. Netw. Technol., vol. 2, pp. 1322–1325, 2011.

[4] T. S. Gunawan, I. Rahmithul, H. Yaldi, M. Kartiwi, and N. Ismail, "Prototype Design of Smart Home System using Internet of Things," Indones. J. Electr. Eng. Comput. Sci., vol. 7, no. 1, pp. 107–115, 2017.

[5] Mohanty, A., Roy, S. (2021). Smart Plant Monitoring System Using BLYNK Application. In: Banerjee, S., Mandal, J.K. (eds) Advances in Smart Communication Technology and Information Processing. Lecture Notes in Networks and Systems, vol 165. Springer, Singapore. https://doi.org/10.1007/978-981-15-9433-5_19

[6] S. Chen, H. Xu, D. Liu, B. Hu, and H. Wang, "A vision of IoT: Applications, challenges, and opportunities with China Perspective," IEEE Internet Things J., vol. 1, no. 4, pp. 349–359, 2014.

[7] N. A. Hidayatullah and D. E. J. Sudirman, "Desain dan aplikasi internet of thing (iot) untuk smart grid power system," 2017.

[8] N. Fajrin, I. Taufik, N. Ismail, L. Kamelia, and M. A. Ramdhani, "On the Design of Watering and Lighting Control Systems for Chrysanthemum Cultivation in Greenhouse Based on Internet of Things," IOP Conf. Ser. Mater. Sci. Eng., vol. 288, no. 1, 2018.

[9] Alqinsi P, Matheus Edward IJ, Ismail N, and Dharmalaksana W. "IoT-Based UPS Monitoring System Using MQTT Protocol". *2018 4th International Conference on Wireless and Telematics (ICWT)*, Nusa Dua, Bali, Indonesia, 2018. Pp. 1-5, doi: 10.1109/ICWT.2018.8527815.

[10] Rigoursyah et al, 2020 : Muhammad Agung Foury Rigours dan Ekki Kurniawan, S, "Perancangan Ups Berbasis Sumber Energi Listrik Terbarukan Dan PLN Termonitor Perangkat Iot," *eProceedings of Engineering*, vol. 3, p. 1, 2020.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 26 - 32, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1006

Comparison of Faster R-CNN ResNet-50 and ResNet-101 Methods for Recycling Waste Detection

Puteri Nurul Ma'rifah
Depart. of Electrical Engineering
State Polytechnic of Malang
Malang, Indonesia

Moechammad Sarosa
Depart. of Electrical Engineering
State Polytechnic of Malang
Malang, Indonesia

Erfan Rohadi
Depart. of Information Technology
State Polytechnic of Malang
Malang, Indonesia

Abstract: Inefficiencies in waste management contribute to the increasing amount of pollution in society, leading to public demands for better waste management and classification. Waste sorting is the beginning of the waste recycling process, which can help reduce the amount of waste in the environment. However, coupled with a lack of awareness of waste sorting due to minimal public education about waste management, the waste sorting system is still carried out manually using human power. Therefore, it is necessary to have a waste classification system to encourage people to manage their waste well. This research aims to design a tool to detect types of waste and classify them into three categories; metal, paper and plastic waste. The system can recognize the shape of trash images using a deep learning method developed using Faster R-CNN with ResNet-50 and ResNet-101 as the network architecture. This research began by collecting 250 datasets of metal, paper and plastic waste which were used as training data and test data in the testing process. The training data for the training process are 80, 120 and 200 datasets respectively. Test data for each training experiment, 30 and 50. Where 20 datasets in 50 test data are taken from the dataset for the training process. In each training process, the number of steps is carried out up to 3000, 4000 and 5000 steps, each of which has a total loss parameter. Based on the test results applied to the Faster R-CNN ResNet-50 and ResNet-101 methods, it produces an average F1 Score of 63% and 77% respectively. The best F1 Score is Faster R-CNN ResNet-101.

Keywords: Deep Learning, Object Detection, Waste Classification, Recycled Waste, Resnet50, Resnet10

1. INTRODUCTION

Waste is leftover goods/materials that are no longer used. Waste is produced from several processes, namely from human activities or natural ecosystems. There are many classifications of waste, one of which is classified into organic and non-organic waste. Based on research, Indonesia is the second largest waste producer after China. President Jokowi revealed that the floods experienced at the beginning of the new year 2020 were caused by damage to the ecosystem and ecology and because many people were still throwing rubbish carelessly. The Ministry of Environment and Forestry (KLHK), also said that the amount of waste generated nationally is 175,000 tons per day. It's pretty clear that waste is a big problem in Indonesia **[1]**

Object detection is concerned with detecting visual objects (such as people, animals, or cars) in digital images [2]. Object detection aims to detect target objects with theories and methods of image processing and pattern recognition, determine the semantic category of objects, and mark the specific position of target objects in the image [3].

Deep learning is part of Artificial Intelligence (AI) [4]. Deep learning allows computational models consisting of several processing layers to learn data representations with various levels of abstraction [5].

Convolutional neural networks (CNN) is one of the most powerful deep learning algorithms that has many applications in image classification, segmentation, and detection. [6]. R-CNN is a convolution-based algorithm or Region-based CNN. R-CNN is a combination of Region Proposal Network (RPN) and CNN proposed by Girshick et al in 2015.

2. METHOD

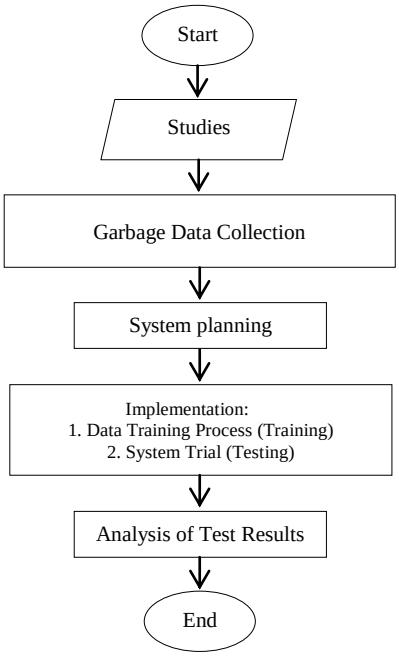


Figure 1. Research Stages

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 26 - 32, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1006

Literature studies are intended to study various reference sources such as journals, articles, papers on related research topics and themes. The topics or themes studied in this research are CNN, ResNet -50, and ResNet -101. Dataset is a collection of data needed in this research. The data needed in this research is plastic, paper and metal waste.

At the dataset collection stage, the dataset collected is images of plastic, paper and metal waste. Dataset collection was carried out by conducting independent data collection. This research carried out several stages of dataset training. The dataset used is 200 images. The data used for the training process includes 80 images for the first process, 120 images for the second process and 200 images for the third process. In each data training or training process is carried out using the ResNet-50 and ResNet-101 architectural models. Training is carried out up to step 5000. This step is carried out to test the algorithm of the system being designed. In the testing stage, the dataset used for test data is 30 images and 50 images for each training process. The 20 images in the test data of 50 images are obtained from the trained images. This step was taken to determine the reliability of the system in detecting recyclable waste, namely plastic, paper and metal waste. The steps taken are by entering test data into the system, then the system will detect and classify the waste contained in the image.

3. RESULT AND DISCUSSION

Computation time is the time required for the calculation process carried out by a computer when solving a problem using the algorithm used. From the experiments that have been carried out, there is computing time or time required during the training process and testing process for each image on the Faster R-CNN ResNet-50 and ResNet-101 network architectures..

3.1 Computation Time for Data Training Process (Data Training)

The computing time required during the training process for each architecture is shown in Table 1.

Table 1. Training Process Computation Time

Network Architecture	Amount of Training Data	Training Computation Time			Average
		3000	4000	5000	
Faster R-CNN ResNet 50	80	15m 24s	20m 26s	25m 19s	20m 23s
	120	15m 15s	20m 41s	25m 03s	20m 26s
	200	15m 40s	20m 39s	26m 45s	21m 01s
	Average				20m 37s
Faster R-CNN ResNet 101	80	18m 52s	25m 18s	31m 55s	25m 22s
	120	19m 39s	26m 20s	32m 42s	26m 14s
	200	19m 43s	26m 50s	32m 49s	26m 27s
	Average				26m 01s

Based on Table 1, it can be seen that the longest time required for the training process is the Faster R-CNN ResNet-101 architecture with an average time reaching 26 minutes 01 seconds, while the fastest is the Faster R-CNN ResNet network architecture. -50 with an average time of 20 minutes 37 seconds.

3.2 Computation Time Data Testing Process (Data Testing)

The computing time required during the testing process for each architecture is shown in Table 2.

Table 1. Test Process Computation Time

Network Architecture	Trial	Test Computation Time			Average
		3000	4000	5000	
Faster R-CNN ResNet 50	1	01m 47s	01m 43s	01m 41s	01m 44s
	2	01m 41s	01m 43s	01m 41s	01m 42s
	3	01m 15s	01m 43s	01m 16s	01m 25s
	4	01m 12s	01m 40s	01m 15s	01m 22s
	5	01m 18s	01m 18s	01m 44s	01m 27s
	6	01m 20s	01m 13s	01m 13s	01m 15s
	Average				01m 29s
Faster R-CNN ResNet 101	1	01m 42s	01m 18s	01m 22s	01m 27s
	2	01m 42s	01m 42s	01m 39d	01m 41s
	3	01m 43s	01m 45s	01m 22s	01m 37s
	4	01m 20s	01m 41s	01m 21s	01m 27s
	5	01m 44s	01m 22s	01m 20s	01m 29s
	6	01m 41s	01m 40s	01m 42s	01m 41s
	Average				01m 34s

Based on Table 2, it can be seen that the longest time required for the testing process is the Faster R-CNN ResNet-101 architecture with an average time reaching 01 minutes 34 seconds, while the fastest is the Faster R-CNN ResNet network architecture. -50 with an average time of 01 minutes 29 seconds. ResNet-101 consists of 101 initial layers so it has a longer computing time than ResNet-50 which consists of 50 layers [7].

3.3 Recyclable Waste Detection Results

Recyclable waste detection results based on each network architecture model are shown in Figure 2 to Figure 9, In this table there are several images from the detection of recyclable waste based on the Faster R-CNN ResNet-50 and Faster R-CNN ResNet-101 architectural models.

3.3.1 Faster RCNN ResNet-50 Detection Results

The results of detecting recyclable waste on the Faster R-CNN ResNet-50 network architecture can be seen in the image below:

Figure 2. Faster RCNN ResNet-50 detection results for metal can waste

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 26 - 32, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1006

Figure 3. Faster RCNN ResNet-50 detection results for paper waste

Figure 4. Faster RCNN ResNet-50 Detection Results of plastic bottle waste

Figure 5. Faster RCNN ResNet-50 Detection Results for recyclable waste classification

3.2.1 Faster RCNN ResNet-101 Detection Results

The results of detecting recyclable waste on the Faster R-CNN ResNet-101 network architecture can be seen in the image below:

Figure 6. Faster RCNN ResNet-101 detection results for metal can waste

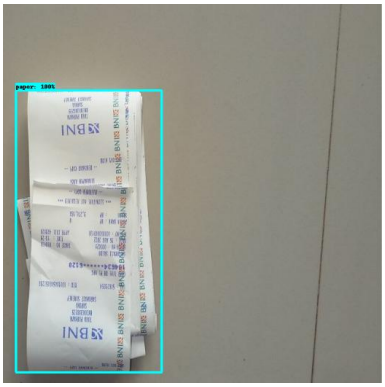


Figure 7. Results of Faster RCNN ResNet-101 detection of paper waste

Figure 8. Faster RCNN ResNet-101 detection results for plastic bottle waste

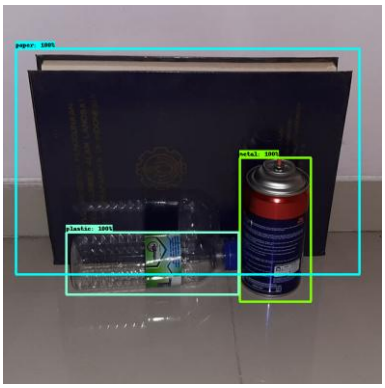


Figure 9. Faster RCNN ResNet-101 detection results for recyclable waste classification

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 26 - 32, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1006

3.3 Total Loss in the Training Process

Table 3. Total Loss Results in Faster R-CNN ResNet-50 and ResNet-101 Network Architecture Models

Network Architecture	Amount of Training Data	Total Loss Step tO-			$\bar{x}$
		3000	4000	5000	
Faster R-CNN ResNet 50	80	0.2101	0.2058	0.1808	0.1989
	120	0.2081	0.1988	0.1798	0.1956
	200	0.2003	0.1805	0.1536	0.1781
				Average	0.1909
Faster R-CNN ResNet 101	80	0.2485	0.2181	0.1883	0.2183
	120	0.22	0.1634	0.1873	0.1902
	200	0.1958	0.1037	0.1834	0.1610
				Average	0.1898

Based on the results obtained in Table 3, the smallest average total loss was obtained in the Faster R-CNN ResNet-101 network architecture of 0.1898, while the largest average total loss was in the Faster R-CNN ResNet-50 network architecture model of 0.1909 [8].

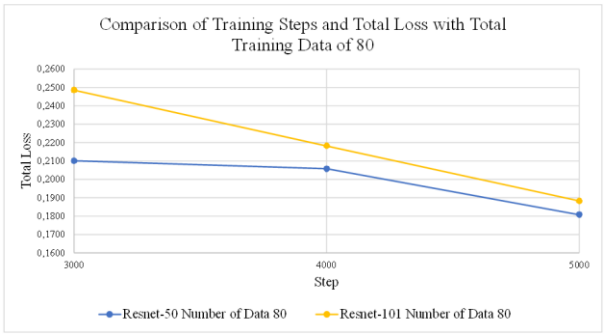


Figure 10. Effect of Training Steps and Total Loss on Faster R-CNN ResNet-50 and ResNet-101 Network Architecture Models with 80 Training Data

Based on the results of Figure 10, it shows that the Faster R-CNN ResNet-50 and ResNet-101 network architecture types at steps 3000 to 5000 show decreasing loss values. However, the smallest average total loss was obtained in the Faster R-CNN ResNet-101 step 5000 network architecture [9].

Figure 11. Effect of Training Steps and Total Loss on Faster R-CNN ResNet-50 and ResNet-101 Network Architecture Models with 120 Training Data.

Figure 12. Effect of Training Steps and Total Loss on Faster R-CNN ResNet-50 and ResNet-101 Network Architecture Models with 200 Training Data

Based on the results of Figure 11 and Figure 12, it shows that the Faster R-CNN ResNet-50 network architecture type at steps 3000 to 5000 shows a decreasing loss value. However, the Faster R-CNN ResNet-101 network architecture type from steps 4000 to 5000 experienced an increase in the total loss value except from steps 3000 to 4000 which experienced a decrease in total loss. The smallest average total loss is obtained in the Faster R-CNN ResNet-101 network architecture at step 4000 [10].

The results obtained in Table 3 to Figure 12 show that the number of steps in the training process influences the total training loss value. As can be seen in the Faster R-CNN ResNet-50 network architecture type at steps 3000 to 5000, the loss value is decreasing. However, the Faster R-CNN ResNet-101 network architecture type from steps 4000 to 5000 experienced an increase in the total loss value except from steps 3000 to 4000 which experienced a decrease in total loss. The smallest average total loss was obtained in the Faster R-CNN ResNet-101 network architecture, while the largest average total loss was in the Faster R-CNN ResNet-50 network architecture model [11].

3.4 Total Loss in the Testing Process

Table 4. Results of Total Loss Testing on Faster R-CNN ResNet-50 and ResNet-101 Network Architecture Models

Network Architecture	Step	Trial-					
		1	2	3	4	5	6
Faster R-CNN ResNet 50	3000	0.493	0.492	0.433	0.430	0.413	0.408
	4000	0.448	0.445	0.406	0.401	0.386	0.386
	5000	0.439	0.436	0.371	0.370	0.367	0.364
Faster R-CNN ResNet 101	3000	0.254	0.251	0.221	0.219	0.211	0.211
	4000	0.226	0.223	0.203	0.202	0.187	0.189
	5000	0.226	0.224	0.219	0.211	0.194	0.193

Based on the results obtained in Table 4, it shows that the number of steps. The training process influences the total training loss value. As can be seen in the Faster R-CNN ResNet-50 network architecture type at steps 3000 to 5000, the loss value is decreasing. However, the Faster R-CNN ResNet-101 network architecture type from steps 4000 to 5000 experienced an increase in the total loss value except from steps 3000 to 4000 which experienced a decrease in total loss. The smallest average total loss was obtained in the Faster R-CNN ResNet-101 network architecture, while the largest

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 26 - 32, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1006

average total loss was in the Faster R-CNN ResNet-50 network architecture model [12].

3.5 Pressicion, Recall, dan F1 Score

At the testing stage for recyclable waste detection, the parameters that will be created to group the detection results are True Positive (TP), False Positive (FP), and False Negative (FN). In this study, test results were obtained from 30 and 50 images which were used as test images for detecting recyclable waste objects with a maximum amount of training data of 200 images [13]. System capabilities based on the Faster R-CNN ResNet-50 and ResNet-101 architectural network models are shown in Table 5 to Table 6 respectively.

Table 5. Experimental Results on the Faster R-CNN ResNet 50 Network Architecture Model

Faster R-CNN ResNet-50				
Trial-		Step		
		3000	4000	5000
1	TP	13	14	15
	FP	14	13	13
	FN	3	3	2
2	TP	22	25	26
	FP	24	23	21
	FN	4	2	3
3	TP	14	19	20
	FP	11	6	5
	FN	5	5	5
4	TP	24	28	35
	FP	18	19	14
	FN	8	3	1
5	TP	20	22	23
	FP	8	4	4
	FN	2	4	3
6	TP	34	38	40
	FP	12	9	7
	FN	4	3	3

Table 6. Experimental Results on the Faster R-CNN ResNet-101 Network Architecture Model [14].

Faster R-CNN ResNet-101				
Trial-		Step		
		3000	4000	5000
1	TP	14	15	17
	FP	13	13	11
	FN	3	2	2
2	TP	24	26	29
	FP	23	22	20
	FN	2	2	1
3	TP	16	20	18
	FP	12	8	8
	FN	2	2	4
4	TP	28	35	31
	FP	18	8	11
	FN	4	7	8
5	TP	21	24	23
	FP	5	4	3
	FN	4	2	4
6	TP	36	42	40
	FP	9	7	8
	FN	5	1	2

Based on the values obtained in Table 5 and Table 6, this can be done by calculating the precision, recall and F1 Score values. To calculate these values, you can look at equations (3.1), (3.2), and (3.3) [15]. The results of precision, recall and F1 Score calculations based on the ResNet-50 and ResNet-101 architectural network models with the Faster R-CNN algorithm are respectively shown in Table 7 and Table 8.

Table 7. Precision, Recall and F1 Score results based on the ResNet-50 architecture network model [16]

Trial-	Step	Precision	Recall	F1 Score
1	3000	48%	81%	60%
	4000	52%	82%	64%
	5000	54%	88%	67%
	Average			64%
2	3000	48%	85%	61%
	4000	52%	93%	67%
	5000	55%	90%	68%
	Average			65%
3	3000	56%	74%	64%
	4000	62%	80%	70%
	5000	80%	80%	80%
	Average			71%
4	3000	57%	75%	65%
	4000	60%	90%	72%
	5000	71%	97%	82%
	Average			73%
5	3000	71%	91%	80%
	4000	85%	85%	85%
	5000	85%	88%	87%
	Average			84%
6	3000	74%	89%	81%
	4000	81%	93%	86%
	5000	85%	93%	89%
	Average			85%
Average				63%

Table 8. Precision, Recall and F1 Score results based on the ResNet-101 architecture network model

Trial-	Step	Precision	Recall	F1 Score
1	3000	52%	82%	64%
	4000	54%	88%	67%
	5000	61%	89%	72%
	Average			68%
2	3000	51%	89%	65%
	4000	54%	93%	68%
	5000	59%	97%	73%
	Average			69%
3	3000	57%	89%	70%
	4000	71%	91%	80%
	5000	69%	82%	75%
	Average			75%
4	3000	61%	88%	72%
	4000	81%	83%	82%
	5000	74%	79%	77%
	Average			77%
5	3000	81%	84%	82%
	4000	86%	92%	89%
	5000	88%	85%	87%
	Average			86%
6	3000	80%	88%	84%

Trial-	Step	Precision	Recall	F1 Score
	4000	86%	98%	91%
	5000	83%	95%	89%
	Average			88%
Average				77%

Based on Table 7 and Table 8, the ResNet-50 and ResNet-101 network architecture models using the Faster R-CNN algorithm obtained average F1 Score values of 63% and 77%, respectively. This shows that ResNet-101 performs better when used in the Faster R-CNN algorithm [17]. The F1 Score value for experiments with 50 test data is better than experiments with 30 test data for each method. This is because, 20 images from the test data of 50 images were obtained from images that had been trained. The best F1 Score value was obtained on the ResNet-101 network architecture model in experiment 6 and step 4000.

3.6 Comparison of Test Results of Two Architectures

Based on several test results that have been carried out, a table summarizing the overall data from the two network architecture models used can be created which is shown in Table 9 [18].

Table 9 Comparison of Test Results from the Faster R-CNN ResNet-50 and Faster R-CNN ResNet-101 Network Architecture Models

Architectural Models	Parameter	Mark
Faster R-CNN ResNet-50	Average Computational Training Time	20 minute 37 second
	Average Test Time Computation	01 minute 29 second
	Average Total Training Loss	0.1909
	Average *F1 Score*	**63%**
Faster R-CNN ResNet-101	*Average Computational Training Time*	26 minute 01 second
	Average Test Time Computation	01 minute 34 second
	Average Total Training Loss	0.1898
	F1 Score Average	**77%**

Based on Table 9, it can be seen that computing the Faster R-CNN Resnet-101 for both the training and testing processes takes longer than the Faster R-CNN Resnet-50 architectural model. However, the average total loss of Faster R-CNN Resnet-101 training is smaller than Faster R-CNN Resnet-50, namely 0.1898 and the average F1 score value is greater, reaching 77% [19]. On the other hand, the Faster R-CNN Resnet-50 computation for both the training and testing processes requires faster time compared to the Faster R-CNN Resnet-101 architecture model. However, the average total training loss produced was the largest, namely 0.1909 and the average F1 score produced was low, namely 63%.

3.7 Effect of Average Loss on Average F1 Score

The influence of the average total loss on the average F1 score results is shown in Figure 13. The graph of the average total loss and the average F1 score is the overall average of the two architectures used.

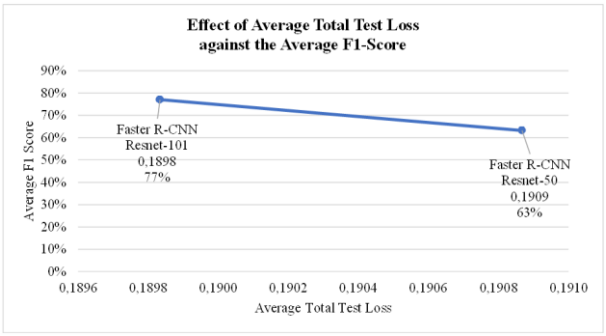


Figure 13. Graph of the influence of the average total loss on the average F1 Score

Based on Figure 13, the smaller the average total loss in the training process, the better the model is at learning object features so that the average F1 score value is greater. Recyclable waste detection errors occur because the object is too small or the image resolution is not good. Apart from that, detection errors are also caused by less complex data annotation processes [20].

4. CONCLUSION

From the series of tests that have been carried out, various things can be concluded that the two network architecture models studied resulted in an average F1 score of 63% and 77% respectively, where the best F1 score is found in Faster R-CNN ResNet-101, but with a longer computation time when compared to the Faster R-CNN ResNet-50 architecture model, where the average computation time for the training process of Faster R-CNN ResNet-50 and ResNet-101 is 20 minutes 37 seconds and 26 minutes 01 seconds respectively. While the average computation time for testing Faster R-CNN ResNet-50 and ResNet-101 is 01 minute 29 seconds and 01 minute 34 seconds respectively. The best F1 Score value was obtained on the ResNet-101 network architecture model in experiment 6 and step 4000, namely 91%.

5. REFERENCES

[1] F. F. Putra and Y. D. Prabowo, "Low resource deep learning to detect waste intensity in the river flow," *Bull. Electr. Eng. Informatics*, vol. 10, no. 5, pp. 2724–2732, 2021, doi: 10.11591/eei.v10i5.3062.

[2] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," pp. 1–39, 2019.

[3] W. Zhiqiang and L. Jun, "A review of object detection based on convolutional neural network," *Chinese Control Conf. CCC*, pp. 11104–11109, 2017, doi: 10.23919/ChiCC.2017.8029130.

[4] D. S. W. Ting *et al.*, "Artificial intelligence and deep learning in ophthalmology," *Br. J. Ophthalmol.*, vol. 103, no. 2, pp. 167–175, 2019, doi: 10.1136/bjophthalmol-2018-313173.

[5] M. Aghilan, A. M. Kumar, M. T. Aafrid, N. A. Kumar, and S. Muthulakshmi, "Garbage Waste Classification Using Supervised Deep Learning Techniques," *Int. J. Emerg. Technol. Innov. Eng.*, vol. 6, no. 3, pp. 77–82, 2020.

[6] R. Sultana, R. D. Adams, Y. Yan, P. M. Yanik, and M. L. Tanaka, "Trash and Recycled Material Identification using Convolutional Neural Networks (CNN)," *Conf. Proc. - IEEE SOUTHEASTCON*, vol. 2020-March, 2020, doi: 10.1109/SoutheastCon44009.2020.9249739.

[7] I. Salimi, B. S. B. Dewantara, and I. K. Wibowo, "Visual-based trash detection and classification system

for smart trash bin robot," *Int. Electron. Symp. Knowl. Creat. Intell. Comput. IES-KCIC 2018 - Proc.*, pp. 378–383, 2019, doi: 10.1109/KCIC.2018.8628499.

[8] D. Atmajaya, N. Kurniati, W. Astuti, Y. Salim, and A. Haris, "Digital Scales System on Non-Organic Waste Types Based on Load Cell and ESP32," *Proc. - 2nd East Indones. Conf. Comput. Inf. Technol. Internet Things Ind. EIConCIT 2018*, pp. 308–311, 2018, doi: 10.1109/EIConCIT.2018.8878667.

[9] Y. Dewilda and Julianto, "Kajian Timbulan, Komposisi, dan Potensi Daur Ulang Sampah Sebagai Dasar Perencanaan Pengelolaan Sampah Kawasan Kampus Universitas Putra Indonesia (UPI)," *Semin. Nas. Pembang. Wil. dan Kota Berkelanjutan*, vol. 1, no. 1, pp. 142–151, 2019, doi: 10.25105/pwkb.v1i1.5270.

[10] M. F. Rahman and Bambang, "Deteksi Sampah pada Real-time Video Menggunakan Metode Faster R-CNN," *Appl. Technol. Comput. Sci. J.*, vol. 3, no. 2, pp. 117–125, 2021, doi: 10.33086/atcsj.v3i2.1846.

[11] Y. Wang and X. Zhang, "Autonomous garbage detection for intelligent urban management," *MATEC Web Conf.*, vol. 232, pp. 1–5, 2018, doi: 10.1051/matecconf/201823201056.

[12] D. Patel, F. Patel, S. Patel, N. Patel, D. Shah, and V. Patel, "Garbage Detection using Advanced Object Detection Techniques," *Proc. - Int. Conf. Artif. Intell. Smart Syst. ICAIS 2021*, pp. 526–531, 2021, doi: 10.1109/ICAIS50930.2021.9395916.

[14] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li, "Computer vision algorithms and hardware implementations: A survey," *Integration*, vol. 69, no. April, pp. 309–320, 2019, doi: 10.1016/j.vlsi.2019.07.005.

[15] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," pp. 1–39, 2019, [Online]. Available: http://arxiv.org/abs/1905.05055.

[16] M. Sarosa, N. Muna, and E. Rohadi, "Performance of faster r-cnn to detect plastic waste," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 5, pp. 7756–7762, 2020, doi: 10.30534/ijatcse/2020/120952020.

[17] A. Wahyudinata and H. B. Dirgantara, "Pengembangan Gim Edukasi 2D Pemilahan Sampah Daur Ulang Berbasis Android," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 20, no. 1, pp. 129–138, 2020, doi: 10.30812/matrik.v20i1.860.

[18] I. Febriadi, "Abdimas : Papua Journal of Community Service Pemanfaatan Sampah Organik Dan Anorganik Untuk Mendukung Go Green Concept Di Sekolah," vol. 1, no. 1, pp. 32–39, 2019.

[19] A. Novianti and I. N. Muliarta, "Eco-Enzym Based on Household Organic Waste as Multi-Purpose Liquid," *AGRIWAR J.*, vol. 1, no. 1, pp. 12–17, 2021. M. Toğaçar, B. Ergen, and Z. Cömert, "Waste classification using AutoEncoder network with integrated feature selection method in convolutional neural network models," *Elsevier Meas.*, vol. 153, 2020, doi: 10.1016/j.measurement.2019.107459.

[20] B. Chitradevi and P. Srimanthi, "An Overview on Image Processing Techniques," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 2, no. 11, pp. 6466–6472, 2014, [Online]. Available: http://www.hindawi.com/isrn/sp/2013/496701/

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 33 - 40, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1007

A Statistical Analysis on the Old Age Homes of Urban Verses Rural Verses Semi-Rural

Swarada Joshi
Student, MSc(Statistics),
Modern College of Arts,
Science & Commerce
(Autonomous), Pune16
India

Rutuja Joshi
Student, MSc(Statistics),
Modern College of Arts,
Science & Commerce
(Autonomous), Pune16
India

Vedant Dusing
Student, MSc(Statistics),
Modern College of Arts,
Science & Commerce
(Autonomous), Pune16
India

Dr. Dipali Meher
Assistant Professor,
Department of Computer
Science,
Modern College of Arts,
Science & Commerce,
(Autonomous), Pune16
India

Prof. Rajashshree Umarani
Assistant Professor,
Department of Statistics,
Modern College of Arts,
Science & Commerce,
(Autonomous), Pune 16
India

Abstract: Due to urbanization and standard of living it has been seen that there is increased demand of old age homes. In this paper we have studied 12 old age homes from urban, rural and semi-rural areas. All types of old ages homes including location, ownership, management, size, capacity, utilization and different facilities provided by are studied. As part of this research, 60 operational senior living community projects were enumerated across India. The study was conducted for 178 units among 12 old age homes. The data is collected and various statistical techniques are applied to the data and various results were found.

Keywords: City old age home; semi-rural old age home; rural old age home; problems of old age peoples

1. INTRODUCTION

India is getting older, and that means more people are becoming senior citizens, aged 60 and above. Right now, about 8.6% of India's population falls into this age group. But by the year 2050, this number is expected to triple, making up about 20% of the total population.

There are some important changes happening in families. More women are working and families are becoming smaller. This means that the traditional way of taking care of elderly family members within extended families is changing. However, taking care of elderly people in India is still not very good, and the system to help them is not very developed. There's not enough good information about how many places there are for senior citizens in India. Some estimates say there are between 600 to 1,700 places where seniors can live comfortably. But we don't have all the facts to know for sure. A study looked at many places where elderly people live, like old age homes and special communities for seniors, in 84 different places across India. They looked at things like how big these places are, how much they cost, and who owns and runs them. Based on this study and information from the 2011 census, they think there might be around 1,150 places for seniors in India. These places can house about 97,000 elderly residents.

2. LITERATURE SURVEY

Shanti Johnson et.al in their paper studied current status of old age home in India along with the facilities provided [1]. Priyanka V Janbandhu et.al in their paper studied about old age home in Pune city and noted that there are more women than men in old age homes[2]. Majority of widowed/widower elderly and elderly who came from nuclear family live in the oldhomes due to lack of care and support and death of their partners, especially women have prolonged widowhood due to longer life expectancy than men[3].The article by S Sudha et.al examined the impact of familial social support ties (indicated by marital status, kin availability, sources of economic support, and frequency and quality of emotional interaction) on subjective health perception among a sample of elderly men and women aged 60 and older in South India [4]. The study by author Gaitri Rajkumari revealed that most of the elderly in the old age homes were from rural background, illiterate, widowed and are economically dependent on others. The female residents were more in number. The findings revealed that the factors that compelled them to join old age homes are verbal abuse of daughter in law, financial constraints, verbal abuse of son, nobody to look after, physical abuse, tarnishing self-respect, health issues etc. [5].

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 33 - 40, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1007

3. RESEARCH METHODOLOGY

The objectives to carry out this research are:

- To determine the degree of association or relation between the two attributes
- Is there any association or relation between staying at old age homes and the reasons
- The finding of study is that there are more tablet seekers in old age homes in urban areas compared to rural areas
- There were eyesight problems among residents of old age homes in rural areas as compared to Semi-rural areas.
- The study has found the statistical significance difference in the gender distribution of residents of urban old age homes
- To determine the correlation between age and disease of old age home

These are the areas which are important to know while doing research, so we made it our priority to look at this too, if we look at more things which matter.

4. VARIOUS PROBLEMS IDENTIFIED

How to enable age with dignity is the biggest challenge that elderly people face today. So following problems were identified.

- Physiological problems
- Psychological problems
- Senile Dementia
- Emotional Problem
- Social Problem
- Financial problem
- Mobility barrier
- Depression

Physiological Problems

Living in old age homes can lead to health problems because there's often no regular medical care or check-ups. This lack of care can result in physical issues. Aging itself can cause physical changes. It depends on things like your genes, how you live, and the environment you're in. Other factors like a bad diet, not eating enough, infections, toxins, overeating, not getting enough rest, stress, working too much, hormonal issues, and extreme weather can also affect your health as you get older.

Your skin may become rough and lose its elasticity, forming wrinkles, and your veins might become more visible. Changes in your nervous system can also affect your brain. Some organs, like the spleen, liver, and soft organs, might shrink. The ratio of heart weight to body weight can decrease over time. These are all common physical changes that can happen as you age, especially if you don't get proper care and attention.

Psychological Problems

Living in old age homes can make people feel lonely and emotionally weak because they are not with their families. This can sometimes lead to mental problems. Mental issues are common in old age. Older people can experience things like severe sadness or confusion in their minds.

Senile Dementia

In old age homes, they often don't have a quiet place to sleep because they share rooms with many people. This lack of good sleep can cause health problems. Older people can suffer from something called "senile dementia".

Emotional Problem

When families don't pay much attention to their older relatives and don't seem to care, it can make the older people very sad. They miss their family a lot and sometimes cry because they feel so alone in the old age home. They really want to be with their family and wait for them to visit.

But when they think about how their family is treating them badly or ignoring them, it can make them feel even sadder, and this sadness can affect their mental health. So, it's important for families to be kind and loving to their older family members to prevent these emotional problems.

Social Problem

In the old age home, older people have to stay there all the time, and they're not allowed to go outside. This means they can't be with others or go places they like. They feel lonely and stuck, and this is the main social problem they face in old age homes.

Financial Problem

Just like everyone else, elderly people have things they want and need. When they live in an old age home, they rely on the organization to meet those needs because they don't have any money to spend on themselves.

Mobility Barrier

The elderly people living in the old age home can't go outside the area surrounded by walls. They have to stay within the home's limits. Even inside, they have specific times for things like lunch, dinner, and watching TV, and they have to follow this schedule.

Depression

Feeling lonely in old age homes can make depression more likely, and this connection between loneliness and depression can affect how long someone lives. So, in very old people, depression only seems to lead to a higher chance of passing away if they also feel lonely.

Depression often goes hand in hand with feeling lonely. Sometimes, when someone is lonely, they might show signs like withdrawing from others, feeling anxious, lacking motivation, or being sad. These signs of depression can look a lot like the signs of loneliness.

5. RESEARCH DESIGN

Nature of the study

The study is descriptive and analytical in nature.

Nature of the data

Primary data collection was done from old age homes. Questionnaire is used to collect this data.

Sample Design:

Nature of the population	finite
Sample Size	178
Sample Unit	Each unit is selected from the old age homes
Sampling method	Purposive sampling

6. SOURCE OF DATA

I. Source of Data

Google form is prepared and questions related to name, age, gender, education, service, children, reasons for living in rural or urban area, any chronic diseases the person will have, medicines taken, economic condition, emotional problems, related to exercise,

II. Tools for Analysis

- Percentage and comparative analysis have been done.
- Pie-chart, Bar graph, Multiple bar graph, Histogram and diagrams are used for presentation.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 33 - 40, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1007

7. DATA ANALYSIS & INTERPRETATION

Data analysis was done and it shows following results.

Table 7.1:- Respondent's Gender.

Gender	No of respondent	Percentage
Male	85	47.75%
Female	93	52.25%
Total	178	100%

Source: - Primary data

The table and graph show the gender classification of respondent 47.75% are male and 52.25% of respondents are female.

Chart 7.1:- It shows the ratio of Male and Female in old age home

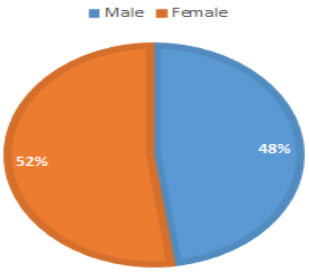


Table 7.2:- shows the difference between the citizens among the Urban vs Semi Rural vs Rural

Places	No. of citizens	Percentage
Urban	73	41.01%
Semi-Rural	48	26.96%
Rural	57	32.02%

Table and graph show the citizens in old age homes are 41.01% of Urban, 26.96% of Semi- Rural, and 32.02% are from Rural.

Chart 7.2: - Chart shows the no of citizens from which area

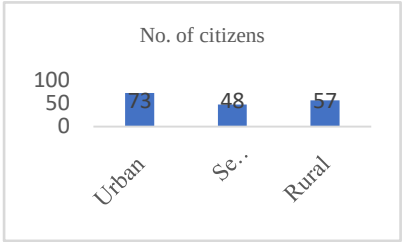


Table 7.3:- shows that how many citizens are Literate and Illiterate

Literate	69
Illiterate	109

Here 38.8% are Literate and 61.2% are Illiterate in old age homes

Chart 7.3:- Chart shows the pie chart of the literacy rate among citizens of old age homes

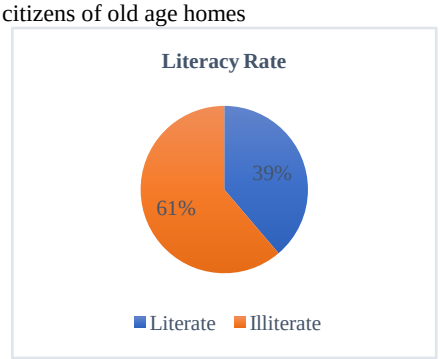


Table 7.4:- Shows how many people are in service and not

In service	43
Not in service	135

Here the ratio of citizens in service is 26.7% and not in service are 76.3%

Chart 7.4:- Chart shows the service ratio of the citizens of old age homes

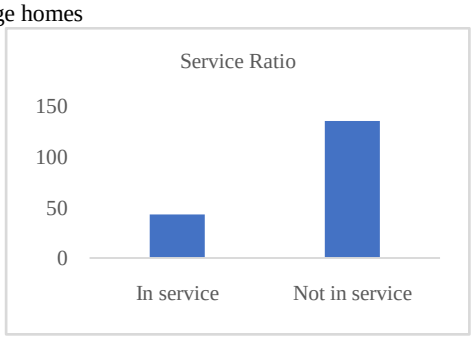


Table 7.5:- This graph shows how many children do they have

Children	Total	Percentage
0	35	19.66%
1	35	19.66%
2	57	32.02%
3	31	17.41%
4	13	0.073%
5	2	0.011%
6	5	0.028%

Table and chart show that the graph of the children, Of the old age home people

Chart 7.5:- Chart show that the ratio of the children of old people

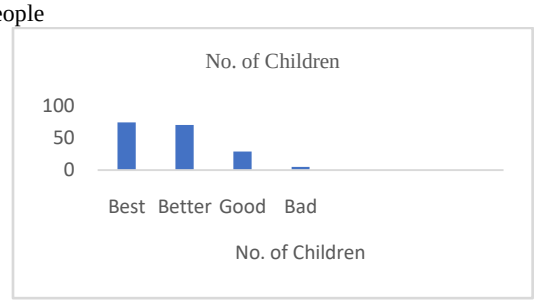


Table 7.6:- Shows how many of them get the pension through their service for the government or by the schemes of government.

Do you get pension	No. of citizen	Percentage
Yes	40	22.47%
No	138	77.53%
Total	178	100%

From the table and figure we have more no of citizens who do not get pension where the numbers are 22.47% of Yes and 77.53% of are No.

Chart 7.6:- Chart shows the ratio of the citizens who get pension

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 33 - 40, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1007

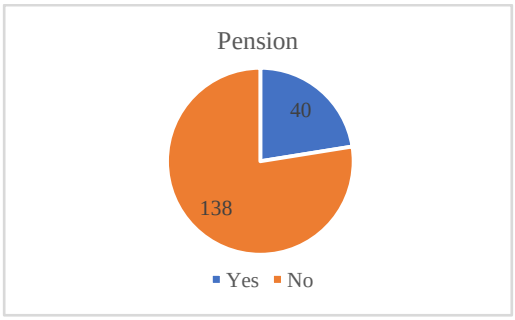


Table 7.7:- shows the health condition of the citizen living there.

Health condition	No. of citizen	Percentage
Best	74	41.6%
Better	70	39.3%
Good	29	16.3%
Bad	5	2.8%
Total	178	100%

Table and graph shows that the health condition among the citizens of the old age homes here whose condition is not good has very less number we can neglect them.

Chart 7.7:- Chart show the ratio of health condition of the citizens of old age home

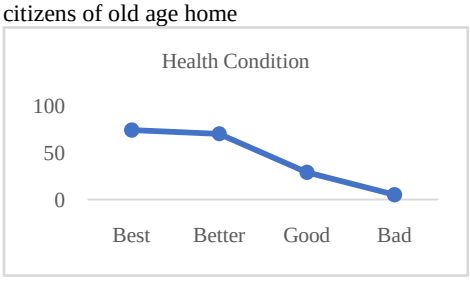


Table 7.8:- Shows that what is the main reason they live here

Reasons	Numbers	Percentage
Daughter-in-law does not behave properly	44	24.71%
Son does not behave properly	48	26.96%
Children in abroad	21	11.79%
No Children	34	19.10%
Other	31	17.41%
Total	178	100%

Table and chart give us the information about the main reason why they people live here where

Chart 7.8:- Chart shows the graph of the reasons and numbers

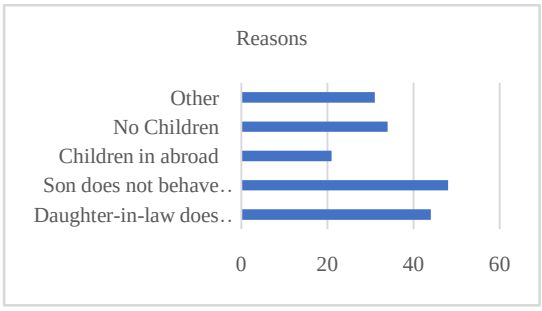


Table 7.9:- Shows that the old people are taking any kind of tablet for their health

Tablet takers	No of citizens	Percentage
Yes	109	61.23%
No	69	38.76%
Total	178	100%

Table and chart shows that the tablets consumed by the old people are 61.23% and 38.76% do not take any type of medicine.

Chart 7.9:- Chart shows the percentage of tablet seekers in old age home

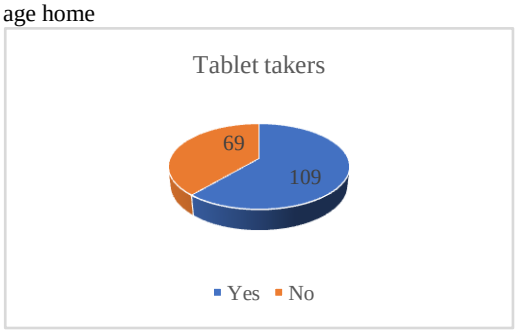


Table 7.10:- Shows that the they have any chronic health disease

Chronic health disease	No of citizens	Percentage
Yes	109	61.23%
No	69	38.77%
Total	178	100%

Chart and Table shows the no of citizens having the chronic health disease where 61.23% are having and 38.77% are not having chronic health disease

Chart 7.10:- Chart shows the representation of the chronic health disease

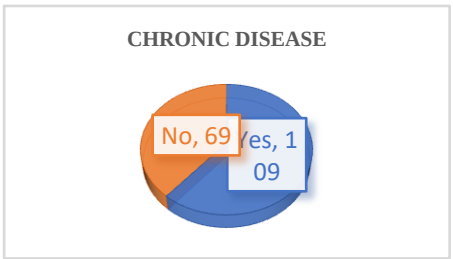


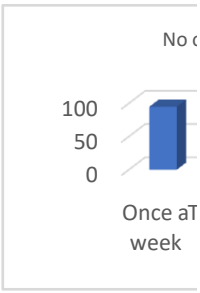
Table 7.11:- Show that the old people get call from their family, relatives, friends or etc.

Calls from Home	No of calls	Percentage
Once a week	95	53.37%
Twice a week	19	10.67%
No calls	64	35.95%
Total	178	100%

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 33 - 40, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1007

Table and Graph shows that how many calls they get from their children, relative, friends.

Chart 7.11:- Charts shows the graphical representation of the people getting calls from home



Age Groups	No of people
60 – 65	40
65 – 70	36
70 – 75	30
75 – 80	29
80 – 85	25
85 – 90	10
90 – 95	6
95 - 100	2

Table 7.12:- Shows that they have eyesight problem or not

Eye sight problem	No of patients	Percentage
Yes	58	33.15%
No	119	66.85%
Total	178	100%

Chart 7.12: -Chart shows the representation of the eye sight problems of old people in old age home.

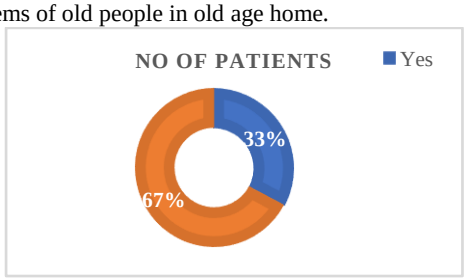
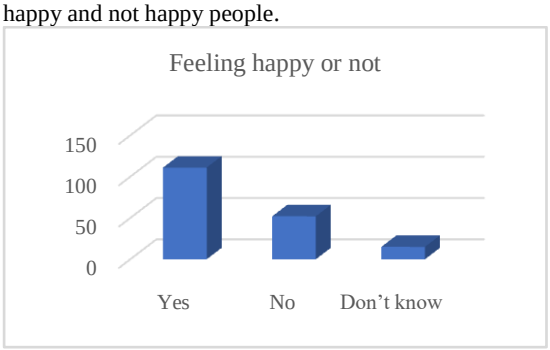


Table 7.13:- Shows that people staying here are happy or not

Happy or Not	No. of citizens	Percentage
Yes	111	62.35%
No	52	29.21%
Don't know	15	8.42%
Total	178	100%

Table and Graph shows that the old people staying there are happy but only 62.35%, 29.21% people are not happy and remaining 8.42% people don't know.

Chart 7.13:- Shows the graphical representation between happy and not happy people.



Fitting of Log-Normal Distribution

Here we took the data of age of old age home which is above 60 years and goes upto 100 years.

1. We fit the log-normal distribution to check whether our data is positively skewed i.e. has long tail on positive side
2. If our data fits the Log-normal distribution, then we can use various test.

Solution:- ESTIMATION OF PARAMETER

X	fi	xi	fi*xi	Fi*xi*xi
Age Groups	No. of people	Midpoint		
60 – 65	40	62.5	2500	156250
65 – 70	36	67.5	2430	164025
70 – 75	30	72.5	2175	157687.5
75 – 80	29	77.5	2247.5	174181.25
80 – 85	25	82.5	2062.5	170156.25
85 – 90	10	87.5	875	76562.5
90 – 95	6	92.5	555	51337.5
95 - 100	2	97.5	195	19012.5
Total	178	-	13040	969212.5

$m1^1 = 73.2584$

$m2^2 = 5445.01404$

S=sqrt(m2)=73.79033

Estimation of σ $LOG_e = LN(m2^1/(m1^1*m1^1))$

Estimation of σ = 0.12029

Estimation of μ = 4.28676

Table to find expected frequency

Age group	No. of people	Upper limit	P (X< li)	Pi	Ei
0 – 60	0	60	0.05484	0.05484	7.7617
60 – 65	40	65	0.17510	0.12026	21.406
65 – 70	36	70	0.37520	0.2001	35.618
70 – 75	30	75	0.60081	022561	40.159
75 – 80	29	80	0.78582	0.185	32.930
80 – 85	25	85	0.90251	0.11669	20.771
85 – 90	10	90	0.96173	0.05922	10.541
90 – 95	6	95	0.986813	0.02508	4.4639
95 -100	2	100	0.99594	0.00913	1.6246
100 - ∞	0	∞	1	0.00406	0.7225
Total	178	-	-	1	178

We wish to test

H_0 = Fitting is good i.e. log normal distribution fits the data.

H_1 = Fitting is not good i.e. log normal distribution does not fit the data.

Under H_0 the test statistic is

Chi – square (cal):- |Oi - Ei |^2=∑Ei

Has Chi – square distribution with

(n – k – 1) = 7-2-1

Table for calculation of value of test statistics

Sr. No	Oi	Ei	(|Oi - Ei |^2)/Ei
1	40	31.1682	2.50255
2	36	65.6181	0.00409
3	30	40.1593	2.57006

4	29	32.9303	0.46909
5	25	20.7712	0.86098
6	10	10.5416	0.027826
7	8	6.81119	0.20749
Total			6.642085

Chi-Square cal = 6.642085
Chi-Square table at d.f = 9.487729
n = 7 no. of class after pooled
k = 2 no. of parameter estimated Comparison
Chi-square cal < Chi-square table value
We accept the null hypothesis is 5% loss and conclude that log normal distribution fits the data.

Testing of hypothesis

RUN TEST:

Here the sequence of Male and Female is,

Area	Daughter-in-law does not behave properly	Son does not behave properly	Children en in abroad	No child	Other	Total
City (Urban)	18.0449	19.6853	8.6123	13.9438	12.7134	73
Semi Rural	12.1123	13.2134	5.7808	9.35955	8.5337	49
Rural	13.8426	15.1011	6.6067	10.6966	9.7528	56
Total	44	48	21	34	31	178

FFF M F M F MMM FFF MMM FF M FFFFFFFFF M F M FF MMM FF M FF M FFF MM FFFFFFFFFFF MMMMMMM F MMMMMMMMM F MMMMMMMM F MMMMM FFFF MMMM FFF M FFF M F M FFFF M FFFF MM FFFFF M FFFFFFFF MMMMM FF MMMM F MMM FFF MM FF MMM FF MMMM F M FFF M F MM F MM

Sample size 178 > 20
So, we will apply Normal Approximation.
Here we set hypothesis as,
H_0: Sample is random
H_1: sample is not random
To test hypothesis, we find the value of r,
i.e. Number of r.v's
r = Number of Runs = 64
n1 = number of Male = 85
n2 = number of Female = 93
Test Statistics: -
$Z_{cal} = r - E(r)/\sqrt{V(r)}$
Consider,
E(r) = {(2*n1*n2)/(n1+n2)} + 1
= {(2*85*93)/(85 + 93)} + 1
E(r) = 89.83
Var(r) = [2*n1*n2*(2*n1*n2-n1-n2)] / [(n1 + n2)^2 * (n1+n2-1)]
= [2*85*93*(2*85*93-85-93)] / [(85 + 93)^2 *(85+93-1)]
Var(r) = 44.069
Now,
Z_{cal} = -3.891
$|Z_{cal}| = |-3.891|$
$Z_{ca}| = 3.891$
At $\alpha = 0.05$
$Z_{tab} = Z\ \alpha/2 = 1.96$
$Z_{tab} > Z_{cal}$
Therefore, Accept H_0

Conclusion: - Sample we took from the old age home is random.

Chi-Square test 1

Health condition is dependent on
Gender
To test the given fact, set up hypothesis as:
H_0 : Two attributes under study are independent.
Observed frequencies

Gender	Best	Better	Good	Total
Male	32	23	42	97
Female	42	16	23	81
Total	74	39	65	178

Expected Frequency

Gender	Best	Better	Good	Total
Male	40	21	35	97
Female	34	18	30	81
Total	74	39	65	178

To test this hypothesis, use p-value
and conclude.
P-value = 0.03374 > 0.01
Hence, Accept H_0
Conclusion: Health condition independent on Gender

Chi-Square test 2

Chi square test Independence of attributes
Relation between staying in an old age home and Reasons.
H_0 : There is no significant relation between staying at old age home and reasons.
H_1 : There is significant relation between staying at old age home and reasons.

Area	Daughter-in-law does not behave properly	Son does not behave properly	Children in abroad	No child	Other	Total
City (Urban)	9	16	8	21	19	73
Semi Rural	14	16	9	6	4	49
Rural	21	16	4	7	8	56
Total	44	48	21	34	31	178

Expected Frequency

Area	Daughter-in-law does not behave properly	Son does not behave properly	Children in abroad	No child	Other	Total
City (Urban)	18.0449	19.6853	8.6123	13.9438	12.7134	73
Semi Rural	12.1123	13.2134	5.7808	9.35955	8.5337	49
Rural	13.8426	15.1011	6.6067	10.6966	9.7528	56
Total	44	48	21	34	31	178

Calculation for Chi-Square test

Observed values	Expected values	(O - E)	(O - E)^2	[(O - E)^2]/E
9	18.04	-9.04	81.7216	4.5300
16	19.68	-3.68	13.5424	0.6881
8	8.61	-0.61	0.3721	0.0432
21	13.94	7.06	49.8436	3.5755
19	12.71	6.29	39.5641	3.1128
14	12.11	1.89	3.5721	0.2949
16	13.21	2.79	7.7841	0.5892
9	5.78	3.22	10.3684	1.7938
6	9.35	-3.35	11.2225	1.2002
4	8.53	-4.53	20.5209	2.4057
21	13.84	7.16	51.2656	3.7041
16	15.1	0.9	0.81	0.0536
4	6.6	-2.6	6.76	1.0242

6	10.69	-4.69	21.9961	2.0576
8	9.75	-1.75	3.0625	0.3141
Total				25.3876

Chi -square calculated = 25.3876
Chi square tabular:-
Degrees of freedom
= (Column – 1) (Rows -1)
= 4*2
= 8
Level of Significance = a = 0.05
Formulae:- =CHISQ.INV.RT(0.05,8)
Chi square tabular = 15.507
Chi square calculated > Chi square tabular
Hence, we reject Null Hypothesis and Accept Alternative Hypothesis
Conclusion: - There is a significant relation between staying at old age homes and the Reason.
Proportion Z test
To obtain the proportion of tablet consumer in urban and rural
Let,
X_1 =The tablet consumers selected from the Urban = 37
X_2 =The tablet consumers selected from the Rural = 12
Given,

n1	73
n2	57

Now, we compute the sample proportion.

p1= X1/n1	37/73	0.5068
p2= X2/n2	12/57	0.2105

We wish to test,
H_0 : p1=p2
i.e. There is no significant difference between two proportion
H_1 : p1>p2
i.e. There is significant difference between two proportion
Under H_0
The test statistic is
P = (p1n1+p2n2) / (n1+n2)
P = |36.9964 + 7.7885|/130
P = 0.3448
Q = 1 – P
Q = 1 – 0.3448
Q = 0.6552
Test statistic: -
$|Z_{ca}| = (p1-p2) / \{PQ[(1/n1)+(1/n2)]\}$
~N(0,1)
Now, the value of P is given by
= (0.5068 –0.2105) / [0.22591(0.0136+0.0175)]
= 0.2963 / 0.00702
$|Z_{cal}|$ = 42.2079 Here, a = 0.05
Z_{tab} = 1.64 $Z_{cal} > Z_{tab}$
42.2079 > 1.64
Hence we reject H_0
Conclusion: - There is significant difference between two Proportion.

Proportion Z test

To obtain the proportion of tablet consumer in urban and rural

Let,
X_1 =The tablet consumers selected from the Urban = 37
X_2 =The tablet consumers selected from the Rural = 12
Given,

Proportion Z test

Eye sight problem in Semi – Rural and Rural

Let
X_1 =Eye sight problem in Semi - rural = 18
X_2 =Eye sight problem in Rural = 21
Given

n1	119
n2	59

Now, we compute the sample proportion

p1= X1/n1	18/119	0.1512
p2= X2/n2	21/59	0.3559

We wish to test
H_0 : p1=p2
i.e. Eye sight problem in Semi- rural and Rural are equal
H_1 : p1<p2
i.e. Eye sight problem in Semi - rural is less that Rural
Under H_0 The test statistic is
P= (p1n1+p2n2) / (n1+n2)
=17.9928 + 20.9981/178
= 0.21905
Q = 1 – P
= 1 – 0.21905
=0.78095
Test statistic: -
$|Z_{cal}| = (p1-p2) / \{PQ[(1/n1)+(1/n2)]\}$ ~N(0,1)
Now, the value of P is given by
= (0.1512 – 0.3559) / [0.1710(0.0084+0.0169)]
= -0.2047/0.004326
$|Z_{cal}|$ = 47.3185
Here, a= 0.05
Z_{tab} = 1.64
$Z_{cal} > Z_{tab}$
47.3185 > 1.64
Hence we reject H_0
Conclusion: - Proportion of Eye sight problem among the residents of old age home is less in Semi – rural than Rural.

To obtain the proportion of Male and Female proportion in Urban old age homes

Let,
X_1 =Female in Urban = 34
X_2 =Males in Urban = 30
Given

n1	93
n2	85

Now, we compute the sample proportion.

p1= X1/n1	34/93	0.3655
p2= X2/n2	30/85	0.3529

We wish to test
H_0 : p1=p2
i.e. Male and Female are equal in Urban
H_1 : p1<p2
i.e. Male are less than Female in Urban
Under H_0 The statistic is
P= (p1n1+p2n2) / (n1+n2)
=33.915+29.9965 / 178
= 0.3594
Q = 1 – P = 1 – 0.3594
=0.6406
Test statistic: -
$|Z_{cal}| = (p1-p2) / \{PQ[(1/n1)+(1/n2)]\}$ ~N(0,1)
Now, the value of P is given by
= 0.3655 – 0.3529/0.2302(0.0224)
= 0.0126/0.00515
$|Z_{cal}|$ = 2.4466
Here, a= 0.05
Z_{tab} = 1.64
$Z_{cal} > Z_{tab}$
2.4466 > 1.64

Hence we reject H_0

Conclusion: - Proportion of Male population is less than Female population in Urban old age homes

If a proportion test shows that there are significant fewer males than females in Urban old age home, conclusion could be Gender population in urban old age are skewed towards females.

FINDINGS

- The analysis does not indicate a significant relation between Health and Sex.
- It has the strong relation between the two factors of staying at Old Age home and the reasons they both are independent.
- The study satisfies the objective that there are more tablet seekers in Urban Old Age home as compared to rural Old Age home .
- The study satisfies the objective that the Eye Sight problem among the residents of Semi – Rural Old Age home is less that than the Rural Old Age homes.
- Analysis indicates that the Females are more than Males in Urban Old Age homes.
- Finding suggest that the relation between age and disease is correlated, it shows that as the Age increases the disease may tend to increase.
- Gender wise classification of respondents exposed that majority are female compared to the opposite gender.
- According to the analysis of data it shows the ratio of old people at old age home are more in Urban > Rural > Semi-Rural
- Analysis says that the literacy rate in Old Age home is more.
- Even the literacy rate is high but the service ratio is less as compared to the literacy rate of Old Age home.
- The Analysis says that the most frequent reason of people staying at old age home is that their Son don't take responsibility of them.
- According to the graph of Health condition we can conclude that the Check ups and the visit of the Doctors is regular. Because most of the people are in good condition.
- Eye sight problems in old people are faced by lot according to data.

SUGGESTION'S

1) HEALTH: The old age home should have a medical facility on- site or have easy access to medical services. The staff should be trained in basic first aid and emergency response procedures.

2) INFRASTRUCTURE: The infrastructure of the old age home should be designed to meet the needs of older adults. This includes accessible facilities, such as ramps, handrails, and grab bars, to promote mobility and prevent falls. The home should have adequate lighting, ventilation, and comfortable living spaces.

3) SECURITY: The old age home should have a robust security system in place to ensure the safety and well-being of residents. This includes physical security measures, such as surveillance cameras and secure entrances, as well as staff training to prevent**.**

4) PHYSICAL NEEDS: The old age home provide the physical needs of the residents, including nutritious meals, clean water, and adequate hygiene facilities. The home should also have facilities for physical exercise and activities.

5) MANAGEMENT OF FACILITIES: The management of old age home should be efficient and effective. This includes proper maintenance of the facilities, regular monitoring of the health and well-being of the residents, and clear communication between the staff, 3.

8. CONCLUSION

Proportion Z test

The study looked at old age homes in cities, semi-rural areas, and rural areas to see how they differ in facilities and healthcare. Here's what we found:

City Old Age Homes: These places have modern facilities and more staff, but they are expensive for residents.

Semi-Rural Old Age Homes: These homes feel more like a home, with personal care. However, they might not have easy access to specialized medical help.

Rural Old Age Homes: These places have a strong sense of community and family involvement, but they might have fewer resources and services.

In simple words, where elderly people live makes a big difference. City homes are modern but costly, semi-rural homes are more personal but lack medical help, and rural homes have a strong community but fewer resources. It's important to consider what an elderly person needs and prefers when choosing a home for them. Also, these homes need support and resources to provide good care for the elderly residents.

9. LIMITATIONS

- There is limitation of sample size and spectrum of respondents.
- Samples taken may be biased.
- The study is limited to old age homes only.

10. REFERENCES

[1] Johnson, Shanthi & Madan, Swati & Vo, Jade & Pottkett, Ansa. (2018). A Qualitative Analysis of the Emergence of Long Term Care (Old Age Home) Sector for Seniors Care in India: Urgent Call for Quality and Care Standards. Ageing International. 43. 1-10. 10.1007/s12126-017-9302-x.

[2] Janbandhu PV, Phad SB, Bansod DW (2022) Lifestyle changes and perception of elderly: A study of the old age homes in Pune city, India. Int J Geriatr Gerontol 6: 138. DOI: https://doi.org/10.29011/2577-0748.100038

[3] Janbandhu PV, and Bansod DW. (2022). A Study on social and psychological conditions of elderly living in old age homes of Pune city in Maharashtra. (Unpublished doctoral dissertation). International Institute for Population Sciences, Mumbai, India.

[4] Sudha S, Suchindran C, Mutran EJ, Rajan SI, Sarma PS. Marital status, family ties, and self-rated health among elders in South India. J Cross Cult Gerontol. 2006 Sep-Dec;21(3-4):103-20. doi: 10.1007/s10823-006-9027-x. PMID: 17242992.

[5] Rajkumari G (2021) A Study of Elderly Living in Old Age Homes in Manipur, India. J Gerontol Geriatr Res. 10: 546

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 41 - 47, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1008

Framework with Enhanced Routing algorithm for Mobile Sensor Network

Hemalatha S
Research Scholar
Bharathidasan University
Tiruchirappalli – 24, Tamilnadu, India

Dr. E. George Dharma Prakash Raj
Associate Professor
School of Computer Science, Engineering and Applications
Bharathidasan University
Tiruchirappalli – 23, Tamilnadu, India

Abstract: Mobile sensor network is subgroup of wireless mobile sensor networks, in which each node are mobile in nature. In heterogeneous sensor network, the network mobility is unpredictable. Routing in mobile sensor network is major challenging with various real time applications. Sensor nodes are mobile in nature, data transmission and data reception process incur packet loss, delay on delivery, packet delivery ratio and energy depletion. To overcome this problem Cluster based routing technique introduced and simulated to acquire. This paper explains about the clustering based routing algorithms in Mobile Sensor Networks that have been proposed earlier by the authors in different journals and further compares the packet delivery rate, energy consumption, latency delay and packet loss of the network based on time interval. This paper analyses the result based on existing routing algorithms like Enhanced perimeter forwarding algorithm (EPFA), partial partitioned Greedy Algorithm (PPGA), Enhanced Efficient routing algorithm (EERA) Mobile Wireless Sensor Networks. Above routing algorithms to be analyzed with set of sensor nodes, simulation results shows an ultimate solution to Mobile Sensor Networks

Keywords – Mobile sensor network, Clustering, Mobility, Energy Efficiency

1. INTRODUCTION:

Mobile sensor network is framed with sensor nodes with different speed of mobility. Mobility factor is changed from to time for a need of a virtual applications. Nodes deployed in randomly, with limited computational capabilities. MSN networks since these networks usually operated in only partially unknown and unpredictable environments with dynamic network structure. Connectivity control, formation control, rendezvous, and coverage control are the usual parameters for mobile sensor network.

This paper has been orchestrated as follows. Section I gives introduction on Mobile Sensor Networks. Section II describes the cluster based routing technique. Section III describes literature study on relevant routing algorithms. Section IV explores the performance analysis of routing algorithms. Section V specifies the conclusion.

2.LITERARURE REVIEW:

According to author routing a mobile nodes are deployed random manner, in these nodes are adapting the frequently changing the climatically conditions. The node coverage and node mobility [1] is the highly challenging to perform routing functions. Intra and inter cluster based routing in Large area WSN, where are analyzed with node mobility and energy factor.

An Optimal Routing with Node Prediction (ORNC) algorithm [2] proposed by the author for the mobile node to improve the routing on mobile network. ORNC algorithm routing is carried out by Node classification, calculating the trust factor, and calculating fitness functions.. Whereas the node classification is based on node initial energy and memory it categorized 4 different types. Based on the Boolean operators AND/OR operators we calculate the trust values. These trust value gives the best path from the n number of from source to destination, so the routing ids performed in an efficient manner.

An inter cluster based routing algorithm is proposed by the author, in the mobile sensors energy consider as major research factor. This proposed work follows certain cluster head selection policy [3]. The sink node is monitored and other parameters are materialized with packet delivery ratio, latency and energy. In CFMS algorithm[3] proposed with mobile nodes, to perform routing on intra and inter cluster based sensor network. The cluster are formed with Directed acyclic graph in intra cluster. In inter cluster node are formulated by the tree structure so that data aggregation is performed in better way, co that delay is reduced substantially. Load balancing is applied while routing so that routing over head is reduced, ultimately the packet delivery ratio is improved.

The propose work Elham Ghorbanifar focus the energy factor of the sensor network[4], whereas concentrate the end to end delay and network life time, the algorithm is executed in the intra cluster based sensor network.

Whale optimization algorithm and tunicate swarm algorithm (TSA) are combined to form whale tunicate swarm algorithm (WTSA)[5], is deal with the multipath routing in Sensor network. Both intra and inter cluster based distance, node degree and energy are the parameters are taken into evaluation of WTSA. A proposed methodology improves the throughput reduces delay.

Then black widow optimized routing protocol[6] find the optimal path to transmit the data with received signal strength indicator (RSSI).In order to create a manual cluster on the network, this task is suggested with a reference node, which lowers the energy used by the cluster head compared to

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 41 - 47, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1008

alternative models. By taking into account congestion and retransmission, which cause the node to lose energy, the reference node determines how many paths are available for the mobile node to transfer the data. This model improves the packet delivery ratio and reduce delay time.
In Modified Ring Routing protocol performs routing in inter cluster based mobile sensor network, here the CH selection based on the residual energy. Each node is connected to its neighbor node with one hop neighborhood list. The sink node is the neighbor of the relay node in ring the corresponding node in the circle performs the routing. Modified Ring[7] Routing Protocol for Mobile Sinks in a Dynamic Sensor Network in

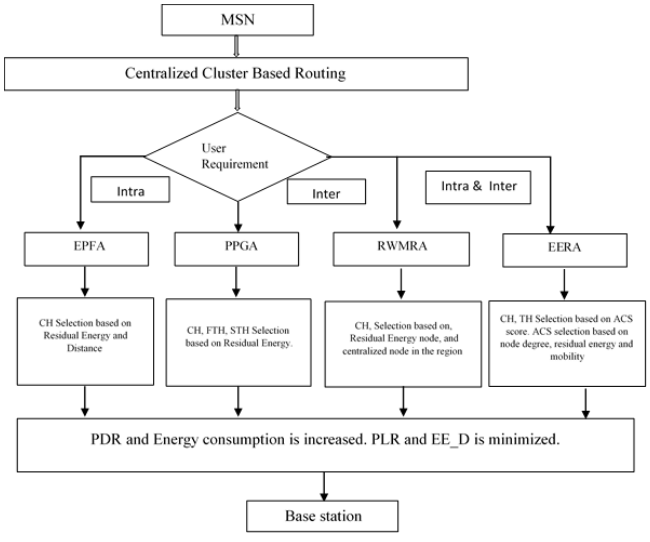
Smart Monitoring Applications.

2. ROUTING ON MOBILE SENSOR NETWORKS:

In mobile sensor network routing is get into clustering of sensor region segmented into n number of overlapping groups. Each group is governed by a node referred to as Cluster Head (CH) with all the nodes within a cluster referred to as Cluster Members (CM). Cluster Head (CH) collects the data from all the member nodes and aggregates the data collected from the member nodes and sends it to the Base Station.

Cluster based routing used to is used to reduce the energy consumption among the sensor nodes thereby increasing the lifetime of the nodes. In Mobile Sensor Networks various cluster based routing algorithms have been proposed. The ultimate aim of all these proposals is to improve the packet deliver ratio, reduce energy consumption, delivery delay and packet loss of network.

From the papers EPFA, PPGA and EERA is analyzed in various routing parameters in which the node density of 120 nodes in dynamic environment, in which various results of the 4 contributions given the ultimate result of the objective is well defined way.



3. RELATED WORK

3.1. Enhanced Perimeter Forwarding Algorithm: EPFA

Enhanced perimeter forwarding algorithm proposed to a cluster region with restricted number of sensor nodes. EPFA [2] is a Cluster based routing Algorithm which improves throughput and reduce energy consumption, packet delay, packet loss Mobile Sensor Networks. EPFA algorithm is designed for the nodes which change the mobility from time to time. It could to restrict to region with radial distance r. The algorithm checks for node existence within the region, data transmission taken carryout by either single hop or may be multi hop fashion. Node degree is determined by one hop neighbors within a transmission range. The distance between the node and the base station, the node's mobility and the energy used in data transmission and reception are all calculated. The node with maximum residual energy is selected as Cluster Head. Once the Cluster Head (CH) is selected the remaining nodes sends a Request to the CH to become a Member to form the Cluster. The nodes which receive an Acceptance message from CH becomes the Cluster Member to form a Cluster. The weight of each node is calculated and the node with maximum weight is considered as the Cluster Head for the regular time interval seconds. As the Cluster is formed the Cluster Members sends data to the Cluster Head which are aggregated and forwarded to the Base Station. Throughput is increased in a Cluster region, when it is compared with CCM_GR Algorithm.

ALGORITHM 1 : SELECTION OF CH

1. Nodes n1, n2, n3..... $\propto$ are the nodes within the network range and forms as a cluster.
2. The density of the network is defined as θ and $\theta = \frac{S^2}{N}$, where S^2 is the area of the Network, N is the number of nodes.
3. All member nodes communicates the information about its current energy and locations to its CH in each
4. The energy value of each node can be calculated and the highest energy node will be selected as a CH.
It acts as a superior in all the clusters.

If N = n1,n2,n3.... $\propto$.} then CH >n1,n2,n3..... $\propto$.

ALGORITHM 2 : SUB-REGION SELECTION AND FORWARDING ALGORITHM

1. Input Neighhop_List, WaitList
2. Add component in WaitList to Neighhop_List
3. For i<NeighHop_List length do
4. Get the area of next hop j
5. if dis_to_nexthop > T r
6. Store j into WaitList, n=0
7. Else
8. Call DivedeSubForwardRegion(List NeighHop_list)
9. End if
10. Call SelectNextHop(List NeighHop_subregion)
11. End for
12. Judge n whether still in NeighHop_list, if not n=0
13. Output n
14. Capacity DivdeSubForwardRegion(List NeighHop_list)
15. Cross the whole NeighHop list
16. Partition these NeighHop into various districts with equivalent region
17. Calculate the normal vitality of each sub-locale ir E
18. Select the greatest the ir E

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 41 - 47, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1008

19. Return the relating rundown of NeighHop_Subregion with greatest ir E
20. End Function

3.2 Partial partitioned Greedy Algorithm: PPGA

Partial-partitioned Greedy Algorithm (PPGA), for mobile sensor networks that consist of frequently moving sensors. PPGA algorithm partitions the sensing field into radial distance r. Each have its mobile sensor nodes with regular time interval. The protocol uses the location information of sensors and the base station to assign a cost function to each sensor node, which is close to the Euclidean length of a sensor node's shortest path to the base station. A packet is forwarded to the base station using greedy forwarding whenever possible. When a packet reaches sensor nodes near local minimums, where greedy forwarding will be impossible after a number of hops, the packet is forwarded following the high-cost to low-cost rule. The node with maximum residual energy is selected as the Cluster Head. The next lower nodes is get elected as transmission Head, and next lower energy node is elected as subordinate cluster head. The other nodes sends a request to the respective Cluster Head and the nodes become Cluster Members as it receives the Acknowledgement from the Cluster Head. Head's to clusters are formed for regular time interval t. As the nodes are in mobility, the nodes move from one cluster to the other, thereby joining the new Cluster. New Cluster Heads are re-elected whenever the Cluster Head leaves the cluster. Similarity the remaining heads get into re-elected for the each time interval t.

The algorithm shows a better performance when compared with EPFA with improves packet delivery ratio, reduces the delay, packet loss and energy. The Energy Consumed during mobility is also less compared to EPFA. PPGA is a Cluster based routing Algorithm which maximize packet delivery ratio, reduce latency delay, packet loss and energy consumption for the Mobile Sensor Networks.

Cluster Head Selection Algorithm ( )

Start_CH_SelectionAlgorithm()
1. singclusweight=w1×E+w2×I
2. isclusterhead=1
3. maxweight=singclusweight
4. timer=1/singclusweight
5. if (timer<0)
6. CH_Announcment(myID,singclusweight)

ReceiveAnnouncment (SNid, weight)

1. If (isclusterhead==1){
2. If(ownweight<weight){
3. isclusterhead=0
4. MyCH=SN id
5. maxweight=weight}}
6. else if(isclusterhead==0){
7. If(maxweight<weight){
8. MyCH=SN id
9. maxweight=weight}}

3.3 Random Waypoint Mobility Routing Algorithm: RWMRA

In RWMRA proposed to high mobility applications, in which the sensor nodes are deployed randomly. Sensor regions are divide into rectangle region. Region head is get based on the residual energy of the nodes in each region. The highest residual energy node is elected as cluster head. Distance between the nodes to its cluster head is for data forwarding. Node degree is calculated by either clock wise or anti-clock wise from the location of the node in a region.

Algorithm 1: Region Creation

1. Randomly deploy sensor nodes.
2. Partition the sensor field into D regions.
3. Select center point and corner points
4. Create region head (SSN) based on selected region
5. Draw circle based on the selected points using radius r_{comm}
6. For each node in Network
7. Get node location n(x,y)
8. For each region in Network
9. Get the location of ssn(x,y)
10. Compute the dist(x,y) = n(x,y) - ssn(x,y)
11. If (dist (x,y) ≤ r_{comm})
12. Set node n coverage = Yes
13. break;
14. Else Set node n coverage = No
15. EndIf
16. EndFor
17. EndFor

Algorithm 2: Announce Mobile Sink Location

1. if distance(dest(x, y),C(xc, yc)) ≤ r1 then
2. send a destinfo_MSG in the opposite direction of the network center
3. else if distance(dest(x, y),C(xc, yc)) ≥ rk then
4. send a destinfo_MSG towards the network center
5. else
6. send two destinfo_MSG; one towards the network center and one in the opposite direction of the network center.
7. for each node receiving destinfo_MSG do
8. if Node i : type == normal sensor node then
9. Ignore the destinfo_MSG
10. else
11. if (Previous destination location != Current destination location) then
12. Save the destination location information, share destinfo_MSG with its clockwise and anti-clockwise neighbouring VS nodes, and other region.
13. stop sharing destinfo_MSG
14. else
15. send the destinfo_MSG to its next neighbouring VS node
16. Ignore the destinfo_MSG

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 41 - 47, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1008

Algorithm 3: Data Transmission

1. compute the distance between X and Y at interval time t

$$dis^{X,Y}(t) = \| X_{loc}(t) - Y_{loc}(t) \|$$

2. if dis X,Y(t) ≤ r_{comm}
(Node X is within the transmission range of Node Y) then
3. NS = Get node status ()
4. If (NS == Coverage)
5. Transmit Data
6. Else If (NS == Non Coverage)
7. Exit routing
8. Else If (NS ==Move)
9. Wait until Node movement
10. Else (Node X is not in the transmission range of Node Y)

// Use Node Y nearest SSN to Compute

11. D1= || SSNloc (t) – Yloc (t) ||
12. D2 = || SSNloc (t) – Yloc (t+1) ||
13. DC = ||D2 – D1|| / D1
14. θ = || θ (t+1) – θ|| / 2π
15. mob = α. DC + β. θ (Where α and β ϵ [0, 1])
16. Find the path based on mob
17. Check all nodes (path) mobility status
18. Update the pause time of node
19. Transmit data

3.4 Enhanced Efficient Routing Algorithm: EERA

EERA is propose for the mobile sensor network it will applied with high mobility environment. In this algorithm is calculates cumulative credit score based on sensor node is energy, mobility and neighborhood node. Based ACS is each region the Cluster head is selected. The next ACS node get elected as the TH.

Algorithm 1: Cluster formation ( )

N nodes randomly deployed in L*L region
Sensor nodes segmented into n clusters with radial factor r
Min_energy = 5 jules
For i = 1 to n do
SN_energy(i) = 30 jules
SN_mobility(i)=

$$\frac{SN\_current\_position(i) - SN\_previous\_position(i)}{time\ interval\ t}$$

Endfor

Procedure Neighbour(t,n,r,Θ)

Let t=200 rounds
For i = 1 to t do
For j= 1 to n do
For k= i+1 to n do
If (SN_neighbour[i]&&SN_neighbour[j][k] ≠ NULL) then
count_neighbour+=1
Endif
Endfor
Endfor
Endfor
Return(count_neighbour)

Procedure Energy(SN, n)

For i=1 to n-1 do
For j=i+1 to n do
if (SN_energy[i]>SN_energy[j]) then
temp_node = SN_energy[i]
SN_energy[i] = SN_energy[j]
SN_energy[j] = temp_node

Endif
Endfor
Endfor
For i=1 to n do // CH and TH selection for n clusters

return(SN_energy[i])

Procedure Mobility (SN, n, t)

For i= 1 to n do
SN_mobility[i]

$$= \frac{SN\_current\_position[i] - SN\_previous\_position[i]}{time\ interval\ t}$$

For i=1 to n-1 do
For j=i+1 to n do

if (SN_energy[i] < SN_energy[j]) then

temp_node = SN_energy[i]
SN_energy[i] = SN_energy[j]
SN_energy[j] = temp_node

Endif
Endfor j
Endfor i
For i=1 to n do
Return(SN_mobility[i])

Algorithm 2: Selection of Cluster head and Transmission head (N, SN_energy, SN_mobility)

1. For i= 1 to N do // for all nodes
2. Assign CH=TH=NULL
3. If (SN_energy(i) > Min_energy and SN_mobility(i)< 20) then
4. CH=SN_id(i)
5. Else
6. For i=1 to n-1 do
7. For j=i+1 to n do
8. if [SN_energy(i)>SN_energy(j)] then
9. temp_node= SN_energy(i)

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 41 - 47, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1008

10. SN_energy(i)= SN_energy(j)
11. SN_energy(j)= temp_node
12. Endif
13. Endfor
14. Endfor
15. Endif
16. For i=1 to n do // CH and TH selection for n clusters
17. CH= SN_id(i)
18. TH= SN_id(i+1)
19. Endfor
20. Endfor

Algorithm 3: Path Announcement (SourceSN_id, SinkSN_id, CH,TH,SN_energy, Θ)

1. t = 200 round
2.Path = null, Tot_time=t // Total time for simulation
3. Do
4.For k = 1 to Tot_time // Time interval for N nodes
5.If (SinkSN_id ≠ null) then // check sink node is availability
6.For i=1 to n do //all nodes in each cluster
7.If (SinkSN_id == neighbour(CH[i])) then

8. TH=path(x,y, with its clockwise and anti-clockwise to the cluster)

9. Else if(SinkSN_id == neighbour(CH[i+1])) then

10. TH=path(x,y, with its clockwise and anti-clockwise to the cluster)

11. Endif
12.Endif
13.Endfor
14.While (Tot_time >t)

Algorithm 4: Procedure pathMaintanence (SSN_id, Sink _id, CH, TH)

1. Path = null

2. Tot_time = t // Total time for simulation

3. Do

4. If (CH || TH == NULL) then

// due to mobility check CH/TH availability

5. Call Procedure Selection_CH/TH( )
6. While (Tot_time > 60)

4. PERFORMANCE ANALYSIS OF ROUTING ALGORITHMS:

The performance of Cluster based routing Algorithms are analyzed. Overall throughput of the Algorithms are analyzed along with the packet delivery ratio, Energy, packet loss and end to end time delay during predefined time intervals. The performance of the various algorithms are analyzed in Matlab 7.0.

The overall Energy Consumption of all the Clusters and the Energy Consumption based on the Mobility of the Nodes are also analyzed.

Table 1 shows the overall Packet delivery ratio of the clusters during Round 200 Rounds at sensor regions.

Rounds	Packet Delivery Ratio			
	EPFA	PPGA	RWMRA	EERA
1	1	1	1	1
50	0.702	0.674	0.692	0.702
100	0.716	0.688	0.706	0.716
150	0.704	0.714	0.72	0.704
200	0.74	0.678	0.7106	0.74

Table 1 : Packet Delivery Ratio Vs No of Rounds at Regions

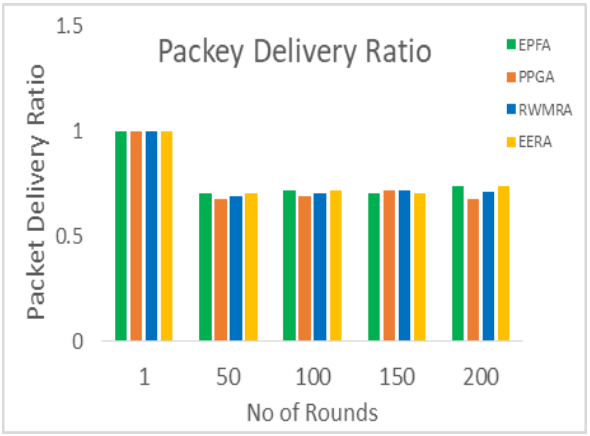


Fig. 1: Packet Delivery ration Vs Time interval at Regions

Fig. 1 shows the Packet Delivery Ratio of the Clusters is more reduced at 50 rounds in EPFA, PPGA, RWMRA and EERA algorithm.

The packet Delivery Ratio of Clusters is almost same in EPFA and EERA in 100 rounds.

PPGA and EERA give the same packet delivery ratio at 150 rounds when it compared with EPFA and RWMRA.

Rounds	End to End time Delay			
	EPFA	PPGA	RWMRA	EERA
1	0.1	0.1071	0.0787	0.0947
50	0.152	0.131	0.1018	0.1085
100	0.181	0.16	0.1177	0.1104
150	0.19	0.213	0.14001	0.1239
200	0.21	0.256	0.1892	0.153

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 41 - 47, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1008

Table 3 shows the overall End to End Delay of the routing during 1, 50, 100,150, 200 rounds atregion.

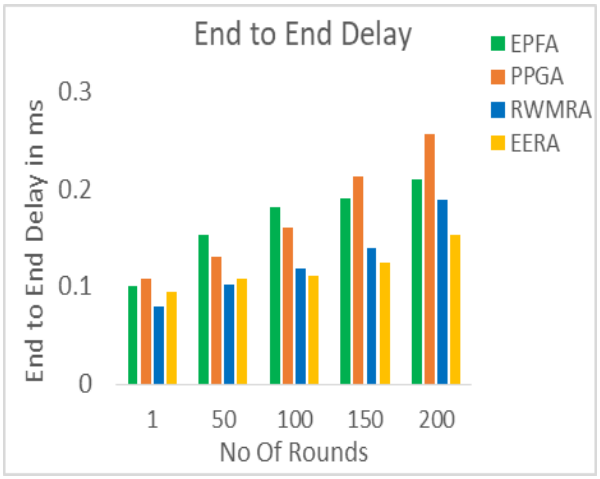


Fig. 2: End to End Delay Vs Time interval at Regions

Fig 2 represents that the EERA algorithm shows reduced End to End time delay of routing than EPFA, PPGA and RWMRA algorithms at 1, 50,100,150,200 rounds.
End to end delay time of the PPGA is increased due to mobility of the sensor node.

Rounds	Packet Loss Ratio			
	EPFA	PPGA	RWMRA	EERA
1	0	0	0	0
50	0.21	0.326	0.308	0.298
100	0.285	0.312	0.294	0.284
150	0.324	0.286	0.2621	0.296
200	0.354	0.322	0.1863	0.26

Table 3 shows the overall Packet Loss Ratio ofthe Clusters during 1 round to 200 Rounds at regions.

Fig. 3 represents RWMRA algorithm shows very less packet loss ratio compared with EPFA, PPGA and EERA Algorithms at 200 rounds.
Packet loss is very high in EPFA because, the sensor node moving out of the cluster region, this algorithm proposed for a cluster region (intra cluster), so the PLRis higher than the other 3 algorithms.

Fig. 3 : Packet Loss ratio Vs Time interval at Regions

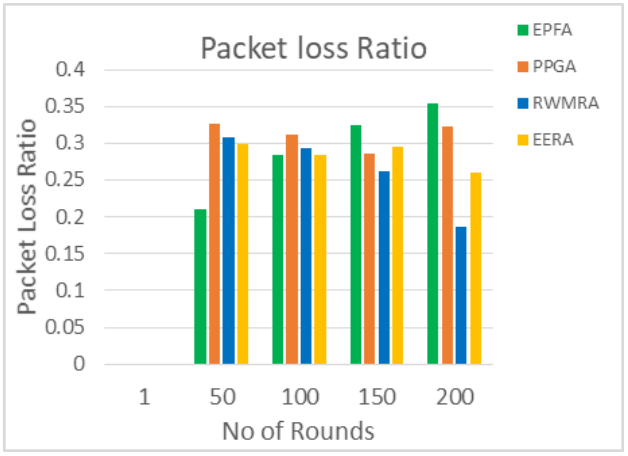


Table 4. Shows the overall Energy consumption of the Clusters during 1 Round to 200 Rounds at regions.

Rounds	Energy Consumption			
	EPFA	PPGA	RWMRA	EERA
1	1.6	1.9575	1.9745	1.9946
50	0.89	1.0638	1.874	1.8988
100	0.68	0.8566	1.7443	1.7981
150	0.51	0.7256	1.6123	1.7002
200	0.43	0.6586	1.477	1.6052

Table 4: Energy consumption Vs Time interval at Regions

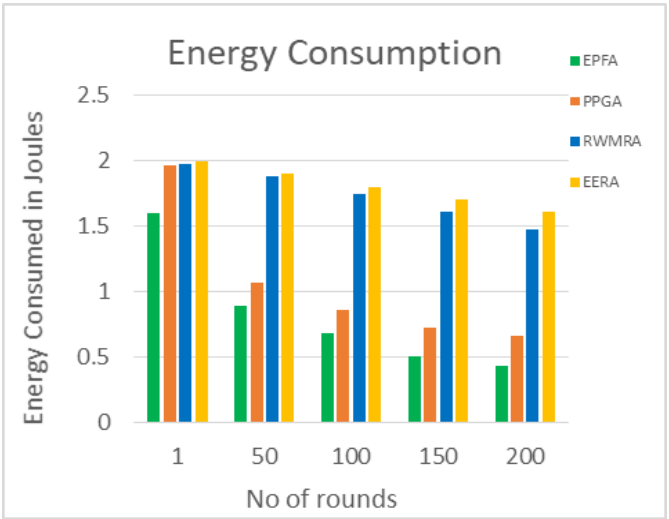


Fig. 4 : Packet Loss ratio Vs Time interval at Regions

Fig. 4 represents that the EPFA Algorithm shows reduced Energy Consumption than in PPGA, RWMRA and EERA Algorithms at 200 Rounds.
The Energy Consumption is high in EPFA, PPGA,

RWMRA and EERA Algorithm at first round.
The Number of Clusters are increased based on the Density of nodes in PPGA, RWMRA and EERA. The Energy Consumption is high in EERA when compared to RWMRA, PPGA and EPFA algorithms.

5. RESEARCH HIGH LIGHT:

Two distinct routing myths propose the cluster-based routing method. Routing based on Intra and inter-cluster factors. To solve the routing inconsistencies in mobile sensor networks, four distinct routing algorithms have been developed.
These algorithms are all used with the Mat Lab tool, and the outcomes are examined.
An enhanced perimeter forwarding algorithm (EPFA) is suggested for intra-cluster routing, and its effectiveness is calculated using the four metrics. Metrics such as end-to-end delay, energy consumption, packet loss ratio, and packet delivery ratio.

Table 5. Overall Performance of theresearch Framework

S.No	Algorithms	Type of Cluster	Performance Evolution
1	EPFA	Intra	PDR, EED, EC
2.	PPGA	Inter	PDR, EED, EC
3	RWMRA	Intra/Inter	PLR, EC
4.	EERA	Intra/Inter	PDR, EED, PLR

6. CONCLUSION:

We have presented a routing algorithms for mobile sensor networks, which is based on the election of a Cluster head and the construction of network with the CH and TH with multiple criteria to obtain the same. On the basis of local information about sensor node, Cluster Head elections are performed in order to restructure the clusters so that message delivery rates could be improved.
The ultimate purpose of those algorithms is to reduce theEnergy Consumption, end to end delay, packet loss and maximize the through put. The results of the existing algorithms are confined to give better result. The performances of the algorithms EPFA, PPGA, RWMRA and EERA discussed in this paper. Here Cluster based routing algorithm reduces reduce the EnergyConsumption, end to end delay, packet loss and maximize the through put.
EERA algorithm is more efficient in terms of reduce the Energy Consumption, end to end delay, packet loss and maximize the through put than EPFA, PPGA and RWMRA.

REFERENCES:

[1] Arpana Mishra, Rashmi Priyadarshini, Shilpa Choudhary, R M Mehra, Qualitative Analysis of Intra-Class and Inter-Class Clustering Routing and Clusterization in Wireless Sensor Network, EVERGREEN Joint Journal of Novel Carbon Resource Sciences & Green Asia Strategy, Vol. 08, Issue 02, pp358-373, June 2021.

[2] Anjaneya Tripathia, Vamsi Kiran Mekathotia, Isha Waghuldea, Khushali Patela and Nithya Balasubramaniana, Neural Network aided Optimal Routing with Node Classification for Adhoc Wireless Network, WCNC Workshop on Computer Networks & Communications, May 01, 2021.

[3] Gokou Hervé Fabrice Diédié1 , Koigny Fabrice Kouassi2 and Tchimou N'Takpé, Cfms: A Cluster Based converge cast Framework For Dense Multi-Sink Wireless Sensor Networks, International Journal of Wireless & Mobile Networks (IJWMN), Vol.15, No.1, February 2023.

[4] Elham Ghorbani Dehkordi, Cluster based routing method using mobile sinks in wireless sensor network, International Journal of Electronics, Jan 2022.

[5] Jatinder Pal Singh and Anuj Kumar Gupta, Energy Aware Cluster Head Selection and Multipath Routing Using Whale-based Tunicate Swarm Algorithm (WTSA) for Wireless Sensor Network, New Review of Information Networking, Front Tailor,| 29 Mar - 2022.

[6] S. Arunkumar ,S. Aravind Kumar, S. Priya,Vinoth kumar P, An Energy-Efficient Routing of Mobile Sensor Nodes using Black Widow Optimized Routing Protocol,YMER, International Open Access Journal VOLUME 21 : ISSUE 11 (Nov) – 2022.

[7] Seli Mohapatra, Prafulla Kumar Behera, Prabodh Kumar Sahoo, Manoj Kumar Ojha, Modified Ring Routing Protocol for Mobile Sinks in a Dynamic Sensor Network in Smart Monitoring Applications, Electronics 12(2):281, January 2023.

[8] Jogendra Kumar, Annapurna Singh, H.S. hadauria, Congestion control load balancing adaptive routing protocols for random waypoint model in mobile ad-hoc networks, Journal of Ambient Intelligence and Humanized Computing ,Springer-Verlag GmbH Germany, part of Springer Nature, April 2020.

[9] Kamalinder kaur, sandeep kang, Game Theory Based Efficient Routing Protocol In Mobile Wireless Sensor Network Using Advanced Learning Mechanism, Journal of Theoretical and Applied Information Technology, 15th July 2023.

[10] H. Esmaeili, B. M. Bidgoli, and V. Hakami, CMML: Combined metaheuristic-machine learning for adaptable routing in clustered wireless sensor networks, Applied Soft Computing, vol. 118, 2022.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 48 - 71, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1009

Green Computing: A Comprehensive Review of Sustainable IT Practices

Mohammed Swaleh

ICTS Department

Technical University of Mombasa

Kenya

Prof. Franklin Wabwoba

Department of Information Technology

Kibabii University, Kenya

Abstract

The objective of the study titled "Sustainable IT Services: Assessing the Impact of Green Computing Practices" is to analyze existing literature on green computing and its effects on sustainable IT services. The aim is to identify crucial concerns and strategic areas that can enhance customer value, business value, and societal value. The paper examines sustainable IT techniques, including power management, virtualization, enhancement of cooling technology, recycling, electronic waste disposal, and optimization of the IT infrastructure to fulfill sustainability criteria. According to the authors, sustainable IT strategies are expanding the concept of sustainability beyond energy consumption and product considerations. This expansion requires the restructuring of both the IT organization as well as the overall practices in order to fully achieve the strategic advantages of green computing.

In essence, this paper provides a significant reference for institutions who wish to understand and adopt sustainable IT practices, particularly in the realm of green computing.

Keywords: Green Computing, Sustainable IT, ICT sustainability, Sustainable Computing, Green IT, Green ICT, Green Technology

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 48 - 71, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1009

INTRODUCTION:

1.1 Background of Study

Green computing, sometimes referred to as green IT or ICT sustainability, is the approach to the design, production, utilization, and disposal of computers and other technological devices with the aim of reducing their negative impact on the environment. The primary objectives of green computing are to optimize energy efficiency throughout the lifespan of the product, and enhance the recyclability or biodegradability of obsolete products and waste generated in factories [Gichuki et al., 2018].

Green computing is crucial for all types of systems, ranging from portable devices to extensive data centers. The environmental impact of the IT sector is substantial, representing 5-9% of global electricity consumption and over 2% of total emissions. Hence, in order to maintain competitiveness, data centers and telecommunications must enhance their energy efficiencyand increase their reliance on renewable energy sources[(*Between 10 and 20% of Electricity Consumption From the ICT* Sector in 2030?*, 2018)]].

There are multiple strategies that both corporations and individuals can adopt to enhance sustainability in the field of information technology. These strategies encompass server collocation, adherence to data center best practices, migration to the cloud, and utilization of cutting-edge IT technologies [(*Sustainable IT: Ways IT Can Be More green and why it matters*, 2019)].

Green computing includes the utilization of energy-efficient central processing units (CPUs), servers, peripherals, power systems, and other IT equipment. Additionally, it emphasizes the reduction of resource consumption and the efficient management of electronic waste[(*What Is Green Computing? | Definition From TechTarget*, 2023)].

Another sustainable IT strategy is server virtualization, a method that enables numerous virtual machines to operate on a single physical server. This approach effectively reduces energy

usage and carbon footprint[(Expert Panel, 2023)].

Green computing offers numerous advantages for both the environment and enterprises. According to Zaveria(2023), adopting energy-saving measures can yield several benefits, including cost savings, decreased greenhouse gas emissions, enhanced health and working conditions, and improved brand reputation.

Green computing is an all-encompassing method for implementing sustainable practices in the field of information technology. It entails the use of several tactics and approaches to minimize the environmental consequences of IT operations. Green computing, as defined by TechTarget in 2023, refers to the practice of using computer resources in an environmentally friendly and sustainable manner.

1.2 Problem Statement

Currently, the majority of operations in Kenya have been digitized, resulting in a significant dependence on the ICT hosting systems that support them. Consequently, there has been an acquisition of more hardware to meet the increasing demand for further digitalization of services. Nevertheless, the absence of sustainable technology for the implementation and deployment of ICT in both commercial and governmental organizations has led to ineffective utilization of ICT resources, heightened energy consumption, and various technical difficulties. Due to these issues, institutions increasingly mandate the implementation of an environmentally sustainable IT deployment strategy.

Sustainable IT techniques, including power management, virtualization, enhanced cooling technology, recycling, electronic waste disposal, and optimization of the IT infrastructure, have the potential to enable intelligent management, facilitate scalability, and promote efficient resource utilization. Institutions can enhance resource allocation efficiency and system scalability by incorporating sustainable IT practices into their service platforms. This can aid in guaranteeing smooth

and continuous operations while simultaneously optimizing the overall efficiency of operations (Koratagere, S. et. al (2023). Hence, it is imperative for Kenyan institutions to adopt sustainable IT practices on their platforms to improve efficiency and adaptability.

The main objective of this project is to implement sustainable IT practices, including power management, virtualization, enhanced cooling technology, recycling, electronic waste disposal, and optimization of the IT infrastructure, in order to fulfill sustainability criteria in both commercial and governmental organizations in Kenya.

In accordance with the findings of Gichuki et al. (2018), institutions are facing mounting pressure to adopt more sustainable practices in their utilization of information and communication technology (ICT). In order to achieve its Vision 2030 development roadmap, Kenya must use green ICT practices to enhance both its organizational and environmental sustainability. The study aimed to examine the environmental impact of green ICT and assess the efficacy of green ICT management in Kenyan institutions.

A combination of quantitative and qualitative approaches was employed, together with an explanatory study design. The technique of purposive sampling was employed to choose specific elements for investigation, whereas random sampling was utilized to conduct a survey of sixty-seven (67) higher education institutions (HEI) in Kenya. The results suggest that despite the implementation of energy-efficient technology, green ICT is predominantly causing adverse effects on the environment due to inadequate electricity management by HEI. The presence of incomplete replacement and unfulfilled optimization of green ICT solutions is evident, alongside the launch of new products and re-materialization. Positive indicators would include increased consolidation, reduced printing, and proper disposal of equipment in compliance with government rules. Kenyan institutions encounter many challenges including

limited awareness, uncertain investment returns, restricted collaboration, and insufficient finance. The research recommends that institutions focus on establishing a framework to implement green ICT policy (Samoei et al, 2021).

1.3 Study Objectives

1.3.1 To investigate environmentally-friendly IT practices, such as efficient power usage, virtualization, advancements in cooling technology, recycling, proper disposal of electronic waste, and optimizing the IT infrastructure, in order to fulfill sustainability criteria in institutions.

1.3.2 Offer suggestions for Kenyan institutions to implement.

The primary research questions in green computing involve evaluating the challenges and solutions in various IT sectors, such as IoT, data centers, and telecommunications, to achieve energy efficiency, waste reduction, and the use of renewable energy sources [Butt, S. et. al. (2020].

2.0 Literature Review:

2.1 History and Evolution of Green Computing

Green computing, often referred to as Green Technology or Green IT, gained significant attention in 1992 with the introduction of the Energy Star program by the U.S. Environmental Protection Agency. The algorithm successfully detected consumer gadgets that complied with energy efficiency regulations (Merritt, 2022, Shanthi, D., & Shalini, M. (2022)]. Over time, the notion of green computing developed and expanded to include a wider array of actions with the goal of diminishing the ecological consequences of technology. This encompasses the utilization of lower amounts of energy, the mitigation of waste, and the advocacy of sustainability [Salama M (2020)].

2.2 Sustainable IT Practices

Key components of sustainable IT practices encompass the utilization of energy-efficient technology,

virtualization, and cloud computing. Energy-efficient hardware minimizes energy consumption by optimizing device utilization according to user choices [Energy5 (2023)]. Virtualization is the process of combining numerous physical servers or resources into a single virtual environment. This allows for better use of resources and minimizes power consumption, cooling needs, and physical space requirements [Energy5 (2023)]. Cloud computing enables remote access to computing resources, hence removing the requirement for energy-intensive on-site servers. The source cited is a publication titled "Energy5" from the year 2023[*Energy5*. (2023)]

2.3 Environmental Impact of Data Centers

Data centers are widely known for their substantial energy requirements. These facilities require a continuous and uninterrupted supply of electricity to operate the servers, cooling systems, and networking equipment. Data centers are predicted to consume around 3% of the global energy supply, and this percentage is projected to increase due to the continuous expansion of the digital ecosystem[(Utilities One, 2023)]. The energy consumed by these facilities not only exacerbates climate change but also presents financial burdens for businesses[(Utilities One, 2023)].

2.4 Energy-Efficient Solutions for Data Centers

In order to address the environmental issues presented by data centers, it is essential to examine and implement eco-friendly alternatives. The options prioritize the enhancement of energy efficiency, the reduction of carbon emissions, and the utilization of renewable energy sources [Utilities One, 2023].

Integrating energy-efficient cooling solutions can substantially decrease energy usage in data centers. Data centers can achieve carbon neutrality and realize long-term cost savings by using renewable energy sources. Virtualization and load balancing are effective methods for maximizing server utilization, resource allocation, and energy efficiency [Utilities One, 2023].

2.5 Key Studies and Findings Related to Green Computing

According to a study conducted by Microsoft, their cloud services demonstrate remarkable energy efficiency, surpassing on-premise data centers by a significant margin of 93%. The cloud is a very sustainable option for enterprises, with a carbon efficiency that is 98% higher than that of traditional data centers[(Svistun. O (2023)].

According to the International Energy Agency (IEA), data centers account for around 1-1.5% of the overall global electricity usage. The global data centers consumed a total of 220 to 320 Terawatt hours of electricity in 2021, which accounted for approximately 0.9% to 1.3% of the world's total electricity consumption. Roundy (2023) reports a significant rise in data center energy use, ranging from 10% to 60%, since 2015.

To conclude, green computing is not merely a trendy term, but rather an essential component of our technological advancement. The advancements in green computing offer both environmental and long-term economic benefits [Energy5. (2023)]. Organizations can effectively mitigate their environmental impact, attain cost savings, and enhance their financial performance by implementing green computing solutions [Energy5. (2023)].

3.0 METHODOLOGY

This paper thus assesses the Impact of Green Computing Practicestowardsthe Sustainable IT Servicesthrough desktop study. Same is done through literature review of various papers done in different countries. The literature is collected basis their relevance and their age, narrowing the same to papers within 10 years of age and related to the topic of green computing and Kenyan institutions.

At the end of it the paper aims to come up with recommendation for sustainable IT services in Kenya for institutions to adopt.

4.0 DISCUSSION OF FINDINGS

Below are the discussion and findings of the current trends by HEI in embracing green computing.

4.1.Creating Awareness of Green ICT and Green IT Plan

This is a critical trend and approach that should be applied at the HEI. Learning institutions should be the first to train society to confront and adopt environmental challenges and environmentally good practices in their use of ICT.

According to a study conducted by (Thomson et al., 2015), the real degree of Green ICT adoption and readiness throughout South African institutions appears to be very low. According to Asabere et al. (2016), several green ICT practices are carried out on a very small scale in Ghanaian institutes. According to the literature, the level of green ICT adoption and implementation at institutions in developing countries is relatively low when compared to institutions in wealthy countries.

In Kenya, this is a move that Higher Educational Institutes (HEI) have already taken; for example, Green ICT is currently one of the courses taught at HEI (Samoei et al, 2021). This is supported by a study by Kirui, et. al. (2023), which reports that Some initiatives have been made by Kenyan universities to guarantee sustainable development. For example, to apply greening IT strategies in universities, both public and private universities took part in creating the "Kenya Green University Network (KGUN). This

entails planning how to incorporate green computing techniques into the institution's operations. It also entails promoting green computing to students and faculty in order to encourage involvement.

4.2. Green Computing Practices and Technologies

Green computing, or sustainable computing, is the practice of creating and utilizing computer systems that cause little harm to the environment. This encompasses the development of hardware that consumes little energy, the enhancement of software for optimal performance, and the advocacy for proper disposal of electronic devices [Energy5. (2023)]. Below are few essential green computing practices:

4.2.1. **Tracking Base Energy Usage**: Monitoring the energy consumption of your data center or IT infrastructure is the first step towards making it more energy-efficient[Borgini. J, 2023].

4.2.2. **Investing in Energy-Efficient Hardware:** Energy-efficient computer hardware significantly impacts energy consumption and overall sustainability. This includes laptops, servers, or peripherals[(*7 Green Computing Best Practices - InApp*, 2023)].

4.2.3. **Power Management:** Optimal power management can contribute to reducing energy consumption, lowering electricity bills, and mitigating environmental impact[(*Utilities One*, 2023)].

4.2.4. **Investing in Renewable Energy Technologies**: Organizations can consider green energy alternatives such as geothermal cooling, wind power, and hydroelectric power to power their data centers[Borgini. J. 2023].

4.2.5. **Server Virtualization**: This technology allows multiple operating systems and applications to run on a single physical server, creating virtual machines (VMs) that share the server's resources. This can improve data center efficiency

and scalability[(*How Does Server Virtualization Improve Data Center Efficiency and Scalability?* 2023)].

Server virtualization enhances data center efficiency and scalability by dynamically distributing resources to virtual machines (VMs) according to demand, hence minimizing idle capacity and optimizing performance [*How does server virtualization improve data center efficiency and scalability?* (2023)].

The majority of institutions have embraced virtualization. Gichuki et al. (2021) found that system virtualization optimizes hardware use, leading to a decrease in hardware components by an average of over 60%. This minimizes expenses related to power consumption, maintenance, carbon emissions, and wasteful disposal. It also decreases the need for additional floor space and facilitates the rapid expansion of hardware resources.

4.2.6. **Cloud Computing**: Cloud computing offers numerous benefits when it comes to sustainability and eco-friendliness compared with traditional on-site IT infrastructure solutions such as reduced energy consumption, lower carbon emissions, efficient resource utilization, scalability options, among others[Utho, 2023].

In order to achieve sustainable development, it is necessary to mitigate the adverse impacts of information technology. This can be achieved by implementing Green IT. Green IT enhances environmental sustainability through the optimization of energy consumption, reduction of greenhouse gas emissions, utilization of less hazardous materials, and promotion of the reuse and recycling of computing equipment components when reaching their renewal level (Owoche, P., et. al. (2019).

4.3. Environmental Impact of Green Computing

4.3.1. Energy Savings - Implementing green computing practices can greatly minimize energy consumption. This is accomplished by employing more efficient hardware and software, including low-power processors, solid-state storage, cloud computing, and virtualization [Zaveria (2023)].Energy-efficient computers or office equipment have the potential to save energy expenses by up to 65%, depending on how they are used (Owoche, P., et. al. (2019). In addition, the utilization of sleep mode and power management functionalities on computers can effectively conserve energy (*Green Computing and E-Waste - Office of Sustainability - University of Maine, n.d.*). According to Bonuccelli (2022), Microsoft Azure and other major public cloud providers have the potential to reduce emissions by 98% when compared to on-premises datacenters. This reduction in emissions can assist organizations in lowering their carbon footprints and decreasing expenses.

4.3.2. Reduction in E-Waste - Green computing is essential for minimizing electronic waste (e-waste). E-waste encompasses outdated, no longer in use, or discarded electrical devices that include both valuable and dangerous substances, necessitating specific procedures for their management and recycling[Prasant, P, 2020]. Green computing advocates for the reuse and recycling of electronic equipment, thus minimizing the production of e-waste (Bonuccelli, 2022). By maximizing the lifespan of existing devices, the environmental impact of electronics usage can be minimized. In addition, implementing strategies that facilitate the reutilization or recycling of current equipment

can effectively decelerate the production of electronic waste [Bonuccelli, G. (2022)].

4.3.3. Other Positive Outcomes

Green computing can also lead to other positive outcomes. For instance, it can enhance the flexibility and productivity of technology users by enabling teleworking and online collaboration, which can reduce travel costs, traffic congestion, and carbon emissions[*Zaveria* (2023)]. Furthermore, green computing can strengthen the brand and public image of businesses that adopt it by demonstrating their social responsibility and innovation[*Zaveria* (2023)].

In conclusion, the implementation of green computing practices can have a significant positive impact on the environment by saving energy, reducing e-waste, and promoting sustainability. It is a vital shift towards a more eco-conscious and efficient tech landscape[El-shamy, H, 2023].

4.4. Case Studies of Green Computing Initiatives

4.4.1. **Google -** Google, a tech giant, has made significant strides in sustainable computing by optimizing data center operations and investing in renewable energy. Their efforts have not only reduced their carbon footprint but also led to substantial energy savings and cost reduction[El-shamy, H., 2023].

4.4.2. **UPS -** UPS, a global logistics company, has been able to reduce its greenhouse gas emissions significantly by leveraging digital technologies and low carbon integrated materials. This has helped the company to reduce the environmental impact of its transportation activities, which make up the bulk of its greenhouse gas emissions[Dilmegani, C., 2023].

4.4.3. **Shenyang Aircraft Corporation -** a Chinese aircraft manufacturer, successfully carried out a green IT transformation by effectively utilizing internal resources, building confidence with

business divisions, and reinforcing their commitment to the transformation process. The integration of production and research enhanced efficiency, while the establishment of a production network supported the efficient allocation of resources[Zeng, et. al.(2018)].

4.4.4. **Arabian Gulf Oil Company -**The Arabian Gulf Oil Company, a prominent corporation in Libya, undertook a study with the aim of promoting awareness regarding green computing. The organization presented various strategies for implementing green computing, including power management, energy efficiency, online work, email usage, virtualization, and cloud computing. The study sought to establish green computing as an effective tool for management [Benamer, W. H, et. (2021)].

4.4.5. **China Mobile -**China Mobile, a telecommunications company, has harnessed collective IT resources for sustainability. Their green leadership strategy has set them apart from the competition, ultimately improving profits. Having a green image has also provided them with a competitive advantage[Li, Y, 2021].

These case studies demonstrate that green computing initiatives can lead to significant energy savings, cost reductions, and environmental benefits. They also highlight the importance of integrating such initiatives into the tech ecosystem to contribute to a more eco-conscious industry.

4.5.Challenges in Implementing Green Computing

Green computing, also known as green technology, aims to reduce the environmental impact of technology by conserving energy, reducing waste, and promoting long-term sustainability[(*Green Computing - Approaches | Challenges*, 2022)][Kirvan, P. 2023].

However, implementing green computing presents several challenges:

4.5.1. **Balancing User Satisfaction and Environmental Impact**: One of the primary difficulties lies in achieving a harmonious equilibrium between attending to the demands of systems, hardware, and software, while simultaneously prioritizing end-user contentment, complying with regulatory obligations, overseeing infrastructure reorganization, and guaranteeing a favorable return on investment [Green Computing - Approaches | Challenges, 2022)].

4.5.2. **Cost of Transition:** Transitioning to green computing often involves replacing existing technology assets with energy-efficient ones, which can be costly[Kirvan, P. 2023].

4.5.3. **Energy Consumption**: Modern technology is based on a diverse set of hardware, complex systems, and networks, all of which require energy to operate. This energy often comes from nonrenewable resources, such as coal, which have a significant environmental impact[*Green Computing - Approaches | Challenges*, 2022)].

4.5.4. **Resource Use**: Nonrenewable resources, such as precious metals like gold, are used to manufacture technology. The extraction and use of these resources can have a significant environmental impact[*Green Computing - Approaches | Challenges*, 2022)].

4.5.5. **Regulatory Compliance**: Green computing initiatives must comply with various regulatory requirements, which can be complex and challenging to navigate[(Kasam, 2023)].

4.6.Future Trends and Emerging Technologies in Green Computing

Several emerging technologies and trends are shaping the future of green computing:

4.6.1. **Cloud Computing**: Cloud computing is a major trend in

green computing. By hosting services on remote servers, cloud computing can reduce the energy consumption and environmental impact of individual devices[The future of Green IT - Tech Mahindra. (2022)]

4.6.2. **Energy-Efficient Data Centers**: Businesses are increasingly adopting green IT to design energy-efficient data centers. This includes using IT equipment for air conditioning, heating, cooling, and ventilation[(Kasam, 2023)].

4.6.3. **Green Software Development**: The development of green software spans the entire lifecycle of software: development, operation, and disposal (reuse). This includes design and coding options, choice of language, selection of AI models, and software development[(*10 Recommendations for Green Software Development*, 2021)].

4.6.4. **AI and IoT:** The convergence of technologies like AI and the Internet of Things (IoT) is greatly impacting Green Supply Chain Management (GSCM) market[Ahmad, S., et. al. (2023)].

4.7. Potential Areas for Further Research

There are several areas where further research could help advance green computing:

4.7.1. **Environmental Impact Assessment**: Further research is needed to assess the environmental impact of different computing technologies and practices[(*10 Recommendations for Green Software Development*, 2021)].

4.7.2. **Standards and Regulation:** More research could help develop standards and regulations for green computing, which could help guide businesses and other

organizations in their green computing efforts[(*10 Recommendations for Green Software Development*, 2021)].

4.7.3. **Harnessing Software for Environmental Sustainability:** Research could explore how software can be used to promote environmental sustainability, such as by optimizing energy use or reducing waste[(*10 Recommendations for Green Software Development*, 2021)].

4.7.4. **Green Algorithms and Practices in Emerging IT Technologies:** Research could concentrate on developing environmentally friendly algorithms, circuits, structures, and practices in emerging IT technologies of the current decade [Shuja, J. (2017)].

4.7.5. **Implementation of environmentally-friendly technologies at the industrial and organizational level**: The price of energy per unit will experience a substantial increase due to a large decline in global energy resources. Consequently, it is imperative for both public and government sectors to develop and implement cutting-edge strategies and plans for green computing [Shuja, J. (2017)].

5. Conclusion:

Green computing, or sustainable computing, seeks to limit energy usage, decrease electronic waste, and advocate for ecologically responsible practices across the whole IT lifecycle. Adopting green computing practices can result in substantial energy conservation, financial savings, and positive environmental outcomes. However, challenges in implementing green

computing include balancing user satisfaction and environmental impact, the cost of transitioning to energy-efficient technologies, and regulatory compliance[Salama, M. (2020)] [*Green IT Factsheet*. (2023)].

Emerging trends and technologies in green computing include cloud computing, energy-efficient data centers, green software development, AI and IoT, and carbon-aware computing[(Merritt, 2022)] [Jacob M, J. (2015)]. These advancements possess the capacity to fundamentally transform the industry by offering both environmental advantages as well as long-term commercial gains.

Potential areas for further research in green computing include environmental impact assessment, standards and regulation, harnessing software for environmental sustainability, green algorithms and practices in emerging IT technologies, and greening of industrial and organizational level technologies[(Education, 2022)].

In conclusion, green computing is a crucial approach to promote sustainable practices in the Institutions, aiming to reduce energy consumption, minimize e-waste, and protect the environment. By embracing energy-efficient technologies, responsible resource use, and proper disposal of electronic waste, organizations can contribute to a more eco-conscious and efficient tech landscape[Safdie, S, (n.d.].

REFERENCES:

Ahmad, S., Mishra, S., & Sharma, V. (2023). Green Computing for Sustainable Future Technologies and Its Applications. In *Contemporary Studies of Risks in Emerging Technology, Part A* (pp. 241-256). Emerald Publishing Limited.

Beardmore, A. (2020, October 12). *Uncovering the Environmental Impact of Cloud Computing*. Earth.Org. https://earth.org/environmental-impact-of-cloud-computing/

Benamer, W. H., Elberkawi, E. K., Neihum, N. A., Anwiji, A. S., & Youns, M. A. (2021). Green Computing case study: calls for proposing solutions for the Arabian Gulf Oil Company. In E3S Web of Conferences (Vol. 229). EDP Sciences.

Between 10 and 20% of electricity consumption from the ICT* sector in 2030? (2018, August 9). Enerdata. https://www.enerdata.net/publications/executive-briefing/between-10-and-20-electricity-consumption-ict-sector-2030.html

Bonuccelli, G. (2022, February 9). Learn the Environmental and Cost Benefits of Green Cloud Computing. Parallels Remote Application Server Blog - Application Virtualization, Mobility and VDI. https://www.parallels.com/blogs/ras/benefits-green-cloud-computing/

Borgini. J. 8 green computing best practices | TechTarget. (2023). Data Center. https://www.techtarget.com/searchdatacenter/tip/8-green-computing-best-practices

Butt, S., Ahmadi, M., & Razavi, M. (2020, November). Green Computing: Sustainable Design and Technologies. In 2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA) (pp. 1-7). IEEE.

Dilmegani, C. Top 10 Sustainability Case Studies & Success Stories in 2023. (2023). AIMultiple. https://research.aimultiple.com/sustainability-case-studies/

Education, I. C. (2022, April 19). What Is Green Computing? - IBM Blog. IBM Blog. https://www.ibm.com/blog/green-computing/

El-shamy, H. Green Computing: Paving the Way for Energy-Efficient IT Solutions. (2023). Green Computing: Paving the Way for Energy-Efficient IT Solutions. https://www.linkedin.com/pulse/green-computing-paving-way-energy-efficient-solutions-hanya-el-shamy

Energy5. (2023.). Energy5. https://energy5.com/green-it-practices-for-energy-efficient-workspaces

Expert Panel. (2023, August 3). Council Post: 20 Sustainable Practices To Help Businesses And Tech Teams 'Go Green.' Forbes. https://www.forbes.com/sites/forbestechcouncil/2023/08/03/20-sustainable-practices-to-help-businesses-and-tech-teams-go-green/

Gichuki, D. K., Nambiro, A. W. and Wabwoba, F. (2018). An Investigation of the Impact of System Virtualization on Green ICT Deployment in Kenyan Public Universities. IJSRSET, 4 (11): 25 – 29, DOI:https://doi.org/10.32628/IJSRSET173810

Green Computing and E-Waste - Office of Sustainability - University of Maine. (n.d.). Green Computing and E-Waste - Office of Sustainability - University of Maine. https://umaine.edu/sustainability/energywise/green-computing-and-e-waste/

Green computing - Approaches | Challenges. (2022, June 1). Green Computing - Approaches | Challenges.

https://adservio.fr/post/green-computing-approaches-challenges

Green IT Factsheet. (2023). Center for Sustainable Systems. https://css.umich.edu/publications/factsheets/built-environment/green-it-factsheet. Center for Sustainable Systems, University of Michigan. 2023. "Green IT Factsheet." Pub. No. CSS09-07.

Harmon, R. R., & Auseklis, N. (2009, August). Sustainable IT services: Assessing the impact of green computing practices. In PICMET'09-2009 Portland International Conference on Management of Engineering & Technology (pp. 1707-1717). IEEE.

How does server virtualization improve data center efficiency and scalability? (2023). Server Virtualization: Benefits, Challenges, and Tips. https://www.linkedin.com/advice/0/how-does-server-virtualization-improve

Jacob M, J. Sr. 2015, Relevance and Impact of Green Computing in IT, International Journal Of Engineering Research & Technology (IJERT) RTPPTDM – 2015 (Volume 3 – Issue 30),

Kasam, A. (2023, July 24). Green Computing: The Future of IT. EnterpriseTalk. https://enterprisetalk.com/featured/green-computing-the-future-of-it/

Kirui, L. K., Nambiro, A. W., & Masinde, E. O. (2023) Green Information Technology (IT) Strategies in Institutions of Higher Learning in Kenya for Sustainability: Challenges and Opportunities.

Kirvan, P. What is green computing? | Definition from TechTarget. (2023, March 1). Data Center. https://www.techtarget.com/searchdatacenter/definition/green-computing

Koratagere, S., Koppal, R. K. C., & Umesh, I. M. (2023). Server virtualization in higher educational institutions: a case study. International Journal of Electrical and Computer Engineering (IJECE), 13(4), 4477-4487.

Li, Y. Doing good while doing well: the case of business IT initiatives. (2021). Doing Good While Doing Well: The Case of Business IT Initiatives. https://knowledge.essec.edu/en/sustainability/doing-good-doing-well-business-it-initiatives.html

Merritt, R. (2022, October 12). What Is Green Computing? NVIDIA Blog. https://blogs.nvidia.com/blog/what-is-green-computing/

Owoche, P., Wabwoba, F., & Wechuli, A. N. (2019). Global Information Technology Infrastructure in Addressing the Problem of Environmental Degradation in Kenya. International Journal of Information Technology Convergence and Services, 9(1), 1-14.

Paul, S. G., Saha, A., Arefin, M. S., Bhuiyan, T., Biswas, A. A., Reza, A. W., ... & Moni, M. A. (2023). A Comprehensive Review of Green Computing: Past, Present, and Future Research. IEEE Access.

Prasant, P. E-Waste Management in Green Computing. (2020). E-Waste Management in Green Computing. https://www.linkedin.com/pulse/e-waste-management-green-computing-paidisetti-prasant-

Pratt, M. K. Cloud computing's real-world environmental impact | TechTarget. (2023). Sustainability and ESG. https://www.techtarget.com/sustainability/feature/Cloud-computings-real-world-environmental-impact

Roundy, J. Assess the environmental impact of data centers | TechTarget. (2023). Data Center. https://www.techtarget.com/searc

hdatacenter/feature/Assess-the-environmental-impact-of-data-centers

Safdie, S. What is Green Computing and What does it Imply? (n.d.). What Is Green Computing and What Does It Imply? https://greenly.earth/en-us/blog/company-guide/what-is-green-computing-and-what-does-it-imply

Salama M (2020), Green Computing, a contribution to save the environment - Lancaster University. Lancaster University. https://www.lancaster.ac.uk/data-science-of-the-natural-environment/blogs/green-computing-a-contribution-to-save-the-environment

Samoei, I., Moturi, C. and Orwa, D. (2021). Green ICT in Institutions of Higher Learning in Kenya. International Journal of Innovative Science and Research Technology. 6(3)

Shanthi, D., & Shalini, M. (2022), Green Cloud Computing–A Review. Stochastic Modeling.

Shuja, J., Ahmad, R. W., Gani, A., Abdalla Ahmed, A. I., Siddiqa, A., Nisar, K., ... & Zomaya, A. Y. (2017). Greening emerging IT technologies: techniques and practices. Journal of Internet Services and Applications, 8(1), 1-11.

Sustainable IT: Ways IT Can Be More Green and Why It Matters. (2019, October 16). BMC Blogs. https://www.bmc.com/blogs/sustainable-it/

Svistun. O, How to Achieve Energy Efficiency and Sustainability in Cloud-Based Solutions. (2023.). How to Achieve Energy Efficiency and Sustainability in Cloud-Based Solutions. https://www.linkedin.com/pulse/how-achieve-energy-efficiency-sustainability-cloud-based

The future of Green IT - Tech Mahindra. (2022).

https://files.techmahindra.com/static/img/pdf/the-future-of-green-it.pdf

The power of green computing. (2023, October 23). MIT Technology Review. https://www.technologyreview.com/2023/10/23/1081969/the-power-of-green-computing/

Utho, 4 Reasons Why Cloud Computing is a Sustainable and Eco-Friendly Choice. (2023). 4 Reasons Why Cloud Computing Is a Sustainable and Eco-Friendly Choice. https://www.linkedin.com/pulse/4-reasons-why-cloud-computing-sustainable-eco-friendly-choice-

Utilities One. (2023). Utilities One. https://utilitiesone.com/the-reliance-on-energy-intensive-data-centers-environmental-concerns

What is green computing? | Definition from TechTarget. (2023, March 1). Data Center. https://www.techtarget.com/searchdatacenter/definition/green-computing

Zaveria. (2023, August 4). Top 5 Advantages of Green Computing. Analytics Insight. https://www.analyticsinsight.net/top-5-advantages-of-green-computing/

Zeng, D., Fu, X., & Ouyang, T. (2018). Implementing Green IT transformation for sustainability: A case study in China. Sustainability, 10(7), 2160.

7 Green Computing Best Practices - InApp. (2023, September 29). InApp - Your Technology Partner for the Long Haul. https://inapp.com/blog/7-green-computing-best-practices/

10 Recommendations for Green Software Development. (2021, November 18). Green Software Foundation. https://greensoftware.foundation/articles/10-recommendations-for-green-software-development

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 71 - 74, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1010

Development of Indonesian History Digital Book Application Based On Hots Literacy as a Portable Teaching Media in the History Education Department

Rosmaida Sinaga
Departement of History Education
Universitas Negeri Medan
Medan, Indonesia

Tappil Rambe
Departement of History Education Universitas Negeri Medan
Medan, Indonesia

Abd. Haris Nasution
Departement of History Education
Universitas Negeri Medan
Medan, Indonesia

Abstract: This research is motivated by the importance of developing applications for Indonesian history e-books for students of history education, which has a limited number. This literature is increasingly needed, especially during the MBKM Curriculum in the R.I 4.0 Era. It is hoped that this product will increase student learning independence in the hybrid learning system in the current pandemic era. This study uses the Research and Development (R&D) method developed by Borg and Gall (2010). Data was collected using a questionnaire, observation, and interviews. The research instrument was tested using content validity. The data analysis technique used quantitative and qualitative descriptive. The literature developed refers to the development stage of Borg and Gall which is grouped into four stages, namely preliminary study, development, field testing, as well as product dissemination, and socialization. Literature that is developed quantitatively and qualitatively is expected to be feasible to apply in history learning, especially Indonesian history.

Keywords: Indonesian History, Digital book, Hots Literacy

1. INTRODUCTION

Every teacher, especially lecturers, has to have access to essential teaching supplies. In light of the advancement of science and the numerous new discoveries resulting from scientific study, it is imperative that every lecturer create current teaching materials. This is corroborated by Hamid Hasan's [1] assertion at the 2017 APPS seminar that the presenters' own scientific research yields the most current teaching materials. Another indicator of the growth of a topic or study is the availability of instructional resources.

In order to identify, create, and assess learning content and strategies, instructional materials are developed. The kind of instructional materials being developed, the quantity of pupils, and the availability of materials are all taken into account. The flexibility principle is used in the creation of educational materials. The ability to accept new ideas that were not already part of the course's content at the time it was implemented is a key component of the flexibility principle [2]. The flexibility principle states that through lectures given by lecturers, students might embrace new contextual values or competencies. The learning process identified by the semester learning implementation plan can help students develop these competencies (RPS). The achievement of learning outcomes is supported by the effectiveness of the implementation of the learning process with the planning of the learning process. Conducting research and development is one strategy that may be utilized to create contextual and results-focused instructional materials.

A new product or an improved version of an existing product that can be accounted for are the goals of research and development. The end outcome can take the shape of software or intangible items rather than having to be physical hardware (software). The end result (in the realm of education) might take the shape of lesson plans, books, worksheets, questions, or other learning tools like these. It can also take the form of the application of learning theory by integrating the creation of learning tools. [3]

In this instance, researchers will study Indonesian history courses at the Department of History Education FIS UNIMED while developing Indonesian history digital book applications. This is driven by the postmodernist development of historical learning theory and how epic historical narratives are now frequently linked to literacy hot topics like critical thinking, problem-solving, and writing skills[4]. The conditions of the world of education, which has now attained an R.I. of 4.0, are becoming more and more appropriate for this Literacy Hots method.

One of the required subjects in the MBKM History Education Curriculum for Higher Education is Indonesian history. However, especially for courses that are digitally based, it is uncommon to find material to accompany these courses. The only literature that is Grand Narrative, which is worthless and primarily consists of cause-and-effect stories, is what is readily available on the market. Even though the Indonesian History material is chock-full of morals and skills, Hots Literacy in particular has the potential to be assimilated by pupils, especially in light of

the current global rivalry. Students unquestionably require literature that is based on Hots Literacy in order to develop life skills in accordance with the requirements of MBKM [5].

The content on Indonesian history holds a significant place in the middle school curriculum, just like it does in universities. The absence of supporting literature that can serve as a foundation for the creation of history teaching materials in secondary schools, however, continues to be an issue in the development of Indonesian history education. In order to maximize the accomplishment of the goals of the history curriculum in middle schools and the history education curriculum in higher education, Indonesian history literature based on Literacy Hots must be developed [6].

2. METHOD

The research and development technique was applied in this study. The development research method is a technique used to create specific goods and evaluate their efficacy [7]. Research and development is a process or series of stages that can be accounted for in the development of new products or the improvement of existing ones [8].

The research and development (R&D) paradigm proposed by Borg and Gall [8] is used in this application development process. The stages of the R&D development model are as follows: a) the needs analysis stage, which entails student analysis, material analysis, and goal specification; b) the design phase of teaching materials products, which entails teaching material selection and initial design; c) the production/implementation stage of initial product development, which entails the production of teaching materials and the preparation of assessment instruments (validation sheets, questionnaires/quators). d) the validation stage, which includes evaluations from material experts (Historical Education Lecturers) and design experts (Historical Education Lecturers); e) the revision stage, which includes media revisions or improvements based on suggestions by the validators; f) the product testing phase, which includes testing the material on students, having students complete questionnaires, and implementing learning outcomes tests; g) the final revision and dissemination The effectiveness test is the last stage of this investigation [9].

3. RESEARCH RESULT

Analysis Stage

Analyzing some of the required requirements is the first step in developing this material. To serve as a guide in the creation of digital books, these demands include user determination and material selection.

1) User (user)

Students from the Department of History Education at the Faculty of Social Sciences, State University of Medan, are the intended users of the produced Digital Book Application. According to interviews with a number of students, learning on mobile devices (such as smartphones) is very engaging for them and can support independent learning whenever and whenever. According to data from a preliminary study by academics that involved observations and interviews, the typical student has an Android smartphone, which of course allows access to newly developed digital books.

2) Material Needs

The learning materials for the Indonesian History course are the ones being discussed here. The researcher then presents students with a number of learning themes she has prepared before conducting interviews. Based on the requirements they faced in the lectures, students were asked to rate how significant these issues were to them. The researcher wanted to know how crucial it was for students to learn the supplied material in order to advance their skills in Indonesian history. [10]

Students also require educational materials to be provided on the Digital Book application. According to data gathered by researchers, the typical student believes that audio and video content are crucial for teaching Indonesian history. Therefore, the researchers will offer a number of additional features, such as practice questions and films connected to the Indonesian History course, to address the needs of students. The movie that is being shown is one regarding occurrences or symptoms connected to the topic of the course on Indonesian history, and it will be used as the basis for student projects in that course.

a. Product Design Phase (Design)

The development of project-based research method teaching materials and the development of the digital book application for Android-based research method teaching materials are the two aspects of the research's product design for the digital book application, respectively [11].

1. Development of Project-Based Learning Research Methods.

At this point, the researcher creates instructional materials based on the following elements of the project-based learning model:

(a) Authentication.

The projects that students will complete in relation to the real-world issues that are, of course, connected to CPM are included in the teaching materials for the Indonesian History course established in this study. In this part of the lesson, students are guided to be able to solve issues or questions that are meaningful to them, to use issues or questions that they encounter in the real world, and to

create something that has historical significance for themselves or for the community.

(b) Maintaining academic rigor.

Students are challenged to complete a project utilizing the inquiry technique in the educational materials created for them.

(c) Expert relationships (expert relationship). The inclusion of experts or experts from outside the classroom is the secret to the effectiveness of project-based learning. Some of the designed educational materials call for students to be able to communicate with professionals in order to address issues pertaining to the project they must finish. [12]

(d) Active study (active exploration). Every topic and subtopic in the instructional materials is intended to motivate students to actively engage in research, discover, analyze, and present project outcomes.

The researchers conducted interviews and sought feedback from their professors, who were thought to be authorities in instructional design and learning models, when creating these teaching materials.

1) The initial design of products for digital book applications.

The next step is to develop a product in the form of a Digital Book Application after the stages of analysis and project-based learning-based textbook production. The front-end design, or the design for users (students taking the Indonesian History course), and the back-end design, or the design for administrators, or the design for lecturers so that they can update at any time, make up the initial design of the Android-based Digital Book Application product. created a digital book app.

2) Design of the digital book application display.

(a) Initial Application Screen

After a user accesses an application, they are presented with the initial view (login). The user must enter their user name and password in the two fields that are required on the display.

The lecturer, who also serves as the admin for this digital book application, will offer a password to each student who has signed up to take the Indonesian History course.

(b) Front View Design

After the learner has successfully completed the Login stage, the front screen is the display that appears. A menu can be found on the front screen. Digital book application items are shown on the "About" menu in the first menu, which is also a menu. The third menu, "6 assignments," will display items under menu 6 of the KKNI. The menu "Materials" contains materials and sub-materials that will be discussed in lecture activities. RPS is found on the fourth menu.

c) "Home" Display Design

The "Home" menu is a menu that includes information about the Digital Book Application, such as: 1) the name of the Digital Book Application, 2) its goal, and 3) how to use the Digital Book Application.

d) The "Home" sub-Material menu's Display Design.

The teaching materials in the online edition of the Indonesian History course, both in written form and in the form of images and videos, are displayed in a display called the material menu. Students can download the RPS (Semester Learning Plan) and PowerPoint slides from this menu.

e) 6 Task Display Design.

This digital book application is made to assist the blended learning approach in addition to the six goals specified by the KKNI curriculum that is applicable to Medan State University. There is a thorough explanatory display of the 6 tasks that students must complete on the 6 task menu.

b. Product Development Phase (Development)

The creation of the Digital Book Application is in the development stage. This Digital Book Application was created with the design in mind. First, the content-filling elements for this digital book application—such as pictures, videos, and animations—are gathered. Google, YouTube, and electronic books are used to find these resources.

c. Implementation Phase (Implementation)

Trials of products that have been created for a number of respondents are included in the implementation phase. The researcher creates a blended learning scenario as part of the implementation [13]. The four meetings that made up the blended learning implementation for the lectures on Indonesian history included two face-to-face meetings and two online meetings using Zoom and Dendroid web meetings. The learning media expert and the material expert were the two specialists the researcher asked to evaluate the Digital Book Application product during the trial [14].

Student Interest in Android-Based Learning

Students were asked to complete a response questionnaire by giving an assessment of each indicator by putting a check mark () on the range of numbers that are considered appropriate, namely (4) to strongly agree, (3) to agree, (2) to disagree, and (1) to disagree. This was done in order to

determine students' interest in learning the Research Methods course using the Android Digital Book Application based on Literacy Hots. The percentage of values received from each aspect, calculated from the results of the questionnaire, was then used to further analyze the data [15]. To do this, the total value of each aspect was divided by the maximum number and multiplied by 100%.

According to the results of the student response survey, the percentages for features of student happiness, interest in utilizing applications, attentiveness, and involvement were 86.49%, 87.67%, 85.81%, and 88.24%, respectively. It is clear from the four elements that students have a highly positive attitude toward learning, and it can be inferred that they are very interested in using the produced Digital Book Application to study.

4. CONCLUSION

The Literacy Hots-based Digital Book application that was developed can be one of the effective learning media for students and help teachers in learning Indonesian history. The Hybrid Learning learning model will certainly be more effective when coupled with the use of the Digital Book Application. In addition, the use of the Digital Book Application can stimulate students in developing a learning and independent culture and also increase students' Literacy Hots abilities.

5. ACKNOWLEDGEMENTS

We would like to thank LPPM Universitas Negeri Medan for funding our research and all participans and supervisors that contributed to the work in this study.

6. REFFERENCES

[1] Amri, S. (2013). Pengembangan dan Model Pembelajaran dalam kurikulum 2013. Jakarta: Prestasi Pustaka.
[2] Arsyad, Azhar. (2011). *Media Pembelajaran*. Jakarta: Rajagrafindo Persada.
[3]Darmadi, H. (2011). Metode Penelitian Pendidikan. Bandung: Alfabeta.
[4] Darmawan, D. (2014). Pengembangan E-Learning Teori dan Desain. Bandung: PT Remaja Rosdakarya.
[5] Effendi, Empy dan Hartono Zhuang, (2005). "E-learning Konsep dan Aplikasi". Penerbit: Andi Offset. Yogyakarta.
[6] Hobri. (2009). *Metodologi Penelitian dan Pengembangan (Dvelopmental Research) (Aplikasi pada Penelitian Pendidikan Matematika).* Jember: FKIP Universitas Jember.
[7] Kadir. (2015). Statistika Terapan. Jakarta: Raja Grafindo Persada.
[8] Lehman, Lynn. (2007). *A Proactive Approach to Employee Training & Development.* Rising Sun Consultants.
[9] Miarso, Yusufhadi. (2007). *Menyemai Benih Teknologi* Pendidikan. Jakarta: Kencana Prenada Media Group.
[10] Muhtadi, Ali. (2013). Pemanfaatan Program *Computer Assisted Instruction* (CAI) *dalam Program Pembelajaran Berbasis Internet*. Universitas Negeri Yogyakarta. Diakses pada tanggal 6 Desember 2013.
[11] Munadi, Y. (2013). Media Pembelajaran Sebuah Pendekatan Baru. Jakarta Selatan: Referensi.
.
[12] Nasution, Abd Haris. (2021). Pembelajaran Sejarah abad 21. Obelia Publisher: Medan.
[13] Sadiman, Arief S.(2010). *Media Pendidikan*. Jakarta:RajaGrafindo Persada.
[14]Sugiyono, P. D. (2017). Metode Penelitian Kuantitatif, Kualitatif, dan R&D. Bandung: Alfabeta.
[15] Supriatna, Nana. (2007). Komstruksi Pembelajaran Sejarah Kritis. UPI Press: Bandung.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 75 - 79, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1011

The Development of Electronic Module of Indonesian History of the Reform Based on the Case Method to Improve High-Order Thinking Skills

Arfan Diansyah
Departement of History Education
Universitas Negeri Medan
Medan, Indonesia

Yushar
Departement of History Education
Universitas Negeri Medan
Medan, Indonesia

Syahrul Nizar Saragih
Departement of History Education
Universitas Negeri Medan
Medan, Indonesia

Abstract : This study aims to produce an electronic module based on the Case Method that is feasible and effective in improving students' higher-order thinking skills in the material of Indonesian History during the Reformation Period. This research method uses research and development methods using the ADDIE development design. The research stages of the ADDIE development model consist of Analysis, Design, Develop, Implement and Evaluate. The research subjects were students of the Department of History Education, State University of Medan class 2019. Research data were obtained using questionnaires and tests. The research instrument used validation sheets, student response questionnaires, and tests of higher order thinking skills. Validation data and student response questionnaires were analyzed using the product trial eligibility criteria. Higher order thinking skills test data were analyzed by comparing the test results before and after using the electronic module. The results of the study show 1) the electronic module developed has a good category based on the assessment of material experts with a content feasibility aspect score of 3.73 and a presentation feasibility aspect score of 3.5. 2) the developed electronic module has a good category based on the assessment of media experts with a score of 3.6. 3) the developed electronic module has a good category based on a student response questionnaire of 3.66. 4) the developed electronic module can improve students' higher order thinking skills with an average score of the initial test before using the electronic module of 64.5 and the average score of the final test after using the electronic module of 80.6.

Keywords: Development, Electronic Module, Cased Method, Higher Order Thinking

1. INTRODUCTION

The era of globalization with the spirit of change and openness has brought many positive and negative changes to national life. Openness and individual freedom as part of the characteristics of western democracy are increasingly dominating the way of thinking, behaving,and acting of the next generation. The era of globalization is also marked by the development of science and technology. To respond to these changes, the younger generation must have resources such as communication, collaborative, critical thinking, and creativity. Higher education is one of the strategic components in shaping human resources. At the higher education level, the learning process is directed so that students can think critically and skillfully in solving problems. On the contrary, the teaching and learning process is still 'theory oriented' with conventional approaches and learning methods, namely lectures, question and answer, and discussions. Likewise, students' mastery of the materials is still very limited in understanding the text without understanding it more deeply, linking it with other concepts or linking it to the real world.

Learning history in the midst of the development of technology and information demands a change in the implementation of historical education. These innovations can be in the form of learning strategy, learning media and teaching materials so that history learning can function properly. One of the innovations that can be done is to develop an electronic module with a case method approach. The main reasons for developing case-based electronic modules are (1) teaching materials require illustrations of real cases in the application of science; (2) the available teaching materials are still based on theory, and (3) to build strong reasons for students to understand, solve problems, and apply materials in the nation and state.

The current availability of teaching materials which are still theoretical unconsciously directs lecturers to employ the lecture method in their teaching and learning process. . This learning condition certainly has not been able to hone students' analytical skills, sensitivity to problems, problem-solving skills and the ability to evaluate problems holistically. When attending lectures, students are limited to understanding while taking notes. Lecturers become the main role in achieving learning outcomes and seem to be the only source of knowledge. The lecturer-centered learning with passive students has low learning effectiveness. Therefore, the development of an electronic module on the Indonesian history during the reformation with a case method approach becomes an alternative in dealing with the current problem of learning history which tends to bememorization.

The module is considered as a teaching material that is compiled in a complete and systematic way supported by a set of learning experiences that are planned and designed to help students master specific learning objectives. The module at least contains learning objectives, learning materials/substances, and evaluation. The module functions as an independent means, so students can learn independently according to their speeds (Daryanto, 2013). Electronic

modules are information and technology (ICT)-based learning materials that are interactive because of the ease of navigation, display of images, videos, and feedback through formative tests (Nina Ikhwati, 2019). It is interactive because the process includes interactions and other activities such as paying attention to pictures, listening to sounds, and even paying attention to videos that aim to increase learning motivation and enthusiasm and has a high graphic value in its presentation. The electronic module is a form of presenting self-study materials that are systematically arranged into the smallest learning units to achieve certain learning objectives which are presented in an electronic format including animations, audios, and navigations that make users more interactive (Sugianto, 2013).

The use of electronic modules in the learning process can foster creativity, productive thinking habits, and active learning conditions so that students can develop literacy skills in the digital era. The advantages of electronic modules compared to print modules are that their interactive nature makes it easier to navigate, allows displaying images, audio, video, and animation and is equipped with formative tests or quizzes that allow immediate automatic feedback (Suarsanaga and Mahayukti, 2013). Based on the research conducted by Diky, et al., (2017), Development of E-Learning Module: Historical Culture Society Based on Local Genius shows that the electronic module on the history of the Using culture is effective in increasing students' knowledge of the cultural history of the Using community.

The case method is learning by using real-world cases implemented in the classroom. The case method describes a real situation related to the material being studied, simulating real-world conditions into a controllable environment in the classroom through discussion in the decision-making process (Jogiyanto, 2006). The case method emphasizes the process of solving cases or problems faced scientifically, placing cases as keywords in the learning process. The case method is implemented by teachers and lecturers by choosing lesson materials that have cases that can be solved. Through learning with the case method, students can develop analytical and innovative thinking skills, critical and analysis, creativity skills and reasoning skills in problem solving. These cases can be taken from textbooks or other sources, for example events that occur in the surrounding or family (Sanjaya, 2006).

Higher-order thinking is a student's thinking activity that involves a high-level hierarchical cognitive level from Bloom's taxonomy of thinking, which includes analyzing, evaluating, and creating (Andersen & Krathwohl, 2015). Higher order thinking skills (HOTS) can be achieved when students actively understand and integrate knowledge with their experiences (Anderson & Krathwohl, 2015). In line with this, Deluca (2011) states that to develop higher-order thinking skills, students must first understand factual, conceptual, and procedural knowledge, apply their knowledge to learn by doing, and then reflect on the process that produces a solution. Teachers can do this by guiding students through observation activities and concept formation, giving responses, analyzing, comparing and providing the necessary considerations. The activeness of students and the guidance of teachers greatly contribute during learning (Zerihun et al., 2012).

Based on the above background, the research questions in this study are 1) how is the feasibility of an electronic module - using the case method based on the assessment of material experts and media experts? 2) how is the feasibility of an electronic module using the case method based on student responses? 3) how to improve students' higher order thinking skills by using electronic modules using the case method?

2. METHOD

This study used research and development methods using the ADDIE development model. The research stages of the ADDIE development model consist of Analysis, Design, Develop, Implement and Evaluate. The participants were 30 undergraduate students of the Department of History Education, State University of Medan, class of 2019. The data were obtained from questionnaires and tests. The research instrument used validation sheets, student response questionnaires, and tests of higher order thinking skills. Validation data and student response questionnaires were analyzed using the product trial eligibility criteria. Higher order thinking skills test data were analyzed by comparing the test results before and after using the electronic module.

3. RESEARCH RESULT

Electronic Module Feasibility Analysis

The feasibility of the electronic module was conductedusing an assessment instrument of material and media experts. The results of the material expert assessment are shown in Table 1.

Table 1. Material Expert Validation Results

No	Aspect	Score	Criteria
1	Content feasibility	3.75	Good
2	Presentation feasibility	3.5	Good
3	Linguistic feasibility	3.8	Good
	average	3.68	Good

Source: Research Results, 2022

As seen in Table 1, the developed electronic module was considered feasible based on the assessment of material experts with an average score of 3.68 (good). In the aspect of content, the substance of the electronic module material was in accordance with learning outcomes. The materials presented in the module used recent cases in order to encourage students' curiosity to read and study it. On the presentation aspect, the electronic module had a good appeal which was placed in several parts. The front cover used the combination of colors, illustrations, and matching font sizes. In the content section of the module, images were provided as stimulation. In addition, the actual problems experienced by students in daily life were presented. Various forms of assignments and exercises were also provided to make them more interesting.

The feasibility of the electronic module was not only based on the assessment of material experts but also based on the assessment of media experts. The results of the media expert's assessment are presented in Table 2.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 75 - 79, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1011

Table 2. Media Expert Validation Results

No	Indicator	Score	Criteria
1	Electronic module size	4	Good
2	Electronic module layout design	4	Good
3	Electronic module content design	3	Enough
4	Color selection accuracy	3	Enough
5	Image usage accuracy	4	Good
	average	3.6	Good

Source: Research Results, 2022

As described in Table 2, an average score of 3.6 (good) was obtained. These results indicated that in terms of the appearance of the media, the teaching materials developed were considered as feasible . The developed electronic module used a font size that was easy to read and in accordance with the students' characteristics. The module also used a proportional letter comparison between the title and content. In the layout, the spacing between columns was well-arranged. Alternating between paragraphs began with a capital letter. The images on the modules were presented according to the substance of the material presented.

Student Response Analysis

To get the students' evaluation regarding the developed electronic module, learning activities were implemented. Students used case method-based electronic modules as teaching materials during learning. The results of student responses to the electronic module are presented in Table 3.

Tabel 3. Student Response Analysis

No	Indicator	Score	Average	Criteria
1	Ease of learning material	94	3.1	Enough
2	presentation according to learning achievement	113	3.7	Good
3	presentation requires students to think actively	104	3.5	Enough
4	presentation requires students to explore information	119	3.9	Good
5	presentation of pictures makes it easier for students to understand	115	3.8	Good
6	presentation contains cases	107	3.6	Good
7	compatibility of language with student development	116	3.8	Good
8	language used is communicative	112	3.7	Good
9	book display attracts interest to read	110	3.6	Good
10	Readable text size	118	3.9	Good
	Average	3.66		Good

Source: Research Results, 2022

Based on the results of the students' evaluation as depicted in Table 3, an average score of 3.66 was obtained. These results indicated that the developed electronic module was considered good.

Analysis of Higher Order Thinking Skills

The results of the students' higher order thinking skills test were obtained by comparing the average scores of the higher order thinking skills test before and after using the electronic module. The description test consisting of 5 questions about the history of Indonesia during the reformation was used. The average score before using electronics is presented in Table 4

Table 4. Initial test scores

Total score	**1935**
Average	64.5
Highest score	70
Lowest score	52

Source: Research Results, 2022

Based on the table above, the higher order thinking skills were still relatively low. The highest score was 70 and the lowest scorewas 52. The average student critical thinking skills test results were 64.5. These results were obtained from the initial test before the learning activities were implemented. In the next activity, learning activities were provided using case method-based electronic modules. At the end of the lesson, students were given a description test of 5 questions. The students' final test results are presented in Table 5.

Table 5. Final test scores

Total score	**2419**
Average	80.6
Highest score	87
Lowest score	72

Source: Research Results, 2022

Based on the table above, it can be seen that the higher order thinking skills increased. The highest score was 87 and the lowest score was 72. The average score of the student's critical thinking skills test on the final test was 80.6. These results indicated that there was an increase in the average score from 64.5 on the initial test to 80.6 on the final test. Thus, it can be concluded that the use of electronic modules based on the case method approach can improve students' higher order thinking skills.

The development of an electronic module on Indonesian history during the reformation based on a case method was chosen to address student problems regarding student activity and learning outcomes. The electronic module was developed in various stages including analysis, planning, development, implementation, and evaluation stages. The analysis phase was carried out by conducting needs analysis, analyzing

learning resources used by students, and identifying learning problems. The planning stage was done by compiling learning objectives, compiling the subject matter in the module, and determining the application used in the module. The development stage was conducted by making prototype modules arranged according to the learning objectives and presenting cases as problem solving materials. The implementation phase was carried out by directly testing the use of the developed electronic module. The evaluation stage was implemented by giving a test of students' high-level thinking skills after using the electronic module.

The electronic module developed with the case method approach contains cases in the form of gaps, difficulties that occur in aspects of life that are adapted to the subject matter. By giving cases to students through the developed electronic module, students are given the opportunity to analyze, propose solutions, evaluate solutions, solve problems, and make decisions.

The electronic module based on the case method developed presents cases focusing on the cognitive, psychomotor, affective, and motivational aspects. Seidel and Godfrey (2005) states that there are four main characteristics of the case method, namely cognitive, psychomotor, motivational, and affective (interpersonal and attitude). The cases selected in the electronic module are cases that are very close to students' daily lives and are meaningful. Therefore, educators are expected to provide meaningful learning. required When students learn something and find meaning, that meaning can be a reason to continue learning (Nazgul et al., 2020).

The materials compiled in the electronic module relate to concrete or real problems in daily life that can ultimately improve higher-order thinking skills. Based on the results of the student's higher order thinking skills test, it was found that the electronic module developed could improve students' higher order thinking skills. With this case-based electronic module, the learning method applied was also case-based learning. The use of case method-based electronic modules in this study became effective when used in learning activities using the group discussion method. It occurs because the complexity of the case method not only based on the scale of the problem, but also the complex cognitive, psychological, social, and behavioral interactions between group members during the problem-solving process (Lightner, Bober, & Willi, 2010). The group discussion method can hone and improve critical thinking skills for problem solving, communication skills, collaboration, and creativity, so learning is more meaningful and students can benefit from learning because the problems solved are directly related to real life; students are more independent and mature, able to give and receive opinions from others, and instill positive social attitudes among students (Endah Andayani et al., 2022).

Student activities using the case method-based electronic module developed were directed at a problem-solving activity. Through problem solving, students were facilitated in a learning that trains higher order thinking skills. Through the electronic module, students could build representations, analyze, and build relationships in problem solving. The results are in line with those of Danilin's research (2021) that the case method can develop analytical skills, critical thinking, creative thinking, practical skills, communication skills, social and reflexive skills. Another study revealed that all levels of thinking in Bloom's Taxonomy can be achieved using case learning methods (Kulshrestha, 2021). In addition, the case method can be used to increase the activities and character of student cooperation (Nugroho, Bramasta, & Pamijo. 2018).

4. CONCLUSION

The development of an electronic module on the Indonesian history during the reformation using a case method was developed through a process of analysis, planning, development, implementation, and evaluation. The materials presented in the module concern about the cognitive, psychomotor, affective, and motivational aspects. The developed electronic module facilitated students to hone and improve higher order thinking skills through problem solving activities based on a case. It is suggested that educators should develop case-based electronic modules to improve students' higher-order thinking skills.

ACKNOWLEDGMENT

Acknowledgments are addressed to LPPM Universitas Negeri Medan for providing research funding support for the 2022 applied research scheme so that the research and writing of this research article can be completed. The author would also like to thank all parties involved in supporting the implementation of the research.

REFERENCE LIST

Andayani, Endah, & Mustikowati, Rita Indah. (2022). *Case Method: Mengoptimalkan Critical Thinking, Creativity Communication Skills dan Collaboratively Mahasiswa Sesuai MBKM di Era Abad 21*. Jurnal Penelitian dan Pendidikan IPS (JPPI) 16 (1), 52-60.http://ejournal.unikama.ac.id/index.php/JPPI.

Anderson, L.W. & Krathwohl, D.R. (Eds). 2015. Kerangka landasan untuk pembelajaran, pengajaran, dan asesmen: revisi taksonomi pendidikan Bloom.(Terjemahan Agung Prihantoro). Yogyakarta: Pustaka Pelajar.

Danilin, R. (2021). Case method in the students university foreign language education. *Tambov University Review. Series: Humanities*, 95-106. https://doi.org/10.20310/1810-0201- 2021-26-195-95-106

Daryanto, 2013. *Menyusun Modul Bahan Ajar untuk Persiapan Guru dalam Mengajar*. Yogyakarta: Gava Media.

Deluca. 2011. The GRIDc Project Developing Students's Thinking Skills in a Data-Rich Environment. Journal of Technolofy Education. 23. https://doi.org/10.21061

Diky Aprianto, Nurul Umamah, Sumardi. 2017. *Development of E-Learning Module: Historical Culture Society Based on Local Genius*. Jurnal Historica vol. 1 no.2 hal 6

Dony, Sugianto, et.al, 2013. *Modul Virtual Multimedia Flipbook Dasar Teknologi Digital*. Jurnal INVONTEC, hal 102.

Jogiyanto. 2006. *Metode Kasus*. Jakarta: Andi.Sanjaya, 2006

Kulshrestha, R. (2021). Effective Use of Case Method as a Pedagogical Tool. *Book Chapter: Case Method for Digital Natives.* India: Bloomsbury Publishing

Lightner, S., Bober, M. J., & Willi, C. (2010). Team-Based Activities To Promote Engaged Learning. *55*(1), 5–18. https://doi.org/10.3200/CTCH.55.1.5-18

Nazgul, K., Anar, B., Baglan, Z., Moldir, S., Sadvakasova, G., & Nishanbayeva, S. (2020). Preservice Teachers ' Opinions on the Use of Technology in Education. *International Journal of Emerging Technologies in Learning*, *15*(23), 182–192. https://doi.org/https://doi.org/10.3991/ijet.v15i23.18831

Nina Ikhwati Wahidah, Nurdin Ibrahim, Suyitno Muslim, 2019. "*E-Module: Design A Learning Material With Rowntree And Hannafin Model For Higher Education"* International Journal Of Scientific & Technology Research Volume 8, Issue 12. Hal 7

Nugroho, A., Bramasta, D., & Pamijo. (2018). The The Implementation of Case Study Method to Develop Student's Activities and Characters. *Jurnal Sains Sosial dan Humaniora*, 2(2), 175-178. https://doi.org/10.30595/jssh.v2i2.3349

Seidel, R., & Godfrey, E. (2005). Project and Team based Learning : An Integrated Approach to Engineering Education. *4th ASEE/AaeE Global Colloquium on Engineering Education*.

Suarsanaga, Mahayukti, 2013. *Pengembangan E-Modul Berorientasi Pemecahan Masalah Untuk Meningkatkan Keterampilan Berpikir Kritis Mahasiswa*. Jurnal ISSN , Vol. 2, No. 3. hal

Zehrihun, Z. Beisshunzein, J & Van Os, W. 2012. Student Learning Experience as Indicator of Teachin Quality. Educational Assessment, Evaluastion and Accountability, 24. http://dx.doi.org/10.1787/19939019

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 80 - 86, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1012

Smart Charge Controller using Buck Converter Topology for Bicycle Power Generator in DC House Electrical Grid

Lutfir Rahman Aliffianto
Electrical Automation
Engineering Department
Institut Teknologi Sepuluh
Nopember
Surabaya, Indonesia

Ilham Agung Wicaksono
Electrical Automation
Engineering Department
Institut Teknologi Sepuluh
Nopember
Surabaya, Indonesia

Rezki El Arif
Biomedic Engineering
Department
Institut Teknologi Sepuluh
Nopember
Surabaya, Indonesia

Abstract: Human energy is a renewable and abundant source. It can be effectively used in DC home electric systems. To convert human energy into electrical energy, a bicycle power generator is required, accompanied by a battery to store the converted energy. Nevertheless, the varying nature of the energy generated through pedaling may lead to swift damage to the battery. The author designed and built a charge controller to ensure a constant flow of voltage and current to the battery. Abbreviations of technical terms are explained the first time they are used. The language is clear, objective, and formal, without filler words or biased language. The writing adheres to conventional structure and uses precise word choice. Grammatical correctness and consistent formatting are employed, with common academic sections included.

The Buck Converter topology regulates the field current (I_f) entering the alternator to maintain a stable alternator output voltage. The current from the source flows into the Buck Converter circuit. The microcontroller adjusts the duty cycle to set the field current (I_f) in the Buck Converter. When the rotation of the alternator is high, the value of the field current (I_f) is decreased, and when the rotation is low, the value of the field current (I_f) is increased in order to maintain a constant output voltage despite the changes in rotation.

Keywords: chargecontroller, topology buckconverter, bicyclepower generator, optimization, charging coordination.

1. INTRODUCTION

Human power can serve as an alternative energy source due to its abundance. By converting energy from human power into electrical energy, bicycle power generators can be utilized as a power source in DC home installations. By converting energy from human power into electrical energy, bicycle power generators can be utilized as a power source in DC home installations. The tool effectively converts human energy into usable power, making it a practical and eco-friendly option. By converting energy from human power into electrical energy, bicycle power generators can be utilized as a power source in DC home installations.

In implementing the bicycle power generator, several issues arise, particularly in the form of heavy pedaling due to the field current (If) required by the alternator to generate electrical energy. This results in people's disinclination to generate energy through this bicycle power generator. Another issue with the bicycle power generator is that the voltage generated by the alternator tends to fluctuate due to changes in human power input. Directly connecting the battery to this fluctuating voltage can shorten its lifespan. This problem can be resolved by using a charge controller that regulates the field current (If) entering the alternator and produces a stable voltage to charge the battery.

This research focuses on the optimization of charge controlling for bicycle power generator in installation system of DC House. This charging control takes advantage produces a stable voltage to charge the battery using buck converter charge controller.

2. LITERATURE REVIEW

2.1 Electrical Installation of DC House

The DC House electrical installation consists of the main components: the source of electrical energy, the charge controller, the battery, the MISO (Multiple Input Single Output) DC-DC converter, and the load in the form of DC equipment. To illustrate the main system in the DC House electrical installation, refer to Figure 1.

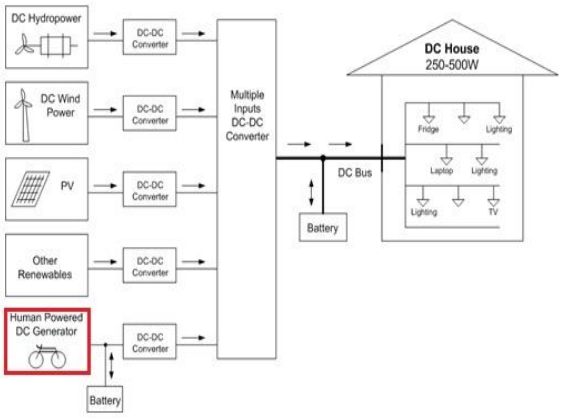


Figure 1. The main system in the DC House electrical installation.

2.2 Bicycle Power Generator

A bicycle power generator is an electricity source for DC houses that converts human pedaling power into electrical energy. The process involves pedaling a stationary bicycle connected to an alternator via a v-belt. The electrical energy produced by the alternator can be stored in a battery or used directly[1].

The main components of the bicycle power generator are bicycle, alternator, bike stand,charge controller, v-belt, battery.

2.3 Buck Converter

The buck converter operates based on two conditions: the switch is in either the ON or OFF state. In the ON state, as

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 80 - 86, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1012

illustrated in Figure 2, the input current flows through the inductor L and capacitor C because the diode is non-conductive. During this time energy is also stored in inductor L. When VC> VOUT, capacitor C discharges and transfers energy to the output. When the switch is OFF, as depicted in Figure 3, the capacitor C transfers energy from the input to the output. The diode is conducting because of the energy previously stored in inductor L. Next, the inductor current iL flows partly to capacitor C, and the energy is then transferred to the output. The value of inductor current iL continuously decreases until the switch is turned ON again during the next cycle.

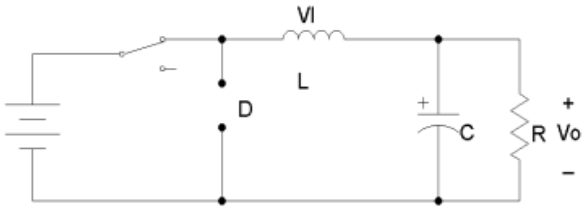


Figure 2.Buck converter circuit when switch is ON

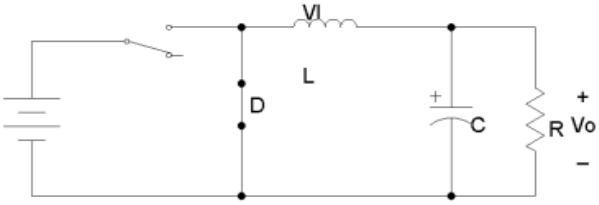


Figure 3.Buck converter circuit when switch is OFF

3. RESEARCH METHODE

This research addresses two main issues: how to create a Charge Controller utilizing the Buck converter topology to regulate the field current (If) that enters the alternator and stabilize the alternator's output voltage that enters the battery, as well as how to determine the appropriate amount of PWM duty cycle given to the Buck converter by the microcontroller control logic. Both issues have practical implications. To implement this research, it is necessary to conduct a literature study on the supporting theory, design and manufacture appropriate tools, perform tests and analysis, and draw conclusions.

3.1 Design and Development of Equipment

The design of the charge controller using a buck converter topology is segmented into various components, including the design of the MOSFET driver, the ATMega8 microcontroller acting as a pulse width modulation (PWM) regulator, a 5-Volt regulator acting as the minimum power source of the microcontroller system, and current and voltage sensors. The software design component of this project is presented in the form of a flow chart for the charge controller system.

3.2 Testing and Analysis

Testing is conducted on each circuit block, with the results observed. Technical terms are explained when first used, and clear, concise language is consistently employed throughout. After testing each block, the testing is then carried out on the entire block, resulting in one charge controller system. The testing process is divided into two stages :

1. Testing the power supply circuit
2. Testing the minimum microcontroller system circuit
3. Testing the MOSFET driver circuit
4. Testing the buck converter circuit
5. Conducting current sensor andvoltage sensor testing
6. Finally, testing the overall charge controller system for the bicycle power generator.

4. Design and Development of Equipment

4.1 Block Diagram of Charge Controller System

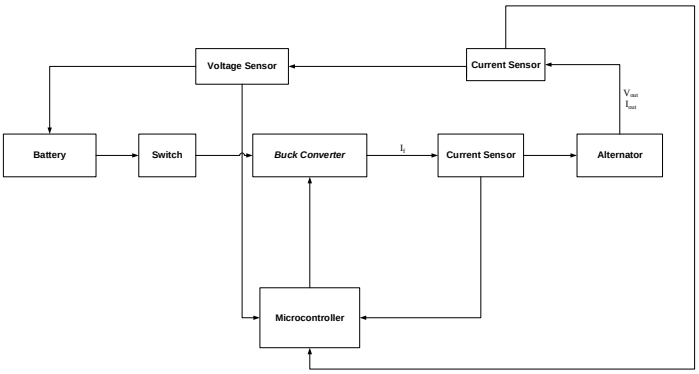


Figure 4. Block Diagram of Charge Controller System

The diagram in Figure 4 illustrates the initial step of the system, where the switch is manually closed to turn the buck converter on. A current sensor measures the output current value of the converter to determine if it meets the minimum field current required by the alternator. If it does not, the microcontroller adjusts the duty cycle value accordingly. The alternator is rotated using a static bicycle, and human pedaling energy is then used to produce electrical energy. The voltage sensor determines the battery charging voltage from the alternator. If the alternator output voltage exceeds or falls short of the battery's optimum charge voltage, the microcontroller adjusts the field current (If) by modifying the duty cycle value. The current sensor placed downstream of the alternator detects the charging current value for the battery. If the minimum current required for charging is met, current flows to the battery. To adjust the field current (If), the microcontroller adjusts the duty cycle as necessary. The MOSFET will act as a switch for the duty cycle set by the microcontroller. Technical term abbreviations will be explained upon their first use. The duty cycle is the ratio between the on time (logic 1) and the switching period and its value will impact the Buck Converter circuit's ability to increase or decrease the field current value (If).

4.2 Hardware Design of Charge Controller

The design of the charge controller hardware is divided into several parts, including:

1. Design of buck converter
2. Design of the power supply circuit
3. MOSFET driver design
4. Design of minimum microcontroller system
5. Design of voltage sensor

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 80 - 86, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1012

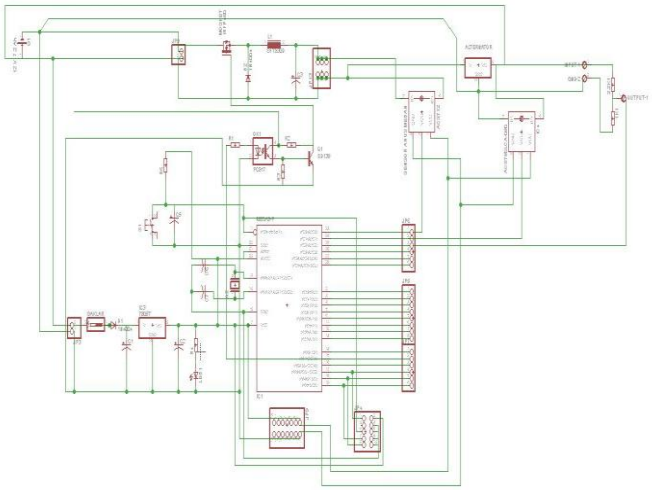


Figure 5. Overall circuit of the charge controller system

4.3 Charge Controller Software Design

The charge controller software is developed using the C language and the CodeVision AVR compiler. The algorithm supporting the operation of the device is shown in the flowchart in Figure 6.

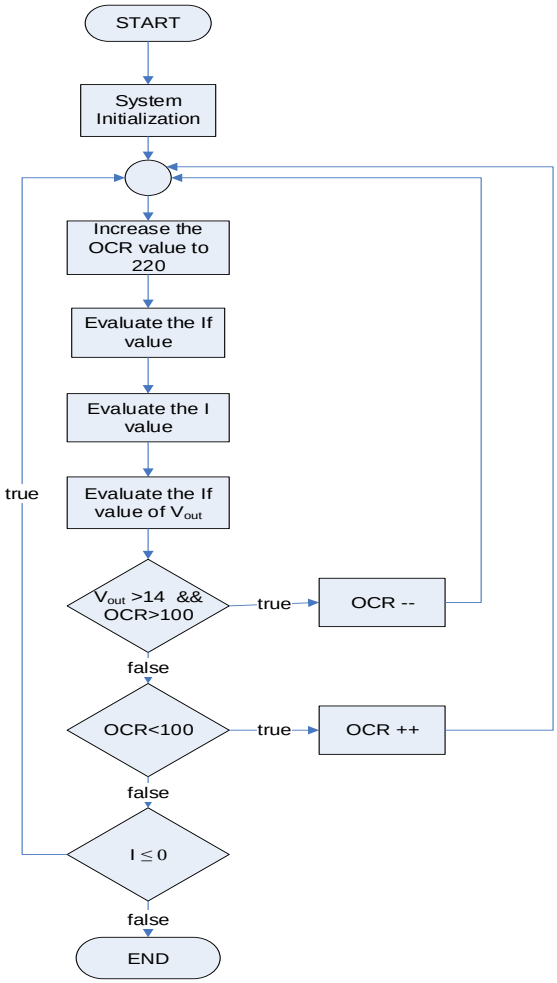


Figure 6. Flowchart of the software

5. RESULTS AND DISCUSSION

5.1 Microcontroller Power Supply Circuit Testing

This test verifies the correct operation of the microcontroller's power supply circuit, which generates a 5V voltage used to power the microcontroller.

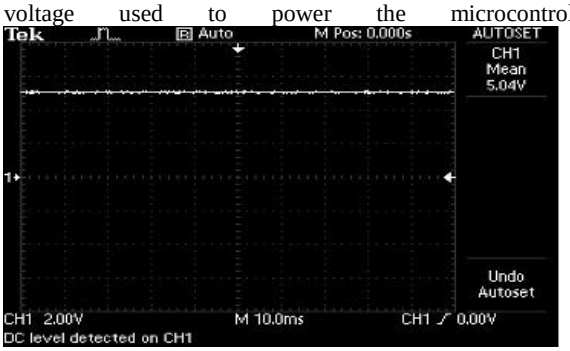


Figure 7. Example of microcontroller power supply circuit test results

The test results for the microcontroller power supply circuit are shown in Figure 7. The results demonstrate that the LM7805 regulator, with an input voltage of 12 volts for the 5 volt power supply circuit, outputs a voltage of 5.04 volts. As per the LM7805 regulator datasheet, the minimum and maximum output voltages are 4.75 volts and 5.25 volts, respectively. Therefore, it can be concluded that the LM7805 regulator operates appropriately.

5.2 PWM Testing on Microcontroller

This experiment aims to observe the PWM waveform and frequency of the Buck Converter circuit. The required frequency for the circuit is 50kHz, as determined by the design. An oscilloscope is necessary to visualize the PWM waveform and frequency.

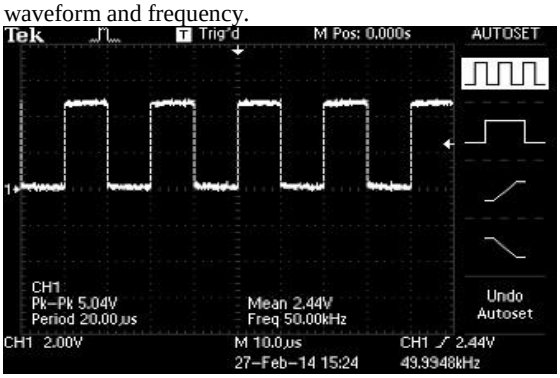


Figure 8. Example of 50% PWM waveform and 50 kHz frequency

The microcontroller is programmed with 50% duty cycle and a 50 kHz frequency. Figure 8displays a PWM waveform example generated by the microcontroller with 5 volts input voltage, the duty cycle set to 50%, and a frequency of 50 kHz.

5.3 MOSFET Driver Circuit Test

The purpose of this test is to verify that the MOSFET driver can drive the output signal from the microcontroller according to the circuit requirements using a 12V input voltage. Table 1 displays the output test results for the MOSFET driver with different duty cycle values.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 80 - 86, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1012

Table 1. MOSFET driver test results table

No.	Duty Cycle (%)	Calculation Voltage (volt)	Measured Voltage (volt)
1.	0	0	0
2.	25	3	3,20
3.	35	4,2	4,47
4.	50	6	5,90
5.	75	9	9,02
6.	100	12	11,80

An example of a mosfet driver output waveform result is shown in Fig. 9.

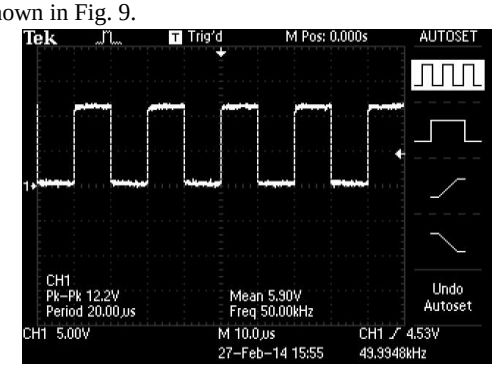


Figure 9. Example of MOSFET driver output waveform with 50% duty cycle

The waveform of the MOSFET driver circuit matches the input that comes from a PWM signal generated by the microcontroller.

5.4 Buck Converter Circuit Test

This experiment aims to observe the efficacy of the Buck Converter circuit in reducing voltage. According to the results, the Buck Converter circuit works effectively. To conduct the experiment, a source voltage of 10.98 volts is supplied from a battery to the input of the Buck Converter. Figure 10 depicts the successful operation of the Buck Converter circuit as demonstrated by the reduction of voltage from 10.98 volts to 5.87 volts. Clearly, this circuit works effectively.

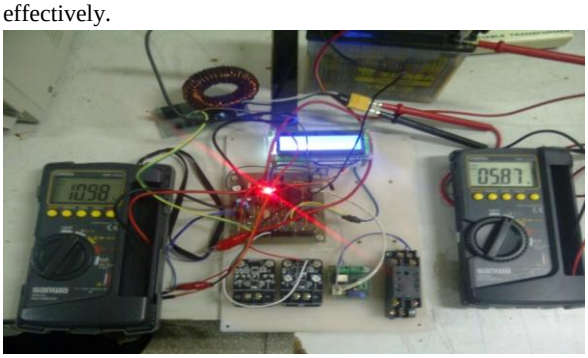


Figure 10. The test results of the Buck Converter circuit to reduce the voltage

5.5 Current Sensor Circuit Test

This is a comparison of current sensor output based on test results and ACS712 current sensor data sheet. In the first experiment, at 0 amperes current, the output voltage of the current sensor was 2.5 volts. Refer to Figure 11 for a visual demonstration of the waveform from the test results of the current sensor.

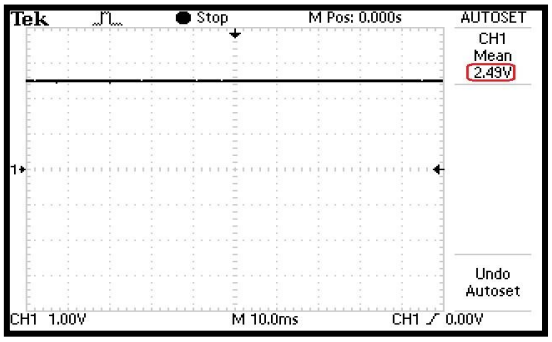


Figure 11. Example of ACS712 current sensor test results

Table 2 shows the test results of the current sensor with various current values.

Table 2. ACS712 current sensor test results table

No.	Input Current (ampere)	Output Voltage (volt)	Voltage based on Datasheet (volt)	Error (%)
1.	0	2,49	2,50	0.4
2.	0,1	2,50	2,51	0,4
3.	0,2	2,51	2,52	0,4
4.	0,3	2,52	2,53	0,4
5.	0,4	2,53	2,54	0,4
6.	0,5	2,54	2,55	0,4

Table 2 shows that the ACS712 current sensor has an average error of 0.4%. This is significantly lower than the maximum reading error of 1.5% stated in the datasheet, demonstrating that the sensor performs well.

5.6 Voltage Sensor Circuit Test

This test examines the functionality of the voltage sensor circuit in measuring voltage. The circuit is evaluated by inputting multiple voltage values to the sensor. Example results of the voltage sensor circuit when reading 11.9 volts are depicted in Figure 12.

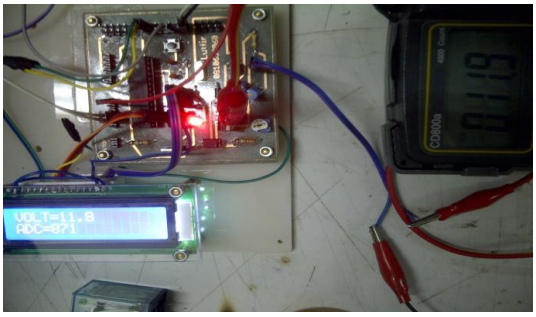


Figure 12. Test result of voltage sensor circuit

The results of the voltage sensor test are displayed in Table 3. The sensor exhibits a low average error rate of 0.13%, indicating its effective performance.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 12, 80 - 86, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1212.1012

No.	Voltage sensor (volt)	Voltage multimeter (volt)	Error (%)
1.	0	0	0
2.	1,2	1,2	0
3.	2,5	2,5	0
4.	3,2	3,2	0
5.	5	5	0
6.	11,8	11,9	0,8

Table 3. Voltage sensor test result table

5.7 Overall System Test

Testing the entire circuit is crucial to determine the success of each individual circuit block that has been incorporated into a single system. The system can be deemed successful if it can showcase a consistent output voltage of the alternator between 13-13.5 volts while subjected to varying alternator rotation speed conditions. Fig. 13 depicts the testing of the complete system.

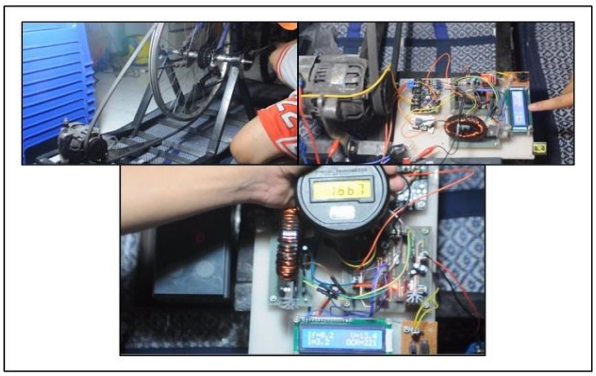


Figure 13. Overall system test

Table 4 displays the data of the comprehensive system test outcomes during alterations in rotational speed.

Table 4 Overall system test data

Rotation speed (rpm)	If (ampere)	Vout (volt)	Iout (ampere)	Torque (Nm)
826	0,3	11,3	1,7	0,207
1424	0,2	12,9	2,3	0,202
1568	0,1	13,1	2,5	0,201
1603	0,1	13,2	2,6	0,198
1668	0,2	13,4	2,8	0,218
1677	0,1	13,2	2,66	0,197
1798	0,2	13,2	3,2	0,220
1872	0,2	13,2	2,86	0,188
2006	0,1	13,3	2,82	0,172
2076	0,1	13,3	2,66	0,162
2196	0,1	13,2	2,81	0,160
2228	0,1	13,4	2,7	0,154
2360	0,1	13,2	3,1	0,166
2412	0,1	13,3	3,09	0,163
2462	0,1	13,5	3,19	0,167

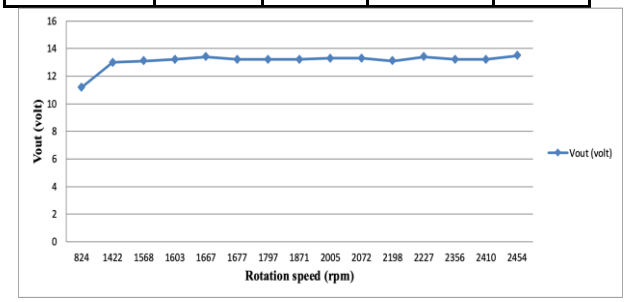


Figure 14. Graph of output voltage versus rotation speed

Figure 14 depicts a graph displaying the relationship between rotation speed and output voltage. It is observed that at high and low rotation values, there is little variation in Vout or it tends to remain constant due to a minimal difference in value. As an illustration, the Vout value remains the same at 13.2 volts for rotation values of 1677 rpm and 2356 rpm. This indicates that the device is functioning as intended - ensuring a stable output voltage with a value ranging between 13 - 13.5 volts.

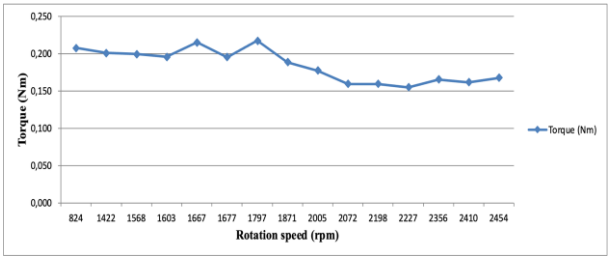


Figure 15. Graph of torque against rotation speed

Figure 15 illustrates the graph of torque generated in relation to rotation speed. Technical term abbreviations such as 'torque' should be defined when first used to improve clarity. The graph shows that the tool produces a relatively stable torque value from low to high rotation values, with only a small difference in value. Additionally, as the rotation speed increases, the torque value gradually decreases. Consistent and formal language throughout the text enhances its academic quality. This indicates that the tool functions as intended by producing an appropriate torque value without excessive increase. This allows the bicycle generator to require less initial power. This reduces the stress on the person pedaling the generator.

The following analysis aims to compare the output voltage and torque values generated with and without a charge controller in the overall system test.

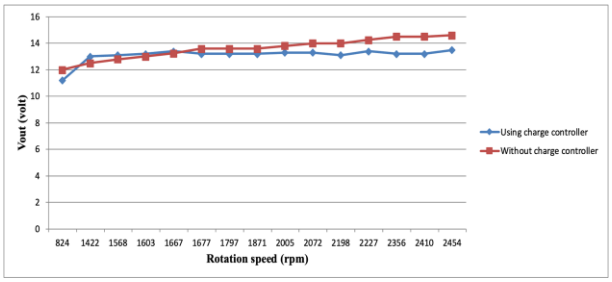


Figure 16. Comparison graph of output voltage using a charge controller and not using a charge controller

Figure 16 displays a comparative graph depicting the output voltage with and without a charge controller. The graph illustrates that as the rotation values increase, the Vout value steadily rises from 12 to 14.6 volts on the line without a charge controller. This indicates an unstable voltage and can quickly damage the battery if used for charging purposes. On the graph line utilizing a charge controller, the Vout value stabilizes between 13-13.5 volts, which guarantees safe battery charging

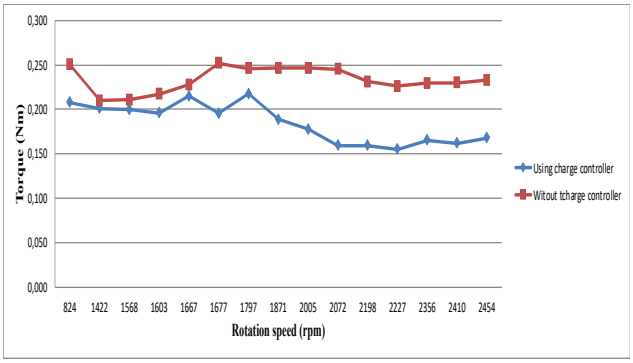


Figure 17 Comparison graph of output torque using a charge controller and not using a charge controller

Figure 17 displays a graph illustrating the correlation between torque output and rotational speed with and without a charge controller. The data reveals that torque output when using a charge controller is consistently lower compared to when one is not used, and tends to decrease. Consequently, initial pedaling of the bicycle power generator without the controller may require greater effort. The line graph shows that the torque value is consistently lower when using a charge controller compared to before its installation. In addition, the torque value generally decreases over time. Consequently, the use of a charge controller reduces the initial stroke power required for an initial bicycle power generator. This makes it easier for people who pedal the generator to experience less resistance.

6. CONCLUSION

After analyzing and testing the Charge Controller utilized in the DC House Bicycle Power Generator, we have concluded that it can effectively govern the field current entering the alternator using the Buck Converter topology. However, some circuit errors were observed in that configuration. Inappropriate component selection caused this issue. Nevertheless, the Buck Converter circuit remains functional, as evidenced by the steady output voltage and reduced torque value when compared to using the charge controller.

7. REFERENCES

[1] Allegro. 2011. ACS712. Massachusetts : Allegro Microsystems, Inc.

[2] Arifianto, Syaiful. 2010. RancangBangun Charger BateraiDengan Buck Boost Converter. Surabaya :PoliteknikElektronika Negeri Surabaya.

[3] Arifin, Zaenal. 2009. *Portable Solar Charger*. Surabaya :PoliteknikElektronika Negeri Surabaya.

[4] Atmel. 2011. ATmega8/ATmega8L, 8-bit AVR Microcontroller with 8 Kbytes in-System Programmable Flash. California: Atmel.

[5] B. E. Noriega, R. T. Pinto and P. Bauer, "Sustainable DC-microgrid control system for electric-vehicle charging stations", *Proceedings of 15th European Conference on Power Electronics and Applications*, pp. 1-10, 2013.

[6] Carter, Bruce and Thomas,R, Brown, 2001. *Handbook of operational Amplifier Applications*. Dallas:Texas Istruments.

[7] Chaidez, Jessica E. 2011. *DC House Modeling and Sistem Design*. San Luis Obispo: California Polytechnic State University.

[8] D. Salomonsson, L. Soder and A. Sannino, "An adaptive control system for a DC microgrid for data centers", *IEEE Transactions on Industry Applications*, vol. 44, no. 6, pp. 1910-1917, 2008.

[9] Hart, Daniel W. 2010. *Power Electronics.* New York : McGraw-Hill.

[10] Hasan, Umar. 2011. Sistem Charging Baterai pada Perancangan Mobil Hibyrd. Surabaya:PoliteknikElektronika Negeri Surabaya.

[11] Hayes, Brandon &Goguely, Louis. 2011. *Bicycle Power Generator Design for DC House: Off Grid Energy Solutions*. San Luis Obispo: California Polytechnic State University.

[12] J. Dominguez and I. Pinedo-Pascua, "GIS tool for rural electrification with renewable energies in Latin America", *Proceedings of International Conference on Advanced Geographic Information Systems & Web Services*, pp. 171-176, 2009.

[13] M. Seiragakis, R. T. Pinto, P. Bauer and M. Damen, "Control of an 80-kW wind turbine connected to a DC microgrid", *Proceedings of 15th European Conference on Power Electronics and Applications*, pp. 1-10, 2013.

[14] Majid, Ahmad Alvin. 2012. Prototype SistemPeneranganRumahDaruratuntukPemulihan pada Daerah PascaBencana. Surabaya :PoliteknikElektronika Negeri Surabaya.

[15] Mazda, Fraidoon. 1997. *Power Electronics Handbook 3rd Edition*.Oxford :Newnes.

[16] Rashid, Muhammad H. 2001. *Power Electronics Handbook*. London: Academic Press.

[17] Salam, Zainal. 2003. *Power Electronics and Drives*. Johor Bahru :UniversitiTeknologi Malaysia.

[18] T. Shibata, K. Sakai and Y. Okabe, "The design and implementation of an on-demand DC Grid in

Home", *IEEE/IPSJ 11th International Symposium on Applications and the Internet*, pp. 152-159, 2011.

INTERNATIONAL JOURNAL OFCOMPUTER APPLICATIONS TECHNOLOGYAND RESEARCH
(IJCATR)

PUBLISHER

ASSOCIATION OF TECHNOLOGY & SCIENCE

E-ISSN: 2319-8656

www.ijcat.com