**Cardholder Segmentation Using K-Means for Effective Marketing Strategies**

**Mike Rasell Carale Dago-oc1, Chris Angelu Bongcawil Jordan2**

Computer Science and Information Technology Department,

Bachelor of Science in Computer Science (BSCS)

Negros Oriental State University (NORSU), Negros Oriental, Philippines

|  |  |  |
| --- | --- | --- |
| **Article Info** |  | **ABSTRACT** |
| ***Article history:***  Received November 14, 2024  Revised January 21, 2025  Accepted January 21, 2025 |  | This study explores the application of K-Means clustering to segment credit card customers based on their demographic, behavioral, and transactional data. Using a dataset encompassing spending habits, payment frequencies, and credit usage, the analysis determined that two distinct customer groups provided the optimal segmentation, as validated through the Elbow Method, Silhouette Score, and Davies-Bouldin Index. The results revealed key patterns: a dominant segment of cost-conscious customers with moderate spending habits and a smaller, high-income group with premium spending behavior. These findings highlight opportunities for businesses to tailor marketing strategies, improve customer engagement, and optimize resource allocation. By demonstrating the scalability and practicality of K-Means clustering, this research provides a framework for leveraging raw data to derive actionable insights in customer analytics.  *This is an open access article under the* [*CC BY-SA*](https://creativecommons.org/licenses/by-sa/4.0/) *license.* |
| ***Keywords:***  Clustering  K-Means  K-Means Elbow Method  Principal Component Analysis  Silhouette Score  Davies-Bouldin Score |

1. **INTRODUCTION**

In the era of digital transformation, businesses have access to extensive data resources that, when utilized effectively, can provide valuable insights into customer behaviors and preferences [1]. Financial institutions, particularly credit card providers, rely on such data to enhance customer engagement, optimize product offerings, and mitigate risks [2],[3]. Clustering techniques have emerged as a cornerstone of customer analytics, enabling businesses to categorize their clientele into meaningful groups based on transactional and behavioral patterns [4].

Credit card usage generates rich datasets encompassing spending habits, payment frequencies, cash advances, and credit limits. These data points, when analyzed effectively, allow for the identification of customer segments, such as high-value users, infrequent spenders, or customers at risk of default [5],[6]. Traditional marketing and risk management strategies often fall short due to their generalized nature, emphasizing the need for data-driven, personalized approaches [7].

This study leverages clustering techniques, particularly K-Means, to segment credit card customers based on attributes such as purchases, payments, and account tenure. By uncovering latent patterns, the research aims to help credit card providers tailor services, improve customer retention, and maximize profitability [8]. This paper also explores how clustering can bridge the gap between raw transaction data and actionable business insights, setting a foundation for scalable and replicable customer analytics frameworks [9].

1. **LITERATURE REVIEW**

**2.1. Clustering**

Clustering is an unsupervised machine learning method that groups data points based on their similarities, aiming to maximize intra-cluster cohesion while ensuring inter-cluster separation [10]. Common algorithms include K-Means, DBSCAN, and Hierarchical Clustering, each suited to specific data structures and objectives [11].

Studies highlight the versatility of clustering in diverse domains. For instance, researchers applied DBSCAN to detect anomalies in financial transactions, demonstrating its strength in handling noise and outliers [12]. K-Means has been widely employed for customer segmentation in retail and finance, providing actionable insights into purchasing behaviors and payment patterns [13]. Hierarchical Clustering, often used in bioinformatics, facilitates the exploration of hierarchical relationships among data points [14].

**2.2. Customer Segmentation**

Customer segmentation divides a user base into groups with similar characteristics, such as spending behaviors, demographics, or engagement levels [15]. This technique allows businesses to design targeted marketing strategies, optimize resource allocation, and improve customer satisfaction [16].

In financial services, segmentation is particularly valuable. A study analyzing telecom customer data used K-Means to segment users based on call duration and data usage, achieving enhanced service delivery and reduced churn [17]. Another example involved clustering e-commerce customers into high-value and low-value groups to guide promotional strategies [18]. Such applications underscore the importance of segmentation in driving data-informed decision-making and operational efficiency [19].

**2.3. K-Means Clustering**

K-Means is a widely adopted clustering algorithm due to its simplicity and computational efficiency. It partitions data into a predefined number of clusters by iteratively updating centroids to minimize within-cluster variance [20].

Various adaptations of K-Means address its limitations. Mini-Batch K-Means enhances scalability for large datasets by processing data in smaller chunks, while Weighted K-Means assigns varying importance to features, making it suitable for customer segmentation tasks [21],[22].

Recent studies have demonstrated the utility of K-Means in analyzing financial datasets. For example, researchers applied Weighted K-Means to segment credit card users based on spending frequency and payment consistency, resulting in more effective resource allocation and personalized offers [23]. Another study highlighted Mini-Batch K-Means’ capability to handle extensive e-commerce datasets while maintaining clustering precision [24]. These advancements reaffirm K-Means’ pivotal role in clustering tasks across industries [25].

1. **METHODOLOGY**

This section provides an outline of the research methodology employed in this study.

**3.1. Hardware and Software**

The study was carried out on a system running Windows 10 with a 64-bit operating system. The system uses an Intel Core™ i5-7400 CPU with 16GB Random Access Memory (RAM). The researchers utilized Jupyter and Python version 3.9.15 as the primary programming language for data analysis and model implementation with the following libraries: Math, NumPy, Pandas, Seaborn, Matplotlib, and Scikit-Learn.

**3.2. Data Acquisition**

This research employs a dataset from Kaggle [26], a .csv file containing 1000 instances and 18 features that can be categorized as financial and behavioral metrics. Financial metrics include monetary values like balances, credit limits, purchases, and cash advances, while behavioral metrics are the frequencies of transactions, purchases, and balance updates, reflecting customer activity. This dataset provides the basis for evaluating the relationship between these elements, and clustering them can help identify patterns in consumer behavior.

**3.3. Data Pre-processing**

Data processing is an essential aspect of model development. Data acquired in their raw form contain noise and anomalies, which can affect the performance and training process of the model being schooled [27]. The researchers employed several techniques to clean the data, which included data normalization and fixing missing values:

Data normalization is a pre-processing technique primarily intended to manage numerical features and is applied to numerical features before the application of classification algorithms. Normalization is crucial to prevent the effect of certain features from being concealed by others, particularly when the ranges of the features are inconsistent [28].

A missing value is a datum that has not been stored or gathered due to issues like faulty sampling procedures, budgetary constraints, or limitations in the data collection process. Missing values are an inevitable aspect of data analysis and can present significant challenges for data practitioners. It is generated due to several reasons, including human mistakes, technical malfunctions, unavailable data, or outdated and inconsistent data [29].

**3.4. Principal Component Analysis (PCA)**

Since the dataset has 18 features PCA reduces the number of features or dimensions in the data into 2 features while retaining the most important patterns or variance in the dataset. The remaining features would be principal components 1 and 2, PCA achieves this by transforming the original features into new, uncorrelated variables called principal components. These components are linear combinations of the original features, ordered such that the first principal component captures the maximum variance in the data, followed by the second, and so on [30][31].

**3.5. K-Means Clustering**

K-means clustering was employed to partition the dataset into distinct groups based on consumer characteristics. The algorithm aims to minimize intra-cluster variance by iteratively adjusting cluster centroids and assigning data points to their nearest cluster. The Elbow Method was used to determine the optimal number of clusters by analyzing the Within-Cluster Sum of Squares (WCSS) values for different cluster counts. Mathematically, the objective function for K-Means is defined as [32]:

(1)

Where is the total number of clusters, the group of data points in the i-th cluster, a single data point, and the center of the i-th cluster.

**3.6. Silhouette Score**

To evaluate the quality of the clusters generated by the K-Means algorithm, the Silhouette Score was calculated. This metric measures how well each data point fits within its assigned cluster compared to other clusters. The score ranges from -1 to 1, where a value closer to 1 indicates that clusters are well-separated and cohesive. A score near 0 suggests overlapping clusters, and negative values indicate that points are assigned to the wrong clusters [33].

(2)

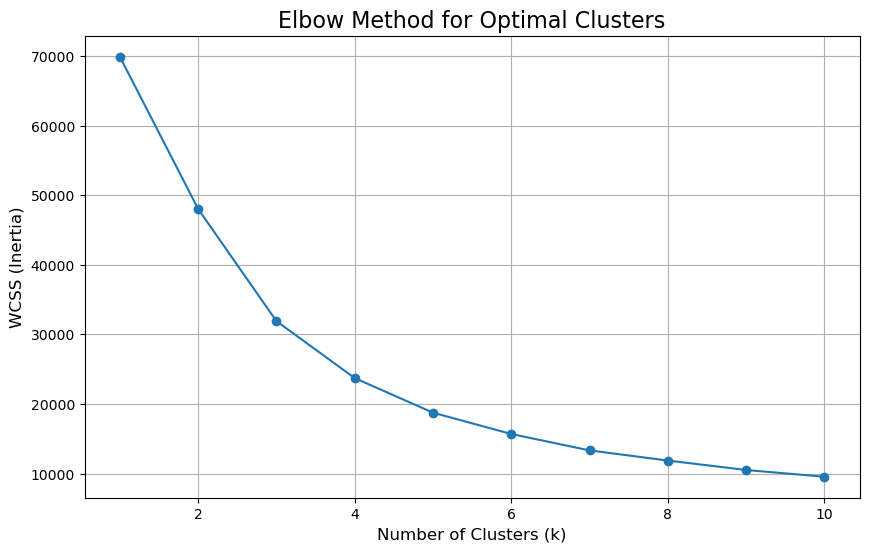
The Silhouette Score measures how well a data point fits into its cluster. It is calculated by , the average distance from the point to all other points in the same cluster, and b, the average distance from the point to all points in the nearest cluster.

**3.7. Davies-Bouldin Score**

To evaluate the quality of the clusters generated by the K-Means algorithm, the Davies-Bouldin Score was calculated. This metric measures the average similarity ratio between each cluster and its most similar neighboring cluster. Lower Davies-Bouldin Scores indicate better-defined clusters, where each cluster is compact and distinct from others [34].

(3)

Where represents the number of clusters, the average distance of all points in cluster to the centroid of cluster , and is the distance between the centroids of clusters and .

1. **RESULTS AND DISCUSSION**

**4.1. Elbow Method for Optimal Clusters**

**Figure 1.** Elbow Method

**3.5.2. Support Vector Machine**

Support Vector Machine (SVM) is a binary linear classifier. As a non-probabilistic supervised learning algorithm, it utilizes training data and employs a high-dimensional space to construct a set of hyperplanes for data classification. While only the features of test data are provided, the model is trained on the training data to predict the target values. For effective classification of problem instances, SVM relies on selecting the optimal hyperplane [7].

**3.6 Model Evaluation Metrics:**

Evaluation measures are metrics used to assess the results of an experiment [24]. In the context of classification models, different evaluation metrics are used to measure their output. In this study, the main performance evaluation metric is “Accuracy”. However, additional metrics such as recall, precision, f-measure, and confusion matrices are also used to supplement the evaluation of the model's performance. Each model identifies learner engagement levels when assessed using these metrics. A brief description of these metrics is provided below.

**Accuracy (AC)** is a common evaluation metric for classification models. It's calculated as the ratio of well-predicted samples to the total sample of prediction. For a balanced dataset, accuracy is a reliable measure of the model's performance [22].

(3)

In this equation, a true number that is positive is denoted by TP while a true number that is negative is denoted by TN, however, FN denotes a false number that is negative and FP denotes a false positive number [7].

**Precision (PRE)** is computed by dividing all the true positive samples by the sum of the predicted positive samples and predicted negative samples [7].

(4)

A high precision score indicates strong class predictions, while a low precision score reflects weak class predictions [25].

**Recall (RE)** is the ratio of correctly predicted positive results to all actual positive samples, also known as the detection rate. It's calculated by dividing the true positive samples by the sum of the positive samples [7].

(5)

**F1-score (FS)** is the mean value for recall and precision. It offers an indicator of mistakenly graded results [22]. It is regarded as the best metric for measuring the performance of models on an imbalanced dataset. It ranges from 0 to 1, with higher values indicating better model performance [7].

(6)

1. **RESULTS AND DISCUSSION**

This section provides the outcome of the analysis, achieved through running the models using 10-fold cross-validation, with accuracy as the primary metric for evaluating the performance of the models. Additionally, the performance of the models is further assessed using f1-score, recall, precision, confusion matrices, and misclassified images.

**4.1. Model Performance**

Table 1 shows the results of the two models after evaluating their performance using precision, accuracy, f1-score, and recall. Figure 1 provides a visualization of the results. Figures 2 and 3 illustrate the performance of two machine learning models through a confusion matrix, which are NB and SVM classifying X-ray images into two categories: PNEUMONIA and NORMAL. Figures 4 and 5 are the misclassified X-ray images of each model of NB and SVM respectively.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models | Accuracy | Precision | Recall | F1-Score |
| SVM | 0.9483 | 0.9528 | 0.9483 | 0.9480 |
| NB | 0.9103 | 0.9186 | 0.9103 | 0.9095 |

**Table 1**. Performance Metrics

In the evaluation of the two machine learning models using a 10-fold cross-validation, in Table 1 it can be observed that the performance of all the models accuracy ranges from 90% to 94%. The result shows that all the models can strongly identify whether X-ray images have pneumonia. And with a thorough analysis of the results, indicates that SVM provided the highest accuracy of 94.83%, with a precision of 0.9528, a recall of 0.9483, and an F1-score of 0.9480. The NB model, while performing well, showed slightly lower accuracy compared to SVM, with an accuracy of 91.03%, a precision of 0.9186, a recall of 0.9103, and an F1-score of 0.9095. These results suggest that SVM handles the dataset very well, using kernel functions like the Radial Basis Function (RBF) shows that SVM is a powerful model for this study. While NB performs relatively well, its performance is lower than the other models. Overall, the findings highlighted the effectiveness of all two models and emphasized their potential for application in automated pneumonia detection.

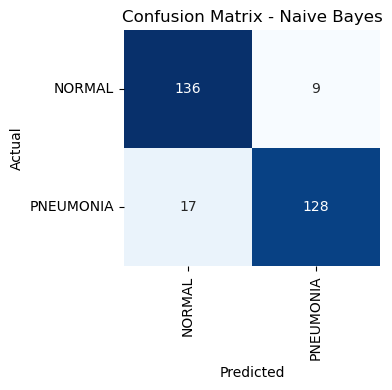
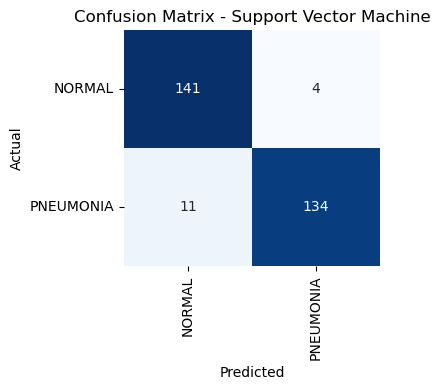
**Figure 1.** Graphical Performance Metrics

Figure 2 presents the performance metrics through a line chart which are: accuracy, precision, recall, and F1-score of SVM and NB.

SVM achieves consistent high across all metrics in both NORMAL and PNEUMONIA classifications, showing the best overall performance among the two models, with performance around 94–95%. The minimal variance between metrics indicates that SVM is a great model in this classification task.

NB shows lower performance than SVM but is still within a decent range for classification tasks, with metrics around 90–92%. Its simplicity and assumptions of feature independence hinder its ability to handle complex interactions in the data. Overall, While NB is effective for simpler patterns, NB is less suitable for handling overlapping or subtle features.

The graph demonstrates that SVM dominates overall across all metrics, SVM outperforms Naive Bayes, making it the better model for this task. This suggests that SVM handles the complex feature extraction from Inception V3 more effectively. While NB shows high precision but low recall, while it avoids false positives relatively well, it sacrifices recall, which makes it miss more pneumonia cases, a critical issue in healthcare tasks.

****

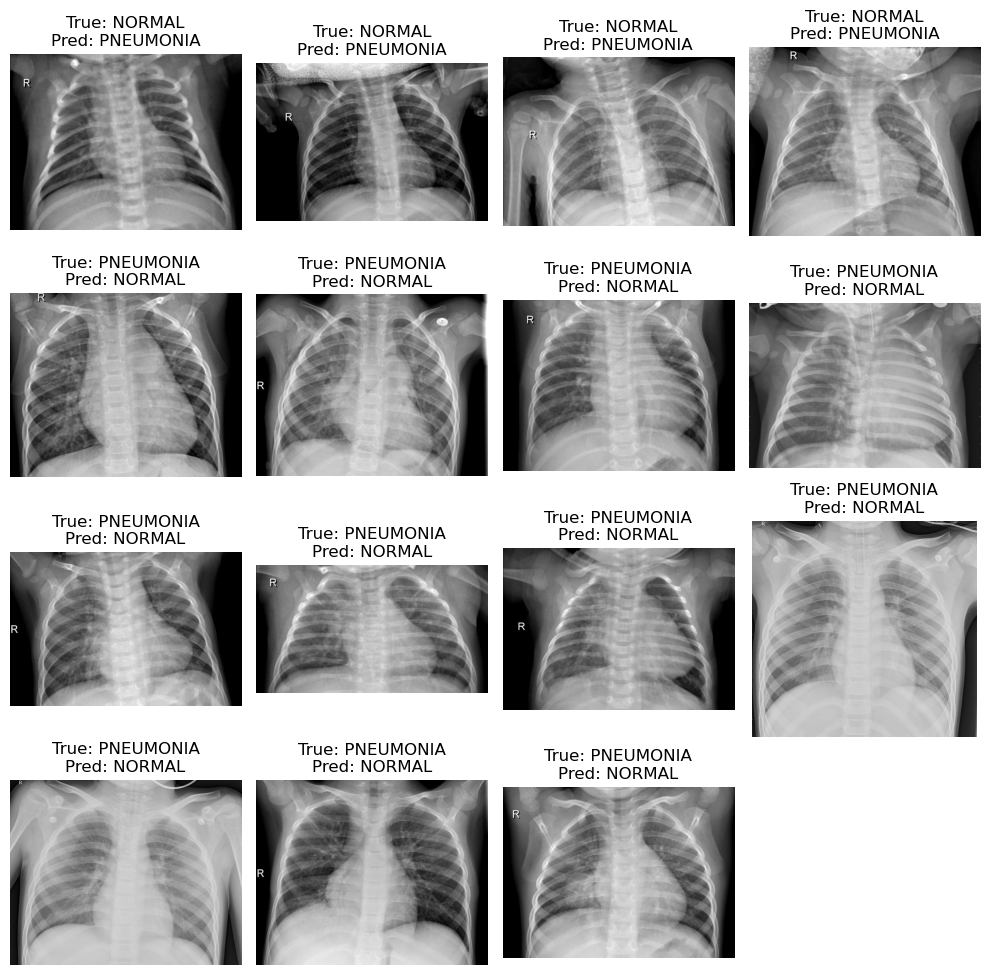
**Figure 2.** SVM Confusion Matrix

**Figure 3.** NB Confusion Matrix

In Figure 2, the confusion matrix for the SVM model shows high accuracy in detecting both NORMAL and PNEUMONIA cases. With only four false negatives, which is critical for pneumonia detection since missing a pneumonia case could have severe consequences. Overall, SVM shows slightly better overall performance compared to Naive Bayes in terms of reducing misclassification.

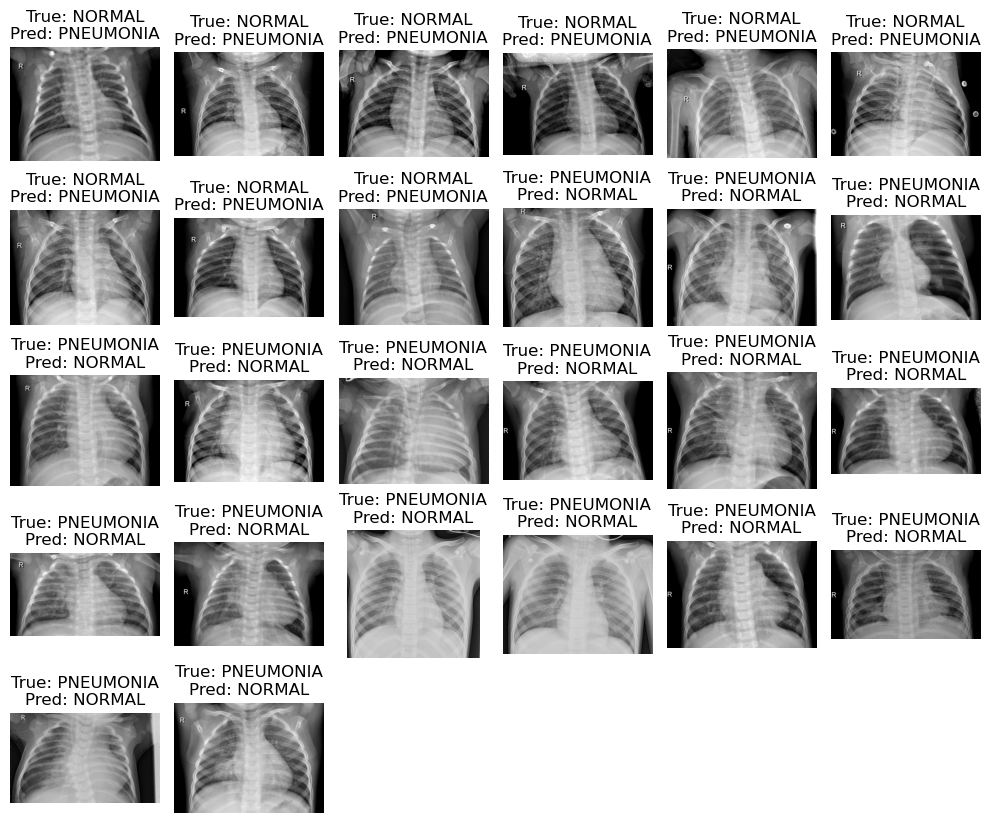
In contrast in Figure 4, NB shows slightly lower accuracy compared to SVM, as evidenced by higher false positives (17) and false negatives (9). Naive Bayes struggles more with classifying pneumonia cases accurately which results in higher false negatives, which is a critical area for improvement. This model may be more prone to noise or incorrect feature distribution assumptions, which could explain the performance gap.

Overall, SVM performs better, particularly with fewer false negatives and positives, making it more reliable for a sensitive task like pneumonia classification. Naive Bayes may not handle complex, non-linear relationships as effectively as SVM, which benefits from Inception V3’s extracted features and its ability to work well with high-dimensional spaces. If the goal is to minimize false negatives which is critical in healthcare, SVM is the better choice. Naive Bayes might be improved with additional preprocessing or feature engineering to better adapt to the data distribution.

**Figure 4.** SVM Misclassified Images

The first row shows chest X-rays where the true label is NORMAL, but the SVM model predicted PNEUMONIA. This suggests slight opacities or irregularities in the lung regions of the misclassified chest X-ray images that resemble pneumonia patterns. The rest of the rows are chest X-rays where the true label is PNEUMONIA, but the model predicted NORMAL. This indicates that the misclassified chest X-rays may lack clear and distinct patterns of pneumonia.

These misclassifications occur in cases where the X-rays are visually less distinct. The model also might be struggling with nuanced or borderline cases where the features extracted by Inception V3 are less pronounced. Overall, the misclassified images suggest that the model performs well for clear cases but struggles with borderline or less distinct cases.

**Figure 5.** NB Misclassified Images

The first and second half rows highlight cases where the model predicted NORMAL as a PNEUMONIA image. The rest of the rows showcase where PNEUMONIA images were predicted as a NORMAL image. It seems that the misclassifications occur for borderline or vague cases which is the same as the SVM model. The NB model might struggle with complex patterns due to its simplicity and assumptions of feature independence, which can lead to limitations in capturing nuanced or overlapping features in X-ray images.

Overall the performance of SVM and NB models suggests that while they handle clear cases effectively, they consistently struggle with less distinct or borderline cases. One of the probable reasons for this is a limitation in the features extracted by Inception V3, the reliance on these features might not fully capture subtle or complex patterns in the X-ray images. This affects NB due to its simplistic assumption of feature independence. In contrast to SVM, which can better model feature interactions demonstrate superior performance.

**4.2. Comparative Discussion**

The researchers in this study have evaluated the performances of the two models namely, Support Vector Machine (SVM) and Naive Bayes (NB) in classifying chest X-ray images if it has pneumonia. The models demonstrate varied strengths and weaknesses, highlighting their effectiveness in this image classification task.

The SVM model demonstrates strong overall performance in the image classification task, achieving consistently high metrics around 94–95%. However, it shows a slight bias toward missing some PNEUMONIA cases, which is evident in the confusion matrix and the misclassified images.

The NB model exhibits lower performance compared to SVM with metrics ranging between 90–94%. The NB model struggles with complex patterns due to its simplicity and assumptions of feature independence. These misclassifications highlight the limitations of using Naive Bayes for a challenging problem like medical imaging.

This study confirms that SVM is the most effective model for classifying chest X-ray images, achieving the best balance between precision and recall. NB, while effective for simpler tasks, falls short in comparison due to its inherent limitations.

1. **CONCLUSION**

Pneumonia is one of the common and fatal diseases in the world, and to treat it, people need access to radiologists. Unfortunately, not everyone has access to these professionals. This study proposes an ML model that can classify chest X-ray images to determine if they indicate pneumonia. Chest X-ray images were used as the dataset to train and test the model. This study also utilizes a Deep Convolutional Neural Network (DCNN), specifically Inception-V3, for feature extraction from the X-ray images. Afterward, the features are pre-processed through data normalization techniques, such as Min-Max Normalization and Z-Score Normalization, and then presented to the classifiers for further processing. Two ML classification algorithms SVM and NB were used to examine the efficiency of the proposed system. Numerous performance evaluation measures, including classification accuracy, precision, recall, and F1-score, were applied. From the experimental results, it is observed that SVM performed exceptionally well, attaining the highest classification accuracy of 94.83%. The findings of this study suggest that SVM is the most effective model for classifying chest X-ray images for pneumonia detection, offering a promising tool for educational and health institutions aiming to classify chest X-ray images for pneumonia detection. Meanwhile, NB provide alternative solutions with varying levels of performance.

**REFERENCES:**

[1] Our World in Data, "Pneumonia," 2025. [Online]. Available: https://ourworldindata.org/pneumonia. [Accessed: 12-Jan-2025].

[2] World Health Organization, "Pneumonia," WHO, Dec. 11, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/pneumonia. [Accessed: Jan. 12, 2025]

[3] World Health Organization, "Pneumonia," World Health Organization, [Online]. Available: https://www.who.int/health-topics/pneumonia#tab=tab\_1. [Accessed: Jan. 12, 2025]

[4] Mayo Clinic, "Pneumonia - Diagnosis and treatment," Mayo Clinic, Dec. 15, 2022. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/pneumonia/diagnosis-treatment/drc-20354210. [Accessed: Jan. 12, 2025].

[5] J. M. Chakaya, E. J. Carter, and P. C. Hopewell, "Pulmonary specialty training to improve respiratory health in low- and middle-income countries. Needs and challenges," Annals of the American Thoracic Society, vol. 12, no. 4, pp. 486–490, Apr. 2015, doi: 10.1513/AnnalsATS.201502-071PS.

[6] P. Mishra, S. Bag, and S. Nandi, "AI-driven computational analysis of COVID-19 genome reveals distinct association patterns," Journal of Biomedical Research & Innovation, vol. 13, no. 6, pp. 112–127, Dec. 2022. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC9799366/

[7] G. Theophilus and C. I. Eke, "Machine learning-based e-learners' engagement level prediction using benchmark datasets," Int. J. Appl. Inf. Syst., vol. 12, no. 41, Sept. 2023.

[8] EIT Health, "Machine learning in healthcare: Uses, benefits and pioneers in the field," EIT Health, 18-Sep-2024. [Online]. Available: https://eithealth.eu/news-article/machine-learning-in-healthcare-uses-benefits-and-pioneers-in-the-field/. [Accessed: 12-Jan-2025].

[9] Y. Muhammad, M. D. Alshehri, W. M. Alenazy, T. V. Hoang, and R. Alturki, "Identification of pneumonia disease applying an intelligent computational framework based on deep learning and machine learning techniques," Mobile Inf. Syst., vol. 2021, Art. no. 9989237, 20 pages, May 18, 2021. [Online]. Available: https://doi.org/10.1155/2021/9989237. [Accessed: 12-Jan-2025].

[10] I. S. Masad, A. Alqudah, A. M. Alqudah, and S. Almashaqbeh, "A hybrid deep learning approach towards building an intelligent system for pneumonia detection in chest X-ray images," Int. J. Electr. Comput. Eng. (IJECE), vol. 11, no. 6, pp. 5530-5540, Dec. 2021, doi: 10.11591/ijece.v11i6.pp5530-5540.

[11] T.B. Chandra and K. Verma, "Pneumonia Detection on Chest X-Ray Using Machine Learning Paradigm," in Proceedings of 3rd International Conference on Computer Vision and Image Processing, B. Chaudhuri, M. Nakagawa, P. Khanna, and S. Kumar, Eds., vol. 1022, Advances in Intelligent Systems and Computing, Springer, Singapore, 2020, pp. 15-22. doi: 10.1007/978-981-32-9088-4\_3.

[12] P. T. Mooney, Chest X-ray Images (Pneumonia), Kaggle. [Online]. Available: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia/data. [Accessed: 12-Jan-2025].

[13] C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” Technical report, Cornell University, New York, NY, USA, 2014.

[14] M. C. M. Pasague, "Classifying Garbage Images: A Comparative Study of Naive Bayes, SVM, and Logistic Regression Models," Bachelor of Science in Computer Science, Negros Oriental State University, 2025. [Online]. Available: cristinapasague27@gmail.com.

[15] C. Szegedy, V. Vanhoucke, S. Ioffe, and J. Shlens, “Rethinking the inception architecture for computer vision,” 2015, https:// arxiv.org/abs/1512.00567.

[16] P. Kragsterman, "The ultimate guide to preprocessing medical images: Techniques, tools, and best practices for enhanced diagnosis," About CMRAD, Aug. 28, 2024. [Online]. Available: https://about.cmrad.com/articles/the-ultimate-guide-to-preprocessing-medical-images-techniques-tools-and-best-practices-for-enhanced-diagnosis. [Accessed: Jan. 12, 2025].

[17] Keras, "InceptionV3," Keras API Documentation. [Online]. Available: https://keras.io/api/applications/inceptionv3/. [Accessed: Jan. 12, 2025].

[18] Pytorch Team, "Inception v3," \*PyTorch\*, 2021. [Online]. Available: https://pytorch.org/hub/pytorch\_vision\_inception\_v3/. [Accessed: Jan. 12, 2025].

[19] B. S. Oliveira, A. M. G. de Oliveira, and A. C. de Albuquerque, "The Effect of Preprocessing Techniques, Applied to Numeric Features, on Classification Algorithms’ Performance," Algorithms, vol. 6, no. 2, pp. 11, 2023. [Online]. Available: https://www.mdpi.com/2306-5729/6/2/11. [Accessed: Jan. 12, 2025].

[20] A. Choudhary, "Feature Scaling in Machine Learning: Normalization vs Standardization," Analytics Vidhya, Apr. 2020. [Online]. Available: https://www.analyticsvidhya.com/blog/2020/04/feature-scaling-machine-learning-normalization-standardization/. [Accessed: Jan. 12, 2025].

[21] E. F. Buraimoh, “Predicting Student Success Using Student Engagement in the Online Component of a Blended-Learning Course,” pp. 1–65, 2021.

[22] A. S. Hashim, W. A. Awadh, and A. K. Hamoud, “Student Performance Prediction Model based on Supervised Machine Learning Algorithms,” IOP Conf. Ser. Mater. Sci. Eng., vol. 928, no. 3, 2020, doi: 10.1088/1757- 899X/928/3/032019.

[23] P. Sun, M. D. Reid, and J. Zhou, “An improved multiclass LogitBoost using Adaptive-One-Vs-One,” Mach. Learn., vol. 97, no. 3, pp. 295–326, 2014, doi: 10.1007/s10994- 014-5434-3.

[24] C. I. Eke, A. A. Norman, and L. Shuib, Multi-feature fusion framework for sarcasm identification on twitter data: A machine learning based approach, vol. 16, no. 6

[25] H. Hassan, N. B. Ahmad, and S. Anuar, “Improved students’ performance prediction for multi-class imbalanced problems using hybrid and ensemble approach in educational data mining,” J. Phys. Conf. Ser., vol. 1529, no. 5, 2020, doi: 10.1088/1742-6596/1529/5/052041.