**Pneumonia Image Classification: A Comparative Study of Naive Bayes, Support Vector Machine, and Logistic Regression Models**

**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 10, 2025  Accepted January 10, 2025 |  | Pneumonia is a common and fatal disease, which is also one of the leading causes of death worldwide. One of the ways to diagnose this disease is through chest X-ray imaging, and with this, there is a need for an expert radiologist who possesses expertise and experience in the desired domain. According to the World Health Organization (WHO) report, about 2/3 of people in the world still do not have access to a radiologist to diagnose their disease. With this problem in mind, this study proposes a machine-learning classification model that diagnoses pneumonia disease efficiently and effectively. The study used a dataset of 290 X-ray images with 145 normal and 145 pneumonia images. The machine learning algorithms used are: Naive Bayes, Support Vector Machine, and Logistic Regression. The performance metrics of each model were done using 10-fold cross-validation. The predictive results show that the Support Vector Machine outperformed other models by attaining an accuracy of 94.83% with precision, recall, and f1-score values of 0.9528, 0.9483, and 0.9480 respectively. LR and NB also demonstrated strong performance, with LR having an accuracy of 94.14% while NB has an accuracy of 91.03%.  *This is an open access article under the* [*CC BY-SA*](https://creativecommons.org/licenses/by-sa/4.0/) *license.* |
| ***Keywords:***  Pneumonia  Naïve Bayes  Support Vector Machine (SVM)  Logistic Regression  Inception-V3  Accuracy  Precision  Recall  F1-Score |

1. **INTRODUCTION**

Pneumonia is one of the leading causes of death worldwide [1], it can affect everyone anywhere, especially children [2]. Pneumonia is a form of acute respiratory infection that is most commonly caused by viruses or bacteria targeting the lungs and it can cause mild to life-threatening illness in people of all ages. When an individual has pneumonia, the alveoli are filled with pus and fluid, which makes breathing painful and limits oxygen intake [3].

Identifying if a person has pneumonia requires general practitioners or doctors specializing in lung conditions [4]. Sadly some countries have not kept pace with the overall economic growth, and their health systems remain weak, which includes doctors who specialize in lung conditions [5]. This lack of general practitioners leads to increased workloads for existing specialists, potential delays in diagnosis and treatment, and overall strain on healthcare systems [6].

This study aims to make a machine-learning model that can classify X-ray images if it has pneumonia. With this tool, it can potentially automate diagnosis which can reduce the time doctors spend on diagnosing pneumonia cases. This allows doctors to focus more on treatment rather than performing or interpreting tests. Additionally, this can also be used as a training and education tool for medical students to learn how to interpret chest X-rays. However, the model is intended solely as a support tool, not a replacement for human judgment.

The rest of this paper is structured thus: Section 2 provides a detailed review of recent literature in the field of pneumonia image classification, Section 3 outlines the methodology employed in this study, and further provides a brief description of the machine learning algorithms used in this investigation. While Section 4 presents the results and discussion, Section 5 concludes the research, and Section 6 offers recommendations.

1. **LITERATURE REVIEW**

Machine learning (ML), once a theoretical concept, has rapidly evolved into a transformative force shaping nearly every aspect of our lives [7]. This technology is also slowly integrated into healthcare, offering innovative solutions for diagnosis, treatment, and operational efficiency [8]. For instance, ML models are being utilized to detect pneumonia from chest X-rays [9]. This section is a summary of existing related papers previously published about detecting pneumonia from chest X-rays that also use Naive Bayes (NB), Support Vector Machine (SVM), and Logistic Regression (LR).

The researchers in [9] tested the predictive performance of 6 machine learning models. In this review, we focus on Naive Bayes (NB), Support Vector Machine (SVM), and Logistic Regression (LR). The other models include Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), and AdaBoost (AB). Their dataset consists of 5,856 X-ray images, which include 1,583 normal cases and 4,273 pneumonia cases and is split into 70% for training and 30% for testing. For image embedding they used 5 Deep Convolutional Neural Networks (DCNN) as feature extractors. However, this study will only focus on Inception-V3, the same as the one used in this study. The other DCNNs are AlexNet, SqueezeNet, VGG16, and VGG19. ANN delivered the best overall performance with a 97.19% accuracy, closely followed by LR at 97.08%. Meanwhile, SVM with a linear kernel achieved an accuracy of 85.02%.

Researchers in [10] used five models including, Convolutional Neural Network (CNN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Random Forest (RF), and Softmax. In image embedding they extracted the features from the fully connected (FC) layer of a modified CNN architecture using transfer learning. The FC layer generated feature vectors representing each class. The CNN model used in this study was pre-trained on Optical Coherence Tomography (OCT) images before being adapted to the pneumonia detection task. This differs from directly using models like Inception V3. Their dataset is comprised of 5,852 anterior-posterior chest X-ray images, which include 1,581 normal and 4,271 pneumonia cases, and is split into 70% for training, 15% for validation, and 15% for testing. However, in this review, we focus on the SVM, which achieved an accuracy of 99.61%. Among the models, Softmax slightly outperformed SVM with an accuracy of 99.72%.

While Researchers in [11] employed several models, including Multilayer Perceptron (MLP), Random Forest (RF), Sequential Minimal Optimization (SMO), Logistic Regression (LR), and Classification via Regression. The research focused on pixels in lungs segmented Region of Interest (ROI) that are more contributing toward pneumonia detection than the surrounding regions, thus the features of lungs segmented ROI confined area are extracted. It also utilized a total of 412 chest X-ray images containing 206 normal and 206 pneumonic cases from the ChestX-ray14 dataset. But this study will only focus on the LR model which got an accuracy of 95.63% which is the highest among the five.

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, PIL, OS, Scikit-Learn, Matplotlib, Seaborn, TensorFlow, and Keras.

**3.2. Data Acquisition**

This research employs a dataset from Kaggle [12], a file with 5856 .jpeg files which are anterior-posterior chest X-ray images. The images are carefully chosen from retrospective pediatric patients with age groups between 1 and 5 years from Guangzhou Women and Children’s Medical Center. The image has 2 normal categories pneumonia and normal which are further divided for testing, training, and evaluation [12]. Because of hardware limitations, this study uses 290 images with 145 normal and 145 pneumonia images.

**3.3. Image Embedding**

The main goal and objectives of the proposed system are to diagnose and make a tool that can identify whether a person has pneumonia through chest X-ray images. This study uses a Deep Convolutional Neural Network (DCNN) which is Inception-V3, for feature extraction of the X-ray images.

Inception models were developed by a researcher [13] for the first time in 2014. The structures of inception models and the conventional CNN model are different because they are inception blocks which means lapping the same input tensor with multiple filters and concatenating their results. In 2015, a researcher [15] proposed a new version of the inception models named Inception-V3, an improved version of the previous versions of inception models which are Inception-V1 and Inception-V2, and possesses 24M parameters. The advancement in Inception-V3 was as follows: first, it factorizes the “n × n” convolution into asymmetric convolutions which are 1 × n and n × 1. Second, it factorizes the 5 × 5 convolutions into two 3 × 3 convolutions and lastly, it replaces 7 × 7 convolutions to a series of 3 × 3 convolutions. It consists of a block of convolutional layers arranged in a parallel manner and each layer consists of different sizes of filters 1 × 1, 3 × 3, and 5 × 5, respectively. Furthermore, 3 × 3 max pooling is also performed. The outputs are concatenated and sent to the next inception module [9].

**3.4. 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 [15]. In addition to this when preparing input data for a DCNN, it's essential to adhere to specific requirements regarding input shape, color channels, and preprocessing steps to ensure optimal performance [16]. The researchers employed several techniques to clean the data, which included image pre-processing and data normalization:

Image pre-processing is an essential part of image classification to ensure that input images align with the model's expectations. Inception-V3 requires input images in 299x299 resolution and should also have three color channels (red, green, and blue) [17][18]. And since the datasets used are X-ray images they only have 1 color channel (black and white) and inconsistent high resolution. With this in mind, the images are turned to have three color channels, and it is also resized to have a resolution of 299x299. After this, the image is preprocessed using TensorFlow Keras API preprocess\_input so that the images align with the model's training conditions.

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 [19]. After feature extraction using Inception-V3, the features are normalized to improve model training [20]. The normalization techniques used are Min-Max Normalization for Logistic Regression (LR) and Naive Bayes (NB), and Z-Score Normalization for the Support Vector Machine (SVM).

**3.5. Machine Learning Algorithms**

This section focuses on the machine learning classification models utilized in this study. After image and data pre-processing, the machine learning workflow progresses to the model training stage, where an algorithm is taught to learn from data and produce predictions. This algorithm is specifically responsible for the classification of whether an X-ray image has pneumonia. To find the best classifier for engagement level prediction, several classifiers, including Naive Bayes (NB), Support Vector Machine (SVM), and Logistic Regression (LR) were tested through a variety of tests.

**3.5.1. Naïve Bayes**

The Naïve Bayes algorithm has its foundation rooted in the Bayes theorem by Thomas Bayes. One of the strengths of this model is its ability to handle missing values. And Unlike other models, Naïve Bayes conserves processing and training time [21][7]. The term ‘naive’ is used due to this algorithm's uncertain independence. With this, researchers in [22] stated that with this ability it's able to converge quicker when compared to several others.

(1)

Where X is the training set of attributes and Y is the given class [7].

**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].

**3.5.3. Logistic Regression**

Logistic regression is often used to analyze and describe the relationship between entities when the outcome has only two possible options, such as 'Yes' or 'No,' and a set of predictive entities [23]. Logistic Regression calculates the odds of several classes employing a boundary rationality distribution as depicted in the expression below [7].

(2)

**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, and confusion matrices.

**4.1. Model Performance**

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models | Accuracy | Precision | Recall | F1-Score |
| SVM (rbf) | 0.9483 | 0.9528 | 0.9483 | 0.9480 |
| LR | 0.9414 | 0.9447 | 0.9483 | 0.9412 |
| NB | 0.9103 | 0.9186 | 0.9483 | 0.9095 |

**Table 1**. Performance Metrics

In the evaluation of the three 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 91% 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 LR model also performed well, it showed slightly lower accuracy compared to SVM with an accuracy of 94.14%, a precision of 0.9447, a recall of 0.9483, and an F1-score of 0.9412. The NB model, while performing well, showed slightly lower accuracy compared to the other two models, with an accuracy of 91.03%, a precision of 0.9186, a recall of 0.9483, 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. LR also performed well but this model is simpler and linear so it might struggle slightly with more complex decision boundaries compared to SVM with the RBF kernel. While NB performs relatively well, its performance is lower than the other models. Overall, the findings highlighted the effectiveness of all three models and emphasized their potential for application in automated pneumonia detection.

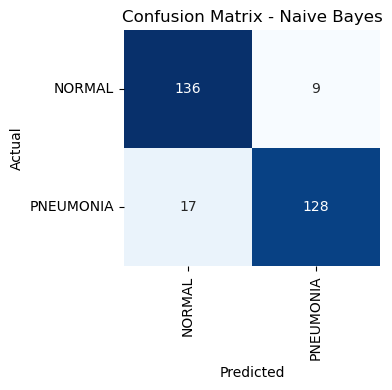
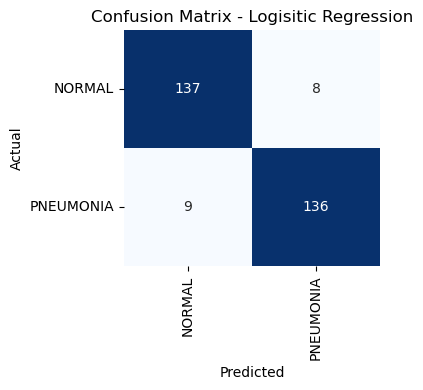
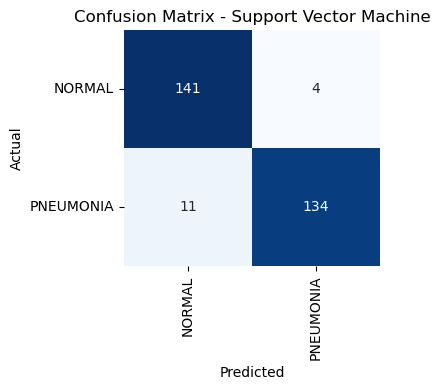
**Figure 2.** Graphical Performance Metrics

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

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

LR Performs slightly below SVM in most metrics but still exhibits stable and consistent performance, with values close to 94%. This indicates that Logistic Regression is also a reliable model for this task, though it is marginally outperformed by SVM.

NB exhibits significantly lower performance, with metrics around 90–94%. There is also a spike in recall, which indicates that NB is particularly good at identifying positive cases but at the expense of misclassifying negative ones.

****The graph demonstrates that SVM outperforms LR and NB in all performance metrics except for recall which all the models have the same value. This makes SVM the best model for identifying pneumonia in chest X-ray images. LR while slightly less effective than SVM, still provides strong results and is a viable alternative. NB on the other hand is less balanced but excels in Recall. NB might still be useful. However, its lower Precision and F1-Score indicate more false positives, which could be problematic. In conclusion, SVM provides the best overall performance, followed by LR. NB, while less balanced, could be valuable in scenarios where Recall is critical.

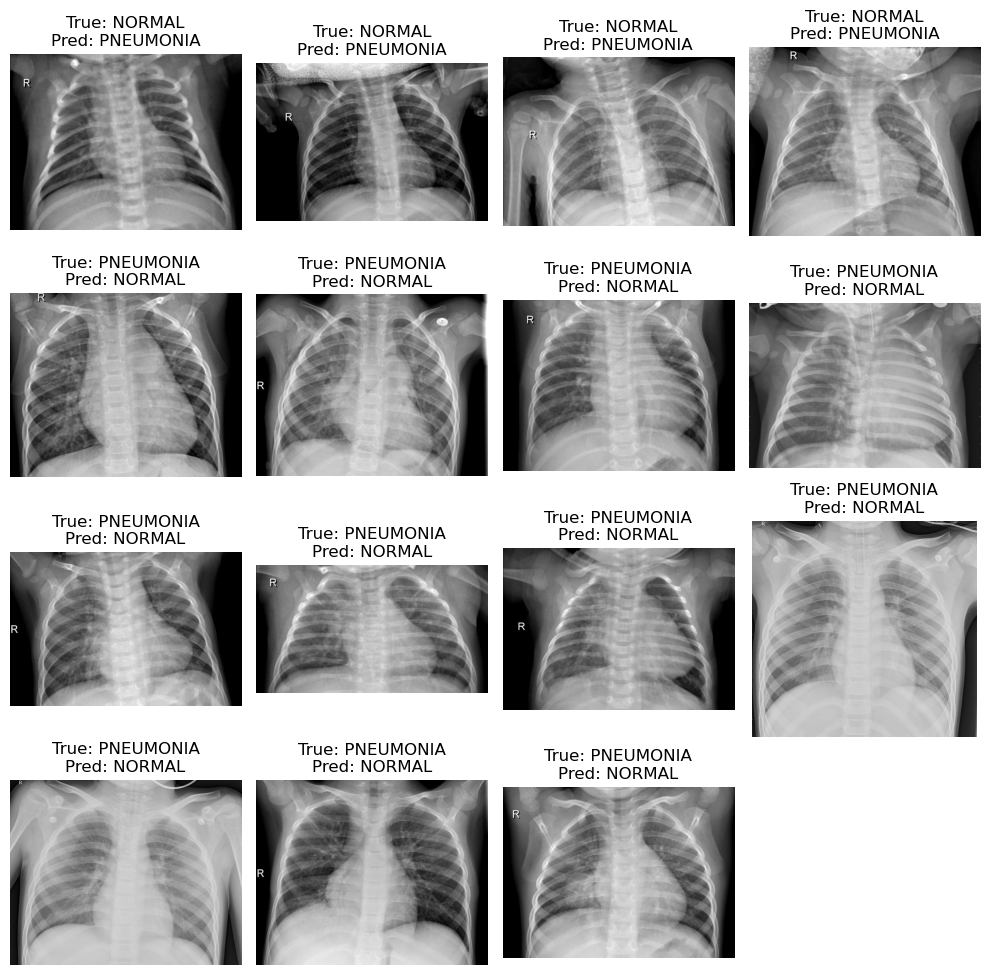
**Figure 3.** Confusion Matrices

The confusion matrix for the SVM model shows that it performs well in classifying both NORMAL and PNEUMONIA cases, with a slight tendency to miss some PNEUMONIA cases. Overall, the model shows strong predictive performance across all two classes.

LR also performed well in both NORMAL and PNEUMONIA classification tasks, with strong precision and recall for both categories. However, it has a bias towards misclassifying a few NORMAL cases as PNEUMONIA and missing some PNEUMONIA cases. This could indicate that LR linear decision boundaries may not capture all complexities in the data compared to SVM.

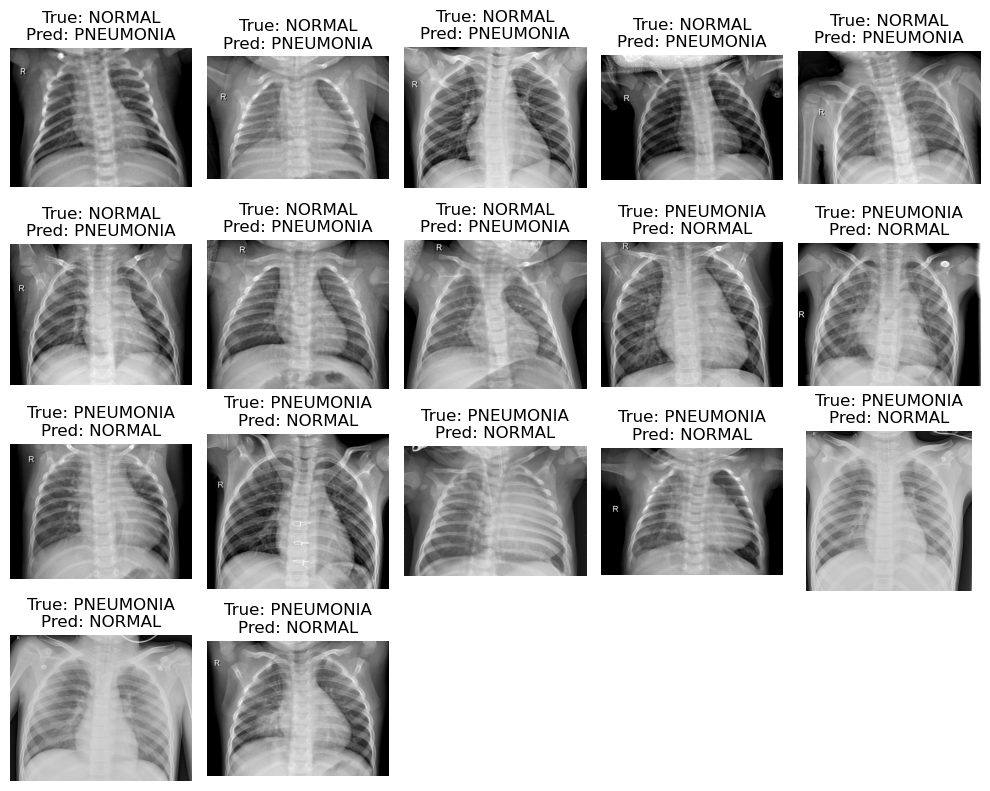
In contrast, NB still performs well in NORMAL and PNEUMONIA classification tasks, but it misclassifies more PNEUMONIA cases as NORMAL. There is a slightly higher precision for the NORMAL class compared to the PNEUMONIA class, but recall for PNEUMONIA is slightly better. Overall NB demonstrates good overall performance with an accuracy of 91.03%.

In general, SVM is the most effective model in identifying if a chest X-ray image has pneumonia, followed by LR, and lastly NB.

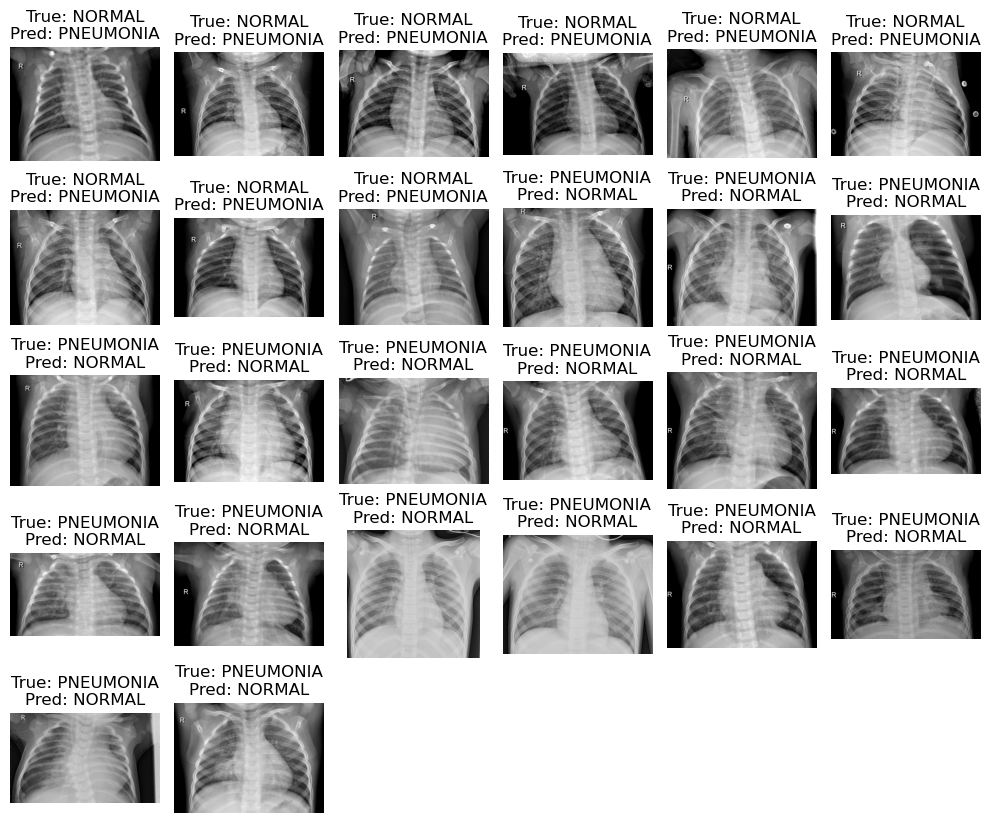
**Figure 4.** SVM Misclassified Images

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.** LR Misclassified Images

The first and second half rows represent cases where the true label is NORMAL, but the model incorrectly classified them as PNEUMONIA. The remaining rows represent cases where the true label is PNEUMONIA, but the model predicted NORMAL. Both SVM and LR models show similar struggles with borderline or less distinct cases. LR being a linear classifier is limited in capturing complex feature interactions, which may explain some of the false negatives.

**Figure 6.** 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 and LR models. 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, LR, 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, SVM and LR, which can better model feature interactions, demonstrate superior performance.

**4.2. Comparative Discussion**

The researchers in this study have evaluated the performances of the three models namely, Support Vector Machine (SVM), Logistic Regression (LR), 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.

LR demonstrates stable and consistent performance across metrics, with values close to 94%, making it a reliable model for this classification task. However, it performs slightly below the SVM model, likely due to its inability to handle complex, non-linear patterns as effectively as SVM. Overall LR is a dependable and robust model for this classification task, providing near-optimal performance. However, its limitations in handling complex feature interactions make it less effective than SVM for edge cases or borderline examples.

The NB model exhibits lower performance compared to SVM and LR, with metrics ranging between 90–94%. Despite its simplicity, NB shows certain strengths, particularly in the recall metric, which indicates its effectiveness at identifying positive cases (PNEUMONIA). However, this comes at the cost of higher false positives, reducing its precision for the NORMAL class.

This study confirms that SVM is the most effective model for classifying chest X-ray images, achieving the best balance between precision and recall. LR provides near-optimal performance and is a robust alternative, though slightly less capable of handling complex patterns. 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. Three ML classification algorithms—SVM, LR, 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, LR and NB provide alternative solutions with varying levels of performance.

1. **RECOMMENDATION**

The researchers in the study suggest several recommendations to enhance the scope and application of this study. Image pre-processing techniques are a crucial aspect of image-processing applications, so methods such as Contrast Limited Adaptive Histogram Equalization (CLAHE) can be explored to improve visual information quality by removing noise, increasing contrast, and deleting high or low frequencies. Other machine learning models, like Convolutional Neural Networks (CNN), can also be explored to improve performance. Additionally, other Deep Convolutional Neural Networks (DCNN), such as AlexNet, SqueezeNet, VGG16, and VGG19, can be utilized to extract useful features from the images. Lastly, due to hardware limitations, the datasets used to train the model were limited. In the future, incorporating more chest X-ray images is recommended to achieve better model 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.