

OPTIMIZED ENSEMBLE MACHINE LEARNING BASED FRAMEWORK FOR EARLY DETECTION OF LIVER DISEASE

Anita Meghta

Research Scholar

Department of Computer Science & Engineering,
Bahara University, Solan, HP, INDIA

Email: anitamegha@gmail.com

Dr. Kapil Sethi²

Professor

Department of Computer Science & Engineering
Bahara University, Solan, HP, INDIA

Email: kapil.7sethi@gmail.com

Abstract: The proposed framework integrates multiple machine learning algorithms, including decision trees, support vector machines, and neural networks, to form an ensemble model that leverages the strengths of each individual technique. By employing advanced optimization techniques such as grid search and genetic algorithms, the framework fine-tunes hyperparameters to achieve optimal performance. The model is trained and validated on a comprehensive dataset comprising clinical and biochemical data of patients, ensuring robustness and generalizability. Through a thorough examination of classification performance, the proposed ensemble model's ability to distinguish between liver disease and non-liver disease instances is scrutinized in comparison to existing machine learning methodologies. The performance metrics, including accuracy, recall, precision, and F1-score, is provided, offering a comprehensive understanding of the ensemble model's proficiency in liver disease diagnosis. The ensemble voting classifier emerges as the top-performing model, exhibiting exceptional accuracy of 97% alongside superior precision, recall, and F1-score, each hovering around 98%. These findings underscore the effectiveness of ensemble techniques in enhancing diagnostic accuracy and advocate for the adoption of the ensemble voting classifier in real-world applications requiring precise liver disease diagnosis.

Keywords: Ensemble machine learning, Liver disease diagnosis, Performance metrics, Ensemble voting classifier, Diagnostic accuracy

1. INTRODUCTION

The liver, the largest internal organ in the body, weighs approximately 1.6 kg (3.5 pounds) and plays a crucial role in maintaining overall health. It processes blood from the digestive system, synthesizes essential proteins for blood clotting and other functions, and metabolizes carbohydrates into glucose to regulate blood sugar levels. Unique among visceral organs, the liver has the remarkable ability to regenerate itself. Healthy liver function is vital, while liver damage can be life-threatening. Liver diseases can be hereditary or result from infections, alcohol use, or environmental factors, leading to conditions such as hepatitis A, B, and C, autoimmune diseases, and certain cancers. Over recent decades, liver diseases like cirrhosis, viral hepatitis, and fatty liver disease have become significant global health concerns, contributing to increased morbidity and mortality [1]. The liver's essential physiological roles underscore the need for a proactive approach to understanding, diagnosing, and treating liver disorders, which often remain undetected due to the liver's capacity to function despite damage. Early detection is crucial for improving outcomes, as many liver diseases are asymptomatic until advanced stages, limiting treatment options. Technologies that enhance diagnostic accuracy are invaluable, enabling physicians to make informed decisions and classify liver diseases efficiently. Liver diseases not only affect individual health but also impose substantial social and economic

10.48047/jocaaa.2024.33.08.190

burdens [2]. They reduce quality of life and strain healthcare systems due to the high costs of long-term care, medical treatments, and hospitalizations. Addressing the growing prevalence of liver diseases is essential for societal well-being and the sustainability of healthcare resources. Early detection facilitates timely medical interventions, allowing healthcare professionals to manage diseases more effectively and prevent progression. Therefore, developing reliable early diagnosis tools is critical for improving treatment outcomes and patient prognosis.

By harnessing the collective power of different models, it reduces errors and biases present in individual models, leading to more precise and dependable predictions. Ensemble techniques such as CatBoost, Gradient Boosting, and Random Forest are extensively evaluated, surpassing conventional diagnostic methods. The study advocates for collaboration between data scientists and healthcare professionals to implement these models in decision support systems. This method is particularly effective in handling data uncertainties and has proven its efficacy across diverse domains, including healthcare, economics, manufacturing, and bioinformatics. In predictive modeling, ensemble learning techniques have gained prominence for their ability to improve prediction accuracy [3-5]. Among these models, the tree classifier coupled with Random Forest demonstrates superior performance, achieving the highest accuracy rates of 91.82% and 86.06% in detecting liver disease. This underscores the practical feasibility of the proposed strategy in real-world healthcare settings [6-7].

2. LITERATURE REVIEW

This study introduces an advanced ensemble learning framework combined with improved preprocessing techniques to forecast liver diseases. The proposed framework enhances prediction and multivariate imputation for missing data. Among these models, the combination of the decision tree classifier and Random Forest (RF) demonstrated the highest accuracy, detecting liver disease with 91.82% and 86.06% precision, proving the approach's effectiveness in real-world applications [8]. Additionally, other studies emphasize innovative ensemble models like CatBoost and multilayer perceptron (MLP), showcasing superior diagnostic accuracy for liver-related diseases, such as Hepatitis C, with performance metrics reaching up to 95.59% accuracy [9][10]. Furthermore, a comparison of various machine learning models revealed that the Voting classifier excelled in accuracy and recall after applying SMOTE and cross-validation techniques [11]. Research on predicting AATD-LD disease progression using UK Biobank data and a stacking ensemble method also demonstrated superior performance compared to individual models [12].

Table 1. Review of literature for liver diseases prediction

Ref No.	Methodology Used	Main Idea
[20-21]	Multi-categorical performance evaluation using machine learning techniques	Develops a performance evaluation system based on multiple factors, using machine learning to assess algorithm effectiveness across different categories.
[22-23]	Convolutional Neural Network (CNN) for image classification	Uses CNN to predict student understanding in classrooms by analyzing facial images, showcasing CNN's effectiveness in image classification tasks.

[24]	AI-based diagnostic models for early detection of breast cancer	Applies artificial intelligence models to early breast cancer diagnosis, emphasizing the importance of early detection using AI and machine learning.
[25]	Predictive risk analysis using machine learning	Analyzes the epidemiology of occupational injuries using predictive models to identify risks and patterns, offering insights into risk prediction methods.
[26]	Machine learning model for predicting chemical harmfulness	Develops a model to predict chemical harm to staff and guests in the hospitality industry, demonstrating the application of machine learning in risk management.
[27]	Predictive models for assessing mental health and employment uncertainty	Focuses on the mental health effects of COVID-19 on hospitality workers, using predictive models to analyze employment uncertainty and its impact.
[28-29]	Machine learning in hospitality industry (review-based study)	Reviews the prospects and challenges of machine learning adoption in the hospitality industry, emphasizing the importance of predictive models in business operations.
[30-31]	Review of sustainable housekeeping practices using machine learning	Focuses on the importance of sustainability in hospitality management, with machine learning-based reviews to optimize housekeeping and hygiene programs.
[32]	AI impact analysis on educational challenges	Investigates the social and cultural challenges in education and how AI can address these, showcasing AI's potential in decision-making environments with multiple variables.
[33]	AI-driven decision-making models in education	Explores the role of artificial intelligence in decision-making within higher education, applying predictive models for effective educational outcomes.
[34]	Machine learning-based computational methods for malaria diagnosis	Reviews machine learning techniques in diagnosing malaria, emphasizing computational approaches to improve diagnostic accuracy in medical applications.
[35-36]	Deep learning models for plant disease prediction	Uses deep neural networks to predict apple diseases, showing how deep learning can enhance the accuracy of plant disease identification and prediction.

3. BACKGROUND

2.1 Logistic Regression

10.48047/jocaaa.2024.33.08.190

Machine learning (ML) methods like logistic regression are transparent in their operations; their mechanisms are well-understood. In binomial logistic regression, the outcome is a binary 0 or 1, and the probability ppp of obtaining a 1 is explicitly known, rooted in the Bernoulli distribution. Logistic regression tackles problems using a modified linear regression approach, where the sigmoid activation function transforms it into a single-layer neural network (NN). This transformation enables the estimation of weight parameters through minimizing cross-entropy loss via stochastic gradient descent (SGD). Instead of the sigmoid function, SoftMax ensures outputs remain as a probability distribution. The loss function shifts to log-likelihood loss under SoftMax for multinomial logistic regression scenarios [13-15].

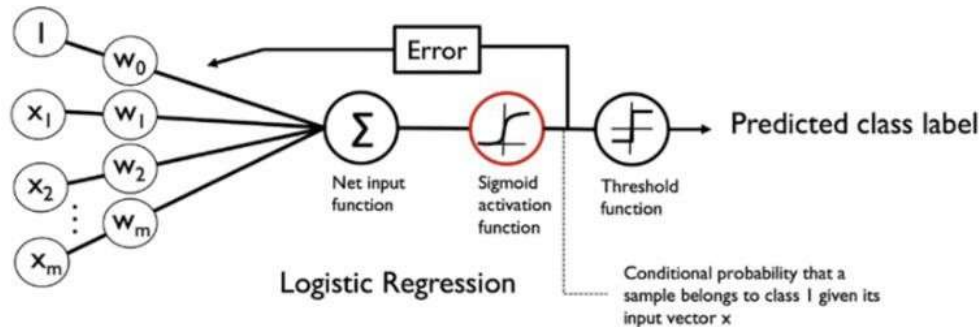


Figure 1. Architecture of a Logistic Regression Model

2.2 Naïve Bayes Algorithm

Naïve Bayes stands as a prime example that simplicity often leads to effectiveness in Machine Learning. Despite the advancements in the field, this algorithm continues to shine due to its straightforwardness, speed, accuracy, and reliability. Particularly acclaimed in natural language processing (NLP), Naïve Bayes leverages Bayes' Theorem, a foundational principle in probability theory, to make probabilistic predictions. It excels in various classification tasks by calculating conditional probabilities efficiently. This theorem underpins the approach ability to classify data by evaluating the likelihood of a particular outcome given certain observed features [16-17]. This article aims to provide a comprehensive meaning of Naïve Bayes, clarifying its core principles and applications, thus ensuring clarity and eliminating ambiguity in its implementation across diverse domains.

2.2 Support Vector Machines (SVM)

Support Vector Machines (SVM) have demonstrated remarkable effectiveness across various two-class classification tasks, such as identifying handwritten digits, detecting faces in images, and categorizing text. This algorithm's exceptional learning capabilities have propelled it to the forefront of machine learning research [16], [18-19]. SVMs are particularly well-suited for binary classification.

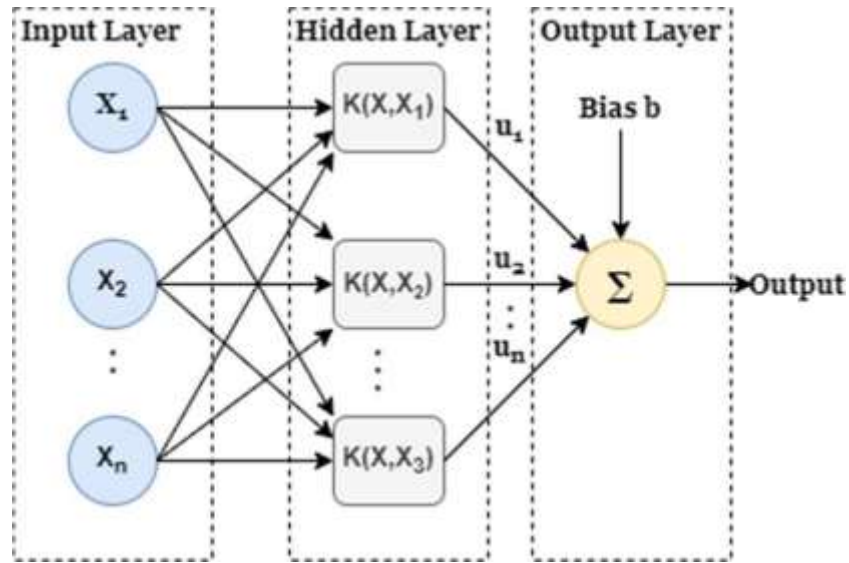
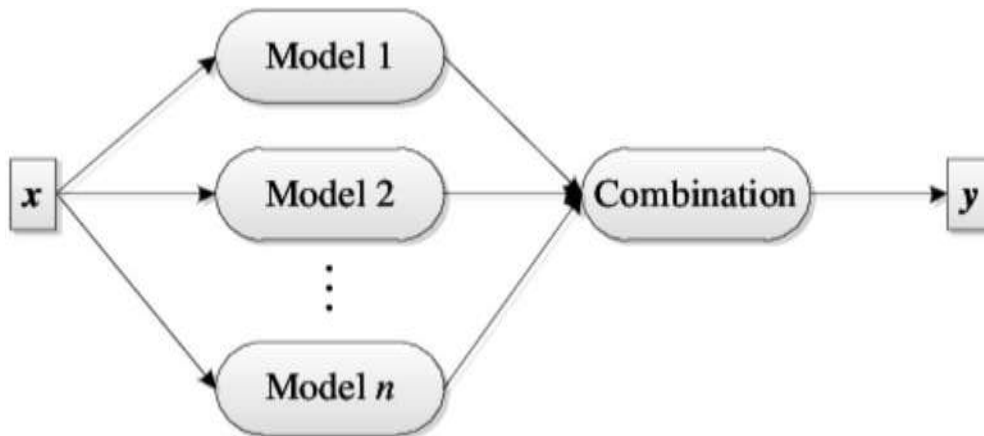


Figure 2: Architecture of SVM Model

4. PROPOSED ENSEMBLE BASED MACHINE LEARNING MODEL

The proposed model achieves higher predictive accuracy than individual models. The core principle is to aggregate various model for achieving higher accuracy by predicting and removing errors arise from noise, variance, and bias. Although noise is unavoidable, ensemble methods effectively reduce variance and bias. A common approach is bootstrapping and aggregation, or "bagging," which resamples the training dataset with replacement to form multiple subsets. These subsets train separate models within the ensemble, and their predictions are aggregated to yield a final, more robust outcome. Bootstrapping enhances the model's variability in the training process, preventing overfitting and allowing the model to capture a broader data distribution. This technique helps reduce bias and improves the model's representation of the dataset's range by incorporating diverse subsets, thus enhancing accuracy.

Figure 3: Ensemble based learning Architecture



10.48047/jocaaa.2024.33.08.190

In bootstrapping, the probability of selecting any instance from the dataset is equal, allowing for a more accurate estimation of the dataset's mean and standard deviation [13]. Boosting, on the other hand, uses a process that applies weighted averages at various levels [14]. Unlike bagging, where models operate separately and predictions are aggregated without preference, boosting sequentially trains weak learners, focusing on correcting errors from previous models. This adaptive approach reduces bias, making it effective for both regression and classification tasks. Stacking [15] combines outputs from two or more models by training a new model that uses these outputs as inputs. It integrates predictions through successive layers to create new predictions, enhancing overall performance compared to individual models [16]. Using diverse base classifiers with weakly correlated predictions can greatly improve results [17].

5. RESEARCH METHODOLOGY

The proposed work develops an Ensemble Machine Learning to identify diagnosis in Liver Disease. The process comprises the collecting of datasets, training of Ensemble Machine Learning using annotated landmarks, and the assessment of performance measures. The research also assesses the effectiveness of Ensemble Machine Learning by comparing it with traditional models like Support Vector Machines, Logistic Regression, and Naïve Bayes. The credibility of results in comparing the usefulness of deep learning to conventional approaches in medical image processing is supported by ethical concerns, strong experimental design, and rigorous statistical analyses.

5.1 Research Design

In order to critically analyze ensemble machine learning models and methodologies in liver disease detection, this work uses a quantitative research methodology. Clinical data and pertinent medical information from individuals with liver disease are collected as part of the study design. In order to access laboratory databases holding patient data and electronic health records (EHRs), cooperation with hospitals and other healthcare facilities is necessary for the data gathering procedure.

5.2 Data Collection

A reputed healthcare facility provided a dataset of clinical data of individuals with liver illness for this investigation. A wide variety of elements are included in the dataset, including imaging results, medical history, demographic data (age, gender), and other pertinent clinical factors. The dataset is meticulously selected to guarantee its quality and correctness. Outliers are handled to preserve data integrity, and missing values are imputed using the proper methods. To uphold ethical standards and safeguard patient privacy, further measures are used to anonymize sensitive patient data. It includes various characteristics related to the diseases. It also contains a binary 'Dataset' column that contains a complete collection of characteristics for studying aspects related with liver health and constructing prediction models for identifying liver illness.

5.3 Preprocessing

Preprocessing the dataset for ensemble machine learning models in liver disease detection entailed resolving missing values using imputation methods, normalizing continuous features, encoding categorical variables, and correcting class imbalance using sampling approaches. These processes assured data integrity, feature scale homogeneity, machine learning algorithm compatibility, and class representation balance. By methodically organizing the information, the following ensemble models could successfully learn from it and

provide accurate diagnostic predictions for liver disease.

5.4 Training Process

By utilizing the training set, the models are optimized and adjusted to reduce errors, enhance performance, and improve their ability to generalize when applied to new, unseen data from the testing set. This approach ensures that the models are well-equipped to provide accurate and reliable predictions when evaluated on the testing set.

6. RESULTS AND ANALYSIS

An F1 score, precision, precision, recall, and recall metrics utilized to examine the digitally controlled system's performance. When juxtaposed with the outcomes derived from the ground truth, these parameters furnish a comprehensive assessment of the model's capacity to consistently detect lower limb misalignments.

- *Accuracy*: Accuracy is a simple metric that quantifies the frequency with which a classifier generates predictions that are accurate [24].
- *Precision*: Precision may be defined as the percentage of instances that have been correctly classified in comparison to the overall number of instances that have been put into classification [25].
- *Recall*: This refers to the percentage of correct positive numbers in comparison to the overall number of true negatives and false negatives.
- *F1-Score*: When calculating the F1 score, the harmonic means of the recall along with accuracy scores is used as the standard deviation [26].

A complete assessment of the proposed Ensemble model in differentiating photographs of Diagnosis in Liver Disease as either misaligned or properly aligned is carried out in the section titled "Findings and Discussion." This evaluation includes a comparison study with existing machine learning models. The confusion matrix is presented at the beginning of the section. This matrix sheds information on the accuracy of the Ensemble Machine Learning model classification as well as places that might potentially be misclassified, so laying the groundwork for model comparisons thereafter. Following that, the ROC curve study investigates the Ensemble Machine Learning model capacity to differentiate between positive and negative instances at a variety of thresholds, with a particular emphasis on a particularly high Area Under the Curve (AUC) value (Figure 4-6).

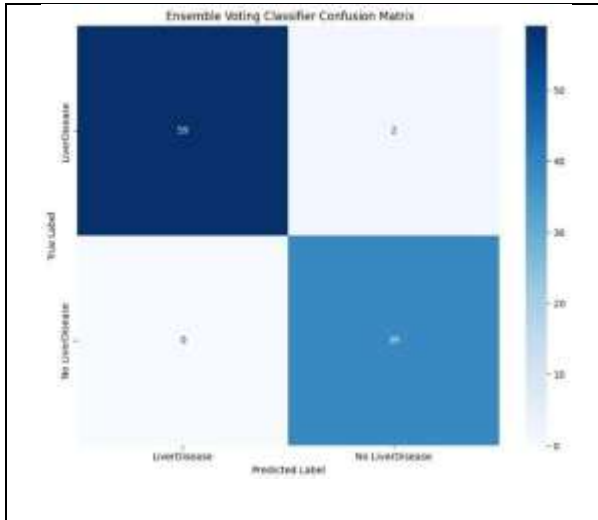


Figure 4: Confusion Matrix of Proposed Ensemble Classifier

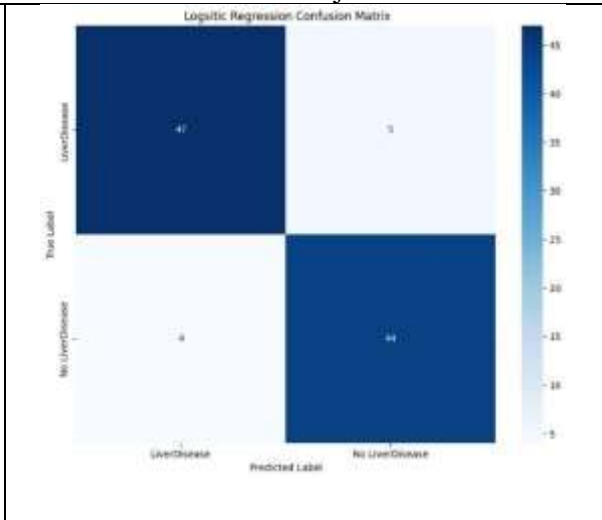


Figure 5: Confusion matrix of Logistic Regression

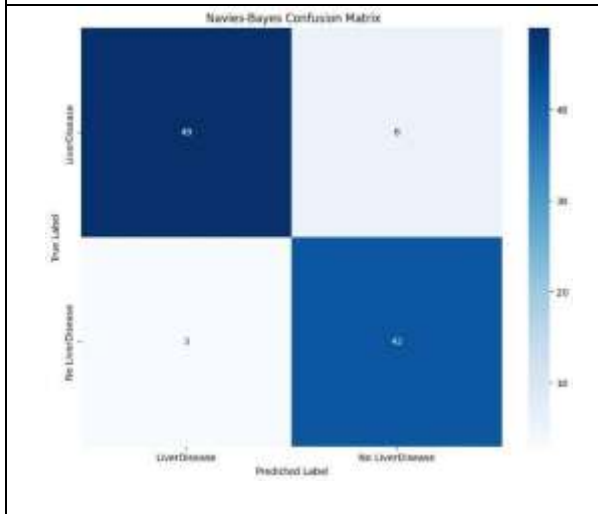


Figure 6: Confusion matrix of Navies-Bayes

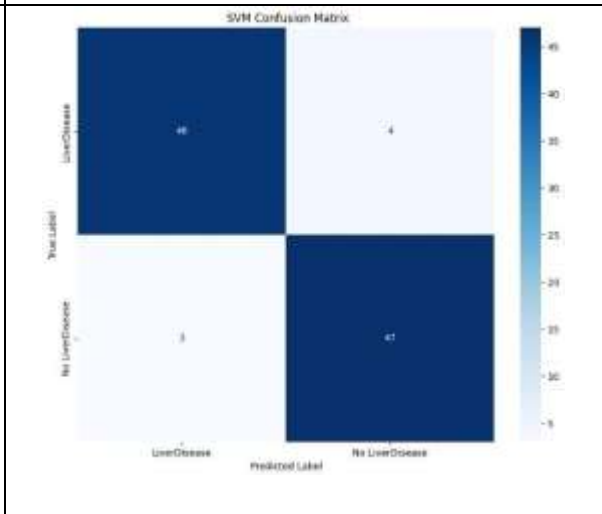


Figure 7: Confusion matrix of SVM

Comparing the confusion matrices of multiple machine learning models shows subtle variations in classification performance. In logistic regression, the true positives (TP) are 46, with 6 false positives (FP), 3 false negatives (FN), and 45 true negatives. Naive Bayes has a comparable TP count of 46, but slightly greater FP at 7 and lower FN at 2, resulting in 45 TN. SVM shows an increased TP count of 50, with 2 FP, 5 FN, and 43 TN. However, the ensemble voting classifier outperforms with 52 TP, just 2 FP, 0 FN, and 46 TN, suggesting a stronger capacity to properly identify positive situations while reducing incorrect classifications. This comparison demonstrates the ensemble voting classifier's ability to achieve high true positive rates and low false positive rates, making it an appealing option for precision-focused classification tasks such as separating liver disease cases from non-liver illness instances.

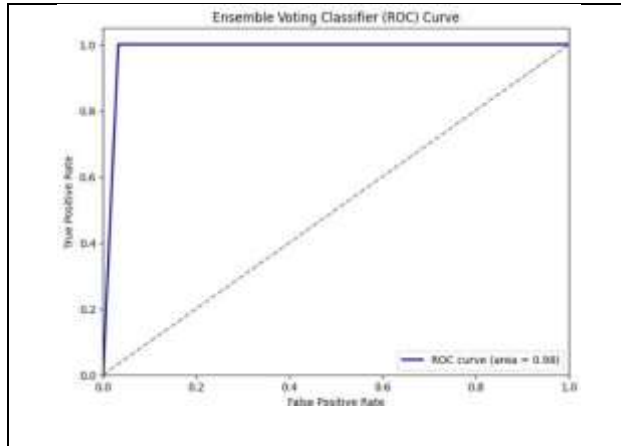


Figure 8: Roc_Curve of Proposed Ensemble classifier

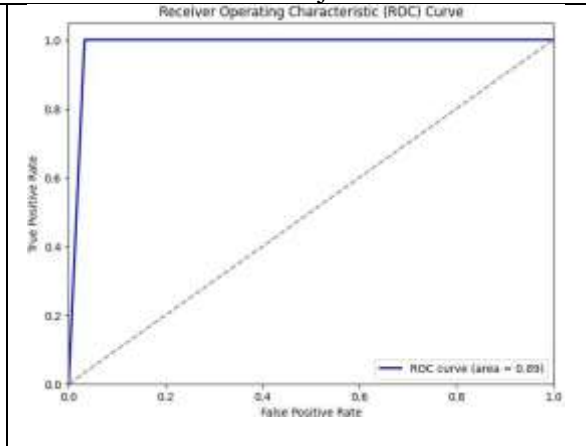


Figure 9: Roc_Curve of Receiver operating

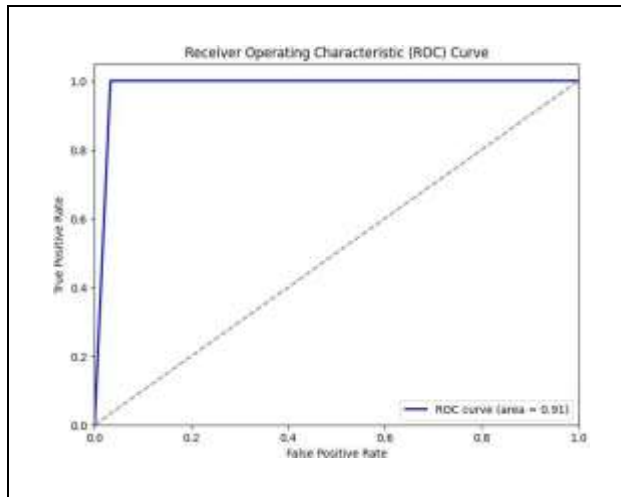


Figure 10: Roc_Curve of Receiver operating characteristic

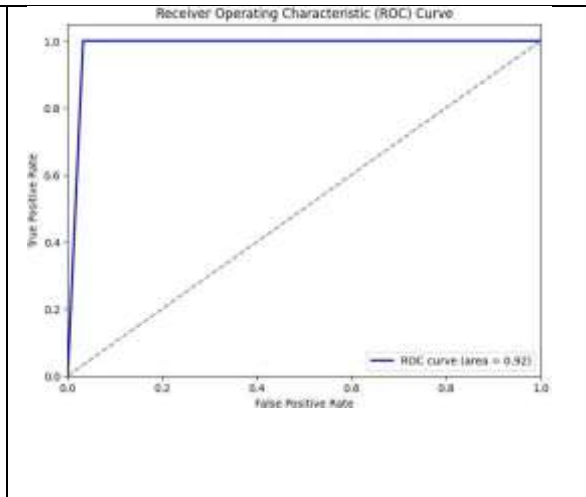


Figure 11: Roc_Curve of Receiver operating characteristic

The comparison of several machine learning models based on ROC curve analysis provides interesting insights into their performance, especially in the context of liver disease categorization. Among the models tested (logistic regression, Naive Bayes, SVM, and ensemble voting classifier), the ensemble voting classifier (Figure 8-11) outperforms the others with an AUC of 0.98, demonstrating exceptional discriminatory power and the ability to accurately classify liver disease and non-liver disease cases. Logistic regression performs well with an AUC of 0.89, followed by Naive Bayes at 0.91 and SVM at 0.92. However, the ensemble voting classifier really shines, harnessing the collective intelligence of its component models to achieve greater predicted accuracy while accurately differentiating between instances of liver illness and non-liver disease. This demonstrates the effectiveness of ensemble approaches in enhancing classification performance, and the ensemble voting classifier emerges as a feasible option for real-world applications needing high accuracy in liver disease diagnostic tasks.

7. PERFORMANCE EVALUATION

10.48047/jocaaa.2024.33.08.190

Performance evaluation in machine learning involves assessing how well a model performs on a given task, typically through various quantitative measures of the system effectiveness in making predictions or classifications. A crucial aspect of performance evaluation is the interpretation of these metrics in context, considering factors like class imbalance, dataset characteristics, and the specific objectives of the task. Evaluating performance often involves techniques such as cross-validation to ensure robustness and generalizability of results. Ultimately, thorough performance evaluation not only validates the model's capabilities but also guides improvements and comparisons with alternative approaches, fostering confidence in its application.

Table 1: Performance evaluation of proposed models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic- Regression	91	91.2	91	91
Naves-Bayes	91.5	91.2	91.2	91.2
SVM	93	93	93	92
Proposed Ensemble Classifier	97	98	98	98

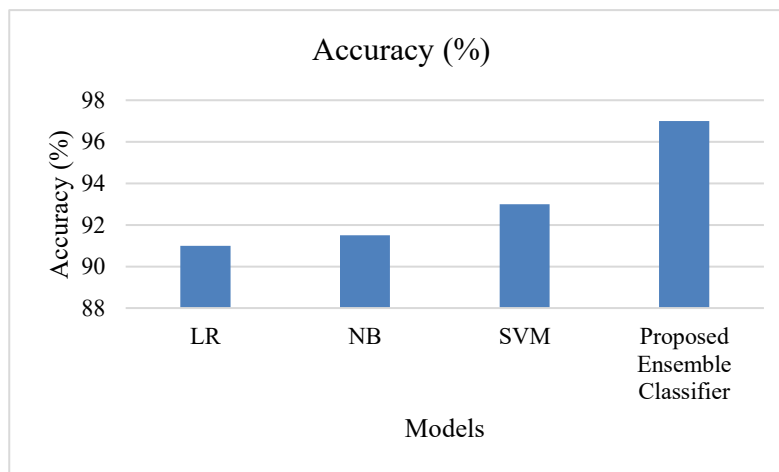


Figure 12: Comparison of Classification Accuracy of various Machine Learning Models

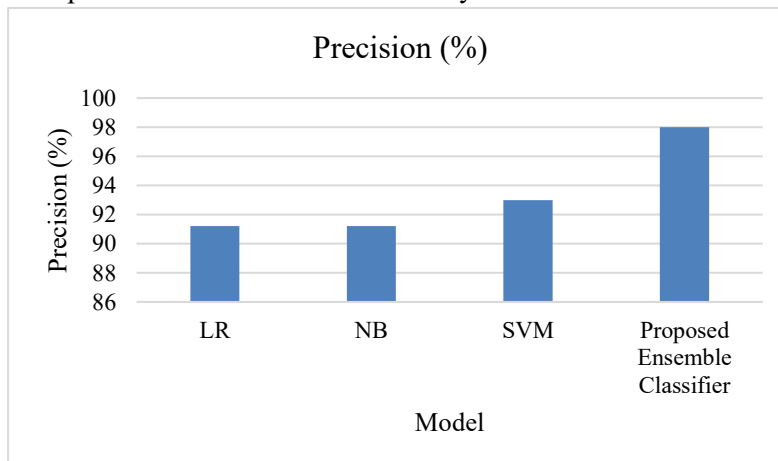


Figure 13: Evaluation of precision of various proposed model

In the evaluation of ensemble models for liver disease diagnosis, several performance metrics were assessed to gauge their effectiveness (Figure 12-15). The logistic regression model exhibited strong performance across all metrics, achieving an accuracy of 91%, with precision, recall, and F1-score also at 91%. Similarly, the Naive Bayes classifier demonstrated competitive results, with an accuracy of 91.5% and consistent precision, recall, and F1-score of 91%. The support vector machine (SVM) model exhibited even higher accuracy at 93%, with matching precision and recall, although a slightly lower F1-score of 92%. Notably, the ensemble voting classifier surpassed all individual models, attaining an impressive accuracy of 97% along with superior precision, recall, and F1-score, each at approximately 98%. These results indicate the effectiveness of ensemble techniques in improving diagnostic accuracy, with the ensemble voting classifier emerging as the top-performing model for liver disease diagnosis.

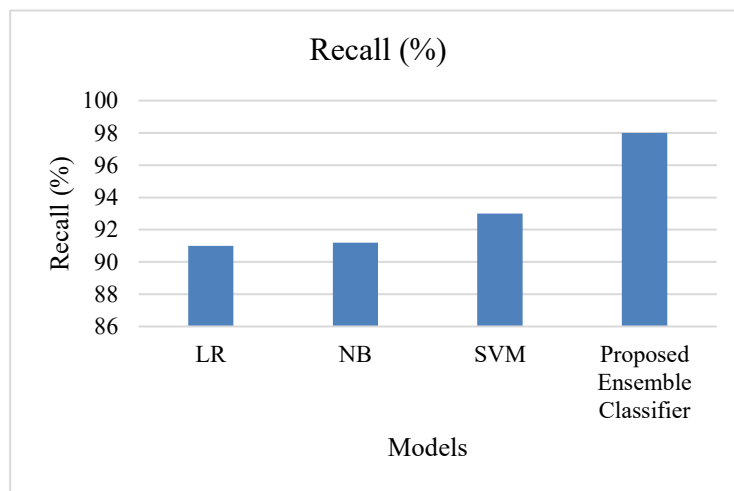


Figure 14: : Evaluation of recall of various proposed model

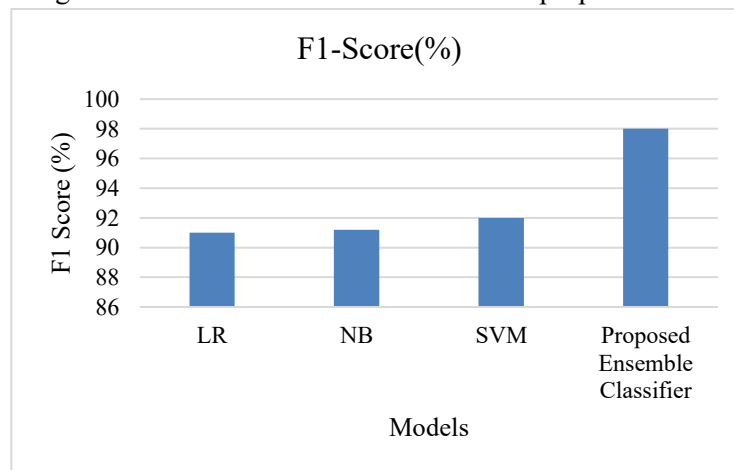


Figure 15: : Evaluation of F1-Score of various proposed model

The assessment of the suggested Ensemble model for the diagnosis of liver illness shows remarkable outcomes that point to its better performance in comparison to other models. After conducting a thorough research that includes ROC curve investigations, confusion matrix inspection, and performance metrics

10.48047/jocaaa.2024.33.08.190

evaluation, the Ensemble Voting Classifier continuously shows strong classification skills with low false positive rates and high true positive rates. The model's remarkable ability to discriminate between cases of liver illness and non-liver disease is shown by its Area Under the Curve (AUC) of 0.98, which highlights its effectiveness. These results highlight the value of ensemble methods in raising diagnostic accuracy and show that the Ensemble Voting Classifier is a strong option for classification tasks requiring precision, with great potential to improve liver disease diagnosis in practical settings.

8. CONCLUSION

In summary, the evaluation of ensemble models in the context of diagnosing liver disease provides substantial knowledge regarding their effectiveness and prospects for clinical implementation. After conducting an extensive examination that includes confusion matrices, ROC curve studies, and a comparison of performance metrics, it is clear that the ensemble voting classifier emerges as the most auspicious model. The ensemble voting classifier outperforms individual models including logistic regression, Naive Bayes, and SVM, as evidenced by its extraordinary accuracy of 97% and astounding precision, recall, and F1-score values of approximately 98% each. The results of this study emphasize the significance of ensemble methods in improving the precision of diagnoses, specifically when it comes to differentiating cases of liver disease from those of non-liver diseases. Therefore, the ensemble voting classifier is demonstrated to be a feasible resolution for practical diagnostic applications involving liver disease.

REFERENCES

- [1] M. Ghosh et al., "A comparative analysis of machine learning algorithms to predict liver disease," *Intell. Autom. Soft Comput.*, Vol. 30, No. 3, pp. 917–928, 2021.
- [2] Jagdeep Singha et al. "Software Based Prediction of Liver Disease with Feature Selection and Classification Techniques" ELSEVIER 2020.
- [3] R. H. Lin, "An intelligent model for liver disease diagnosis," *Artif. Intell. Med.*, Vol. 47, No. 1, pp. 53–62, 2009.
- [4] Greese Gupta et al. "A Web Based Framework for Liver Disease Diagnosis using Combined Machine Learning Models" © IEEE 2020.
- [5] A. K. M. S. Rahman et al. "A comparative study on liver disease prediction using supervised machine learning algorithms," *Int. J. Sci. Technol. Res.*, vol. 8, no. 11, pp. 419–422, 2019.
- [6] P. Sug, "On the optimality of the simple Bayesian classifier under zero-one loss", *Machine Learning* 29 (2–3) (1997)103–130.
- [7] Michael J Sorich, "An intelligent model for liver disease diagnosis", *Artificial Intelligence in Medicine* 2009;47:53-62.
- [8] C. J. A. Kannan, "Comparative Analysis of Machine Learning Techniques for Indian Liver Disease Patients," 6th Int. Conf. Adv. Comput. Commun. Syst., 2020.
- [9] Maria Alex Kuzhippallil, Carolyn Joseph, Kannan A, "Comparative Analysis of Machine Learning Techniques for Indian Liver Disease Patients" © IEEE 2020.
- [10] A. Singh, "A Comprehensive Guide to Ensemble Learning (with Python codes)," 2023.
- [11] A. Quadir et al. "Enhanced Preprocessing Approach Using Ensemble Machine Learning Algorithms for Detecting Liver Disease," 2023.
- [12] R. Ashtagi et al. "International Journal of Intelligent Systems and Applications in Engineering Revolutionizing Early Liver Disease Detection: Exploring Machine Learning and Ensemble

10.48047/jocaaa.2024.33.08.190

- Models,” Orig. Res. Pap. Int. J. Intell. Syst. Appl. Eng. IJISAE, vol. 2024, no. 13s, pp. 528–541, 2024.
- [13] M. O. Edeh et al., “Artificial Intelligence-Based Ensemble Learning Model for Prediction of Hepatitis C Disease,” *Front. Public Heal.*, vol. 10, no. April, 2022.
- [14] E. Dritsas and M. Trigka, “Supervised Machine Learning Models for Liver Disease Risk Prediction,” *Computers*, vol. 12, no. 1, p. 19, 2023.
- [15] L. Meng, W. Treem, G. A. Heap, and J. Chen, “A stacking ensemble machine learning model to predict alpha-1 antitrypsin deficiency-associated liver disease clinical outcomes based on UK Biobank data,” *Sci. Rep.*, vol. 12, no. 1, pp. 1–18, 2022.
- [16] Z. Jin et al. “RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis,” in *Lecture Notes in Computer Science*, vol. 12343 LNCS, 2020, pp. 503–515.
- [17] B. V. b Ashfaqur and Rahman a, “Ensemble classifier generation using non-uniform layered clustering and Genetic Algorithm,” *Knowledge-Based Syst.*, 2013.
- [18] Y. Freund and R. E. Schapire, “Experiments with a New Boosting Algorithm,” *Proc. 13th Int. Conf. Mach. Learn.*, pp. 148–156, 1996.
- [19] S. Džeroski and B. Ženko, “Is combining classifiers with stacking better than selecting the best one,” *Mach. Learn.*, vol. 54, no. 3, pp. 255–273, 2004
- [20] Sethi, K., Gupta, A., Gupta, G., Jaiswal, V., “Comparative Analysis of Machine Learning Algorithms On Different Datasets” *International Conference on Innovations in Computing (ICIC 2017)*.
- [21] Sethi, K., Gupta, A., Gupta, Jaiswal, V., “Machine Learning Based Performance Evaluation System Based on Multi-Categorical Factors” *IEEE, Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC) 2018*.
- [22] Sethi, K., Jaiswal, V., & Ansari, M. D. (2020). Machine learning based support system for students to select stream (subject). *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science)*, 13(3), 336-344(Scopus Indexed).
- [23] Sethi, K., & Jaiswal, V. (2022). PSU-CNN: prediction of student understanding in the classroom through student facial images using convolutional neural network. *Materials Today: Proceedings*, 62, 4957-4964.
- [24] Panwar, N., & Sethi, K. (2022, December). Breast cancer: Early diagnosis by using artificial intelligence. In *2022 2nd International Conference on Innovative Sustainable Computational Technologies (CISCT)* (pp. 1-4). IEEE.
- [25] Rohit, Sethi, K., & Milan, S. (2023). Epidemiology of occupational injuries and hazards in, Indian hospitality: a study report. *Eur. Chem. Bull*, 12, 2586-259.
- [26] Rohit, M., Sethi, K., Khan, M., & Raina, A. (2023). Machine Learning Model for Prediction of the Chemicals Harmfulness on Staff and Guests in the Hospitality Industry: A Pilot Study. *Data and Metadata*, 2, 161-161.
- [27] Sharma, M., Sethi, K., Dhiman, K., & Gusain, A. (2023). Impact of COVID-19 on hospitality workers in India due to perceived employment uncertainty and mental health issues. *International Journal for Multidimensional Research Perspectives*, 1(3), 151-161.
- [28] Rawat, N., & Sethi, K. (2023, September). Early Detection of Breast Cancer By Using Machine Learning Algorithms. In *2023 3rd International Conference on Innovative Sustainable Computational Technologies (CISCT)* (pp. 1-3). IEEE.
- [29] Sethi, K., Sharma, M., & Gusain, A. (2023, June). The Acceptance of Machine module in the

10.48047/jocaaa.2024.33.08.190

- Hospitality Industry with Prospects and Challenges: A Review. In 2023 2nd International Conference on Computational Modelling, Simulation and Optimization (ICCMO) (pp. 283288). IEEE.
- [30] Anita, M., Sethi, K (2023, Dec). Review of Machine Learning Based Techniques For Liver Disease Diagnosis. In 2023. Stochastic Modelling and Computational Sciences, Vol. 3 No.2,1212 – 1221.
- [31] Rohit, Sethi, K.Sharma, M (2023). Importance of Sustainable Housekeeping And Hygiene Programs In Hospitality: A Review. 2023 2nd International Conference on Computational Modelling, Simulation and Optimization (ICCMO).
- [32] Sethi, K., Sharma, A., Jaiswal, V., (2019) “Impact of social and cultural challenges in the education world using AI” Book chapter published in IGI global, Ch-7, 130-151.
- [33] Sethi, K., Chauhan, S., & Jaiswal, V. (2020). Artificial Intelligence in Higher Education: First Attempt. In Impact of AI Technologies on Teaching, Learning, and Research in Higher Education (pp. 1-29). IGI Global.
- [34] Sethi, K., Vandana., Gupta, A., (2020) “Computational Approaches for Malaria Diagnosis Using Machine Learning: A review”, Accepted book chapter for MHARA.
- [35] Meghta, P., and Kumar, A. Choudhary (2021), “International Journal of Mechanical Engineering, Particle Swarm Optimization and Ensemble Learning based Apple Leaf Disease Detection, Pg. 278-291, Vol. 6 No. 2 September 2021, ISSN: 0974-5823.
- [36] Meghta, P., and Kumar A. Choudhary (2022), “An Efficient Deep Learning-based Model for Apple Diseases Prediction”, NEUROQUANTOLOGY, Volume 20, Issue 6, Page101294-101302, June 2022, eISSN:1303-5150.