Jurnal Informatika dan Rekayasa Perangkat Lunak ISSN 2656-2855
Vol. 7, No. 1, Maret 2025, Hal. 31-41 e-ISSN 2685-5518

Leveraging SAMME for Improved Multi-Class Cirrhosis Diagnosis in Clinical Settings

Arum Kurnia Sulistyawati[1], Mohammad Diqi[2*], Dyan Avando Meliala[3], Marselina Endah Hiswati[4], Ajie Wibowo Soejono[5], Dini Sari[6]
[1,3,5,6] Dept. of Information System, Universitas Respati Yogyakarta, Indonesia
[2,4] Dept. of Informatics, Universitas Respati Yogyakarta, Indonesia
[*]Email: diqi@respati.ac.id

Abstract

This study explores the use of the SAMME algorithm to develop a predictive model for identifying various stages of cirrhosis. The dataset includes 418 records with 20 attributes, targeting the classification of cirrhosis stages: C (censored), CL (censored due to liver transplantation), and D (death). The model achieved an overall accuracy of 94%, demonstrating high precision and recall for classes C and D. However, the precision for class CL was lower, indicating a tendency to over-predict this stage. These results validate the SAMME algorithm's potential to enhance diagnostic accuracy while highlighting the need for further refinement to address class imbalance and feature overlap. This research underscores the value of machine learning in early diagnosis and personalized treatment, suggesting future work on larger, balanced datasets and advanced feature engineering to improve model robustness and reliability in clinical applications.

Keywords*: cirrhosis, predictive modeling, SAMME algorithm, machine learning, healthcare*

INTRODUCTION

Cirrhosis is a chronic liver disease characterized by the progressive replacement of liver tissue by fibrosis and scar tissue, disrupting blood flow and liver function (Iyer et al., 2019). Optimizing non-invasive indices like APRI and FIB-4 has proven crucial for managing chronic hepatitis B by accurately ruling out cirrhosis and reducing misclassification rates (Kuo et al., 2019).

Early detection of liver fibrosis is essential for therapeutic decision-making, and serum N-glycan fingerprinting provides a high-throughput method to stage liver fibrosis non-invasively (Huang et al., 2021). Machine learning techniques applied to MRI images can effectively classify liver tissue as fibrotic, aiding in the early identification of fibrosis and potentially preventing its progression to cirrhosis (Schawkat et al., 2020). The Delta-4 fibrosis score is a novel, non-invasive fibrosis score that can accurately detect cirrhosis in patients with chronic Hepatitis D, showing superior performance to established non-invasive tests (Da et al., 2020).

The diagnostic accuracy of human microfibrillar-associated protein 4 (MFAP4) for detecting fibrosis in patients with alcoholic liver disease highlights its potential as a biomarker, emphasizing the need for effective biomarkers in improving outcomes for such conditions (Madsen et al., 2020). Innovative techniques like the Hybrid Coupled Dictionary Pairs on Longitudinal Domain (HCDPLD) approach offer substantial improvements in the resolution and accuracy of ultrasound imaging for diagnosing cirrhosis, potentially reducing the reliance on more invasive methods (Kirubakaran et al., 2019). A hierarchical algorithm using administrative healthcare data has been validated to accurately define cirrhosis etiology, which could significantly assist health services research in understanding and managing this growing patient population (Philip et al., 2020). Population screening for liver fibrosis using serological tests, transient elastography, and radiological methods has shown promising results in identifying significant undetected liver fibrosis or established cirrhosis, potentially altering the approach to diagnosing chronic liver diseases from a later to an earlier and more treatable stage (Rajesh et al., 2020). Deep learning models utilizing ultrasonography images have

demonstrated superior performance in classifying liver fibrosis stages, providing a more accurate assessment than traditional radiological evaluations (Lee et al., 2019).

Early detection and classification of cirrhosis are critical in the medical management of the disease (Lee et al., 2019). Early intervention can significantly reduce the risk of severe complications such as variceal bleeding, hepatic encephalopathy, and liver cancer (Rajesh et al., 2020). Classifying cirrhosis based on the severity of the disease helps in determining the most appropriate treatment strategy for patients (Kirubakaran et al., 2019).

Previous research has shown that using clinical variables and laboratory results in multi-class classification of cirrhosis allows for more accurate detection and more effective management (Philip et al., 2020). However, the main challenges faced in this classification include the vast diversity of symptoms among patients and the complexity in interpreting test results, requiring sophisticated analytical solutions to accurately differentiate between disease stages (Shu et al., 2019).

The SAMME algorithm (Stagewise Additive Modeling using a Multi-class Exponential loss function) is an 'ensemble learning' technique that enhances classification accuracy by combining predictions from several weak models into a more robust model through an iterative process (Lee et al., 2019). This approach gives more weight to incorrectly classified examples, allowing the algorithm to focus more on the more complex cases and, consequently, improve overall prediction accuracy (Q. Wang & Sun, 2021). The application of the SAMME algorithm, especially in a medical context like cirrhosis diagnosis, offers a promising yet relatively unexplored new avenue (Philip et al., 2020).

Furthermore, the use of the SAMME algorithm in clinical studies is still limited, providing a unique opportunity for further exploration in this research (Philip et al., 2020). Effective research using SAMME could potentially improve patient clinical outcomes by providing deeper insights into evidence-based classification and management of cirrhosis (Kirubakaran et al., 2019). Therefore, this study aims not only to develop a robust model but also to demonstrate the applicability of this technique in a clinical setting, providing a reliable diagnostic tool for healthcare practitioners to manage cirrhosis (Y. Wang et al., 2023).

The primary objective of this research is to develop a predictive model using the SAMME algorithm to accurately identify the various stages of cirrhosis, specifically focusing on the classification of stages C (censored), CL (censored due to liver transplantation), and D (death). By leveraging a comprehensive dataset with 20 clinical and biochemical attributes, the study aims to enhance diagnostic precision and recall, thereby improving early detection and personalized treatment plans for cirrhosis patients. The research seeks to address the current challenges in cirrhosis identification, such as class imbalance and overlapping features, and demonstrate the effectiveness of machine learning techniques in medical diagnostics.

This research contributes significantly to the field of medical diagnostics by demonstrating the effectiveness of the SAMME algorithm in improving the accuracy of cirrhosis stage identification. It provides valuable insights into handling multi-class classification challenges in medical datasets, particularly addressing class imbalances and feature overlaps. By integrating machine learning techniques into the diagnostic process, the study enhances early detection and supports personalized treatment strategies for cirrhosis patients. Furthermore, it extends the existing body of literature on predictive modeling in healthcare, offering practical recommendations for refining and implementing these models in clinical settings. This research lays the groundwork for future studies to build upon, ultimately contributing to more effective and reliable diagnostic tools in the medical community.

LITERATURE REVIEW

The identification techniques for cirrhosis used in previous studies, such as conventional statistical analysis, medical imaging, and machine learning models based on regression and classification, have significantly contributed to improving the accuracy of early-stage disease diagnosis (Blanes-Vidal et al., 2022). The use of medical imaging methods like ultrasound and MRI enables more precise detection of structural changes in the liver, while biochemical and biomarker analyses provide essential data on liver function (Singh

et al., 2023). On the other hand, machine learning models have enhanced predictive capabilities by leveraging various clinical and demographic parameters, allowing the identification of more complex patterns and providing more accurate predictions compared to conventional methods (Y. Chen et al., 2022). The combination of these various techniques has aided in the earlier recognition of cirrhosis, enabling quicker and more effective medical intervention (Singh et al., 2023).

Previous research has compared the effectiveness of conventional techniques and predictive model-based approaches in the context of cirrhosis identification (Blanes-Vidal et al., 2022). Studies have shown that while conventional methods, such as clinical examinations and standard diagnostic tests, provide valuable insights, predictive models often outperform them in terms of accuracy and early detection (Kim et al., 2023). Predictive models, leveraging machine learning algorithms, can analyze vast amounts of data and identify complex patterns that may not be apparent to human experts (Nia et al., 2023). These models can incorporate multiple variables simultaneously, offering a more comprehensive and precise diagnosis (Alowais et al., 2023). Consequently, predictive models have demonstrated higher sensitivity and specificity in identifying cirrhosis at its early stages compared to traditional diagnostic techniques (Y. Chen et al., 2022).

The quality of the SAMME algorithm can be optimized to enhance its accuracy in identifying various stages of cirrhosis, especially when dealing with complex datasets, by incorporating several strategies (Naseem et al., 2020). Firstly, advanced preprocessing techniques such as normalization, handling missing values, and feature engineering can be employed to improve data quality (H. Ding et al., 2022). Secondly, hyperparameter tuning through methods like grid search or random search can help in finding the optimal parameters for the SAMME algorithm (Sarker, 2021). Additionally, integrating ensemble learning techniques, such as boosting or bagging, can further enhance model performance by combining the strengths of multiple models (N. Ding et al., 2022). Moreover, leveraging cross-validation techniques ensures that the model generalizes well to unseen data. Lastly, utilizing domain-specific knowledge to select relevant features and applying dimensionality reduction techniques can help in managing the complexity of the dataset, ultimately leading to improved accuracy in cirrhosis identification (Mostafa et al., 2021).

To minimize errors and enhance the ability of a predictive model using SAMME to recognize multi-class cirrhosis, several specialized strategies and techniques can be applied in line with the latest advancements in machine learning (Tanwar & Rahman, 2021). One approach is to implement feature selection and extraction methods to identify the most relevant attributes, reducing noise and improving model accuracy (Zhang et al., 2020). Additionally, employing techniques such as SMOTE (Synthetic Minority Over-sampling Technique) can address class imbalance by generating synthetic samples for underrepresented classes, ensuring better model performance across all classes (R. Haluška et al., 2022). Hyperparameter tuning through methods like Bayesian optimization can help in finding the optimal settings for the algorithm. Incorporating cross-validation ensures the robustness and generalizability of the model (Shekhar et al., 2021). Moreover, leveraging ensemble methods, such as stacking or blending, can combine multiple models' predictions to enhance overall performance (Shahhosseini & Pham, 2022). Utilizing advanced regularization techniques like L1 and L2 regularization can prevent overfitting, thus improving the model's ability to generalize from the training data to unseen data effectively (Nguyen et al., 2021).

The fundamental concept behind developing a predictive model for multi-class cirrhosis identification using the SAMME algorithm involves leveraging the unique characteristics of the dataset to improve accuracy and performance (Menegotto et al., 2021). This includes identifying and utilizing critical parameters such as age, bilirubin levels, albumin levels, platelet count, and prothrombin time, which are key indicators of liver function and disease progression (H. Ding et al., 2022). The SAMME algorithm, being an ensemble method, benefits from these parameters by iteratively adjusting weights to focus on misclassified instances, thereby improving

model accuracy (Akter et al., 2021). Additionally, preprocessing steps such as normalization, handling missing values, and feature engineering are essential to enhance data quality and ensure that the model captures the relevant patterns (Gupta et al., 2021). By systematically incorporating these critical parameters and preprocessing techniques, the SAMME algorithm can effectively differentiate between the various histologic stages of cirrhosis, providing a robust predictive model for clinical application (K. Chen et al., 2022).

The SAMME algorithm, a boosting technique for multi-class classification, is applied in the development of a predictive model for identifying cirrhosis by sequentially training weak classifiers on weighted versions of the dataset (Nam et al., 2020). Initially, each instance in the dataset is assigned equal weight, but in subsequent iterations, the weights of misclassified instances are increased, directing the model's focus towards more challenging cases (Duhayyim et al., 2022). This process continues until the model achieves optimal performance. In the context of cirrhosis identification, SAMME leverages critical features such as age, bilirubin levels, and platelet count to discern between different histologic stages (A. Singh et al., 2020). By emphasizing misclassified instances and combining the outputs of multiple classifiers, SAMME effectively handles the complexity of varying cirrhosis stages, ensuring accurate and robust predictions across all classes (Taylor-Weiner et al., 2021). This iterative focusing mechanism allows the model to adapt to the nuances of the disease's progression, thereby enhancing its diagnostic capability (Shi et al., 2023).

RESEARCH METHOD

3.1 Dataset

The dataset used from Kaggle consists of 418 records and includes 20 attributes, with the target variable being the status of the patient: C, CL, or D. The attributes provide a comprehensive overview of each patient's condition and treatment, covering aspects such as a unique identifier (ID), the number of days between registration and the earlier of death, transplantation, or study analysis time (N_Days), the type of drug administered (Drug: D-penicillamine or placebo), and various clinical features including age (in days), sex, presence of ascites, hepatomegaly, and spiders. Additionally, it records the presence of edema, serum levels of bilirubin and cholesterol, albumin levels, urine copper, alkaline phosphatase (Alk_Phos), SGOT, triglycerides, platelet count, and prothrombin time. This rich dataset is essential for developing a predictive model to identify the status of cirrhosis patients, leveraging the diverse clinical and biochemical features it encompasses.

3.2 Data Preprocessing

Data processing involved several crucial steps to prepare the dataset for modeling. Initially, missing values were handled by filling them with appropriate replacement values: mode values were used for categorical variables such as 'Drug', 'Ascites', 'Hepatomegaly', 'Spiders', and 'Stage', while median values were applied to numerical variables like 'Cholesterol', 'Copper', 'Alk_Phos', 'SGOT', 'Triglycerides', 'Platelets', and 'Prothrombin'. This approach ensures that the dataset remains complete without introducing significant bias. Next, normalization was applied to scale numerical attributes, ensuring that they are on a comparable scale and improving the performance of machine learning algorithms. Additionally, categorical data were converted into numerical format using techniques such as one-hot encoding, enabling the algorithms to process them effectively. Finally, the dataset was split into training and testing subsets, with 80% allocated for training and 20% for testing, ensuring that the model can be evaluated on unseen data to assess its generalization capability.

3.3 Concept of the SAMME Algorithm

The SAMME algorithm (Stagewise Additive Modeling using a Multi-class Exponential loss function) is a boosting method designed to handle multi-class classification problems. Boosting is an ensemble technique that combines multiple weak learners to form a more robust model by iteratively improving prediction accuracy. SAMME extends the concept of Adaboost (Adaptive Boosting) from binary to multi-class classification using an exponential loss function.

SAMME uses an exponential loss function to measure prediction errors. This function is expressed in Equation 1.

$$L\big(y, F(x)\big) = \exp\big(-yF(x)\big) \quad (1)$$

where y is the true class label and $F(x)$ is the model's prediction.

Each data instance is initially assigned equal weight. After each iteration, the weights of misclassified instances are increased so that the next weak learner focuses more on these problematic instances. The new weight w_i for each instance i is calculated in Equation 2.

$$w_i^{(t+1)} = w_i^{(t)} \exp\Big(\alpha_t I\big(y_i \neq h_t(x_i)\big)\Big) \quad (2)$$

where α_t is the strength of the weak learner at iteration t, $I(\cdot)$ is an indicator function that is 1 if the condition inside is true, and $h_t(x_i)$ is the prediction of the weak learner at iteration t.

The strength of each weak learner α_t is calculated based on the accuracy of that learner on the weighted data. Mathematically, this is expressed in Equation 3.

$$\alpha_t = \frac{1}{2}\ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right) + \ln(K-1) \quad (3)$$

where ε_t is the error rate of the weak learner at iteration t, and K is the number of classes.

The final prediction of the SAMME model is a combination of all the weak learners, taking into account their respective strengths. The final prediction $F(x)$ is computed in Equation 4.

$$F(x) = \sum_{t=1}^{T} \alpha_t h_t(x) \quad (4)$$

where T is the total number of iterations (or weak learners).

In the context of cirrhosis identification, the SAMME algorithm is applied to classify patients into various histologic status of the disease. By using clinical and biochemical attributes from the dataset, SAMME aims to build a robust predictive model through iterations and the strengthening of weak learners. Consequently, SAMME can handle the complexity and variability in medical data to provide more accurate and reliable predictions.

3.4 Evaluation Metrics

To evaluate the predictive model, we will use two key metrics: the confusion matrix and the classification report. The confusion matrix is a summary table that helps visualize the performance of the classification model by comparing the actual and predicted labels. It consists of four main components for each class: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). For a multi-class classification problem like cirrhosis stage identification, the confusion matrix will be extended to include each class. This matrix will provide a detailed insight into the types of errors the model is making, showing how often the model correctly predicts each class and where it tends to misclassify.

The classification report provides additional metrics that summarize the performance of the model for each class, including precision, recall (sensitivity), and F1 score, as calculated in Equations 5-7.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (7)$$

Precision measures how many of the predicted positive instances are actually positive, while recall measures the model's ability to identify all positive instances. The F1 score, being the harmonic mean of precision and recall, offers a balanced measure of the model's accuracy, especially useful for imbalanced datasets. By examining these metrics, we can assess the model's effectiveness and robustness, identify specific patterns of misclassification, and determine areas for improvement. This comprehensive evaluation helps us understand the model's strengths and weaknesses, guiding further enhancements to achieve better performance.

RESULT AND DISCUSSION

4.1 Model Performance

The confusion matrix, presented as Figure 1, provides a detailed overview of the model's performance in predicting the status of cirrhosis patients across three classes: C, CL, and D. The matrix highlights the number of correct and incorrect predictions made by the model for each class. For instance, the model correctly predicted 41 instances of class C, 4 instances of class CL, and 34 instances of class D. However, it misclassified 3 instances: predicting 2 instances of class CL and 1 instance of class D as class C, and predicting 1 instance of class C and 1 instance of class D as class CL. This confusion matrix serves as a crucial tool for identifying specific patterns of misclassification and areas where the model's predictive accuracy can be improved.

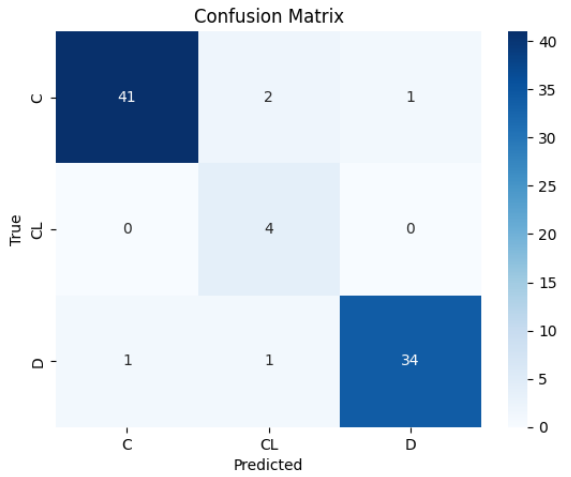

Figure 1. Confusion matrix

The analysis of the confusion matrix reveals several key insights into the model's performance. For the class C, the model correctly identified 41 instances but misclassified 2 instances as CL and 1 instance as D. For class CL, the model demonstrated perfect precision with 4 correct predictions and no misclassifications. However, for class D, while the model correctly identified 34 instances, it misclassified 1 instance as C and 1 instance as CL. The high number of true positives for each class indicates a generally strong performance, but the presence of misclassifications suggests areas for improvement, particularly in distinguishing between classes C and CL. These misclassifications could be due to overlapping features or insufficient differentiation in the dataset, necessitating further refinement of the model or additional feature engineering to enhance its predictive accuracy.

The classification report, presented as Table 1, provides a comprehensive evaluation of the model's performance across three classes: C, CL, and D. The report includes precision, recall, and F1-score metrics for each class, along with the overall accuracy, macro average, and weighted average scores. Class C shows high precision (0.98), recall (0.93), and F1-score (0.95) with a support of 44 instances, indicating strong performance in identifying this class. Class CL, despite a lower precision of 0.57, achieves a perfect recall of 1.00 and an F1-score of 0.73, albeit with a smaller support of 4 instances. Class D also demonstrates high performance with a precision of 0.97, recall of 0.94, and an F1-score of 0.96 with 36 instances. The overall accuracy of the model stands at 0.94, reflecting its robustness in correctly classifying the majority of instances. The macro average and weighted average metrics further underscore the model's balanced performance across all classes.

Table 1. Classification report

	precision	recall	f1-score	support
C	**0.98**	**0.93**	**0.95**	**44**
CL	**0.57**	**1.00**	**0.73**	**4**
D	**0.97**	**0 .94**	**0.96**	**36**
accuracy			**0.94**	**84**
macro avg	**0.84**	**0.96**	**0.88**	**84**
weighted avg	**0.95**	**0.94**	**0.94**	**84**

The classification report indicates that the model performs exceptionally well overall, with an accuracy of 0.94. Class C has the highest precision at 0.98 and a recall of 0.93, resulting in a strong F1-score of 0.95, reflecting the model's reliability in identifying this class correctly. Class D also shows high performance with a precision of 0.97, recall of 0.94, and F1-score of 0.96, indicating that the model is similarly effective in predicting this class. However, Class CL has a lower precision of 0.57 but achieves a perfect recall of 1.00, suggesting that while the model identifies all

true instances of this class, it also misclassifies other instances as CL. The macro average scores of 0.84 for precision, 0.96 for recall, and 0.88 for F1-score reveal a disparity in performance across classes, particularly affecting precision. The weighted averages, closely aligned with the overall accuracy, demonstrate that despite some variability, the model maintains robust performance across the dataset, although further refinement is needed to improve precision, especially for the CL class.

4.2 Summarization of Key Findings

The research addressed the critical problem of accurately identifying cirrhosis stages in patients, a process that is currently challenged by the complexity of the disease and limitations in existing diagnostic methods. Using the SAMME algorithm to develop a predictive model, the study aimed to enhance identification accuracy for multi-class cirrhosis classification. The key findings revealed that the model achieved an overall accuracy of 94%, with particularly high precision and recall for the C and D classes. However, while the model showed a perfect recall for the CL class, its precision for this class was notably lower, indicating a tendency to over-predict CL instances. These results demonstrate the model's potential to significantly improve cirrhosis identification, though they also highlight areas for further refinement, particularly in balancing precision and recall across all classes.

4.3 Result Interpretations

The results of the study indicate several noteworthy patterns and relationships among the data. The high precision and recall for the C and D classes suggest that the model is highly effective at distinguishing these stages of cirrhosis. However, the perfect recall but low precision for the CL class indicates a pattern of over-prediction, where the model frequently misclassifies other classes as CL. This result partially met the expectations, as the model was anticipated to perform well overall but showed an unexpected imbalance in classifying CL instances. Possible alternative explanations for this could include an inherent class imbalance in the dataset or overlapping features between the classes that confuse the model. Further investigation and refinement, such as enhancing feature selection or applying advanced balancing techniques, might be required to address these issues and improve the model's precision for the CL class.

4.4 Research Implications

The relevance and implications of this research are significant for both clinical practice and the broader field of medical data science. The study's findings, which demonstrate the efficacy of the SAMME algorithm in accurately identifying multiple stages of cirrhosis, align with and extend previous literature that highlights the potential of machine learning models to improve diagnostic accuracy. By addressing the specific challenge of multi-class classification in cirrhosis, this research provides new insights into handling complex medical datasets and improving precision in underrepresented classes, such as CL. These contributions underscore the potential for predictive models to enhance early detection and personalized treatment plans in clinical settings, paving the way for further advancements in medical diagnostics and patient care.

4.5 Research Limitations

While this study successfully demonstrates the potential of the SAMME algorithm in enhancing the accuracy of cirrhosis stage identification, several limitations must be acknowledged. The relatively small dataset size and potential class imbalance, particularly in the CL category, may have influenced the model's precision and overall performance. These factors could lead to overfitting or underrepresentation of certain stages, affecting the generalizability of the results. Despite these limitations, the results remain valid for answering the research question as the model still achieved high overall accuracy and provided meaningful insights into improving diagnostic processes. The findings underscore the importance of further research with larger, more balanced datasets to validate and refine the model, ensuring its robustness and applicability in real-world clinical settings.

4.6 Recommendations for Future Research

For practical implementation, it is recommended that the predictive model be

integrated into clinical decision support systems to aid healthcare professionals in early and accurate diagnosis of cirrhosis. To enhance the model's robustness, future research should focus on acquiring larger and more balanced datasets, ensuring that all classes, especially underrepresented ones like CL, are adequately represented. Additionally, exploring advanced feature engineering techniques and incorporating domain-specific knowledge could further improve model performance. Future studies might also investigate the application of other boosting algorithms or hybrid models to compare effectiveness and identify the most suitable approach for this medical context. Continuous collaboration with medical experts will be crucial to validate and refine the model, ensuring its practical relevance and reliability in diverse clinical environments.

CONCLUSION

In conclusion, this study highlights the potential of using the SAMME algorithm to develop a predictive model for the multi-class identification of cirrhosis stages. The model demonstrated high overall accuracy, particularly excelling in predicting the C and D classes with impressive precision and recall. However, the results also indicated an area for improvement in the precision of the CL class, suggesting the need for further refinement in handling class imbalances and overlapping features. These findings underscore the feasibility of leveraging machine learning to enhance diagnostic accuracy and provide a more nuanced understanding of cirrhosis progression.

The implications of this research are significant, offering valuable insights into improving clinical decision-making and patient outcomes through advanced predictive modeling. While some limitations, such as dataset size and class imbalance, were identified, the study's results remain robust and valid, paving the way for future research. By addressing these limitations and exploring new techniques, future studies can build on this foundation, ultimately contributing to more effective and reliable diagnostic tools in the medical field. The continued integration of machine learning models into healthcare systems holds promise for early detection and personalized treatment, benefiting patients and healthcare providers alike.

REFERENCES

A. Singh, P. Nath, V. Singhal, D. Anand, Kavita, S. Verma, & T. -P. Hong. (2020). A New Clinical Spectrum for the Assessment of Nonalcoholic Fatty Liver Disease Using Intelligent Methods. *IEEE Access*, *8*, 138470–138480. https://doi.org/10.1109/ACCESS.2020.3011289

Akter, S., Shekhar, H. U., & Akhteruzzaman, S. (2021). Application of Biochemical Tests and Machine Learning Techniques to Diagnose and Evaluate Liver Disease. *Advances in Bioscience and Biotechnology*, *12*(06), 154–172. https://doi.org/10.4236/abb.2021.126011

Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., Aldairem, A., Alrashed, M., Bin Saleh, K., Badreldin, H. A., Al Yami, M. S., Al Harbi, S., & Albekairy, A. M. (2023). Revolutionizing Healthcare: The Role of Artificial Intelligence in Clinical Practice. *BMC Medical Education*, *23*(1). https://doi.org/10.1186/s12909-023-04698-z

Blanes-Vidal, V., Lindvig, K. P., Thiele, M., Nadimi, E. S., & Krag, A. (2022). Artificial Intelligence Outperforms Standard Blood-Based Scores in Identifying Liver Fibrosis Patients in Primary Care. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-06998-8

Chen, K., Sun, J., Shen, J., Luo, J., Zhang, X., Pan, X., Wu, D., Zhao, Y., Bento, M., Ren, Y., & Pu, X. (2022). *GCN-MIF: Graph Convolutional Network with Multi-Information Fusion for Low-dose CT Denoising* (arXiv:2105.07146). arXiv. http://arxiv.org/abs/2105.07146

Chen, Y., Lin, C.-Y., Yen, H., Su, P., Zeng, Y.-H., Huang, S., & Liu, I.-L. (2022). Machine-Learning Algorithm for Predicting Fatty Liver Disease in a Taiwanese Population. *Journal of Personalized Medicine*, *12*(7), 1026. https://doi.org/10.3390/jpm12071026

Da, B. L., Surana, P., Kleiner, D. E., Heller, T., & Koh, C. (2020). The Delta-4 Fibrosis Score (D4FS): A Novel Fibrosis Score in Chronic Hepatitis D. *Antiviral Research*, *174*, 104691. https://doi.org/10.1016/j.antiviral.2019.104691

Ding, H., Fawad, M., Xu, X., & Hu, B. (2022). A Framework for Identification and Classification of Liver Diseases Based on Machine Learning Algorithms. *Frontiers in Oncology*, *12*. https://doi.org/10.3389/fonc.2022.1048348

Ding, N., Qin, Y., Yang, G., Wei, F., Yang, Z., Su, Y.-S., Hu, S., Chen, Y., Chan, C.-M., Chen, W., Yi, J., Zhao, W., Wang, X., Liu, Z., Zheng, H., Chen, J., Liu, Y., Tang, J., Li, J., & Sun, M. (2022). *Delta Tuning: A Comprehensive Study of Parameter Efficient Methods for Pre-Trained Language Models*. https://doi.org/10.21203/rs.3.rs-1553541/v1

Duhayyim, M. A., Mengash, H. A., Marzouk, R., Nour, M. K., Mahgoub, H., Althukair, F., & Mohamed, A. (2022). Hybrid Rider Optimization With Deep Learning Driven Biomedical Liver Cancer Detection and Classification. *Computational Intelligence and Neuroscience*, *2022*, 1–11. https://doi.org/10.1155/2022/6162445

Gupta, N., Mujumdar, S., Patel, H., Masuda, S., Panwar, N., Bandyopadhyay, S., Mehta, S., Guttula, S., Afzal, S., Mittal, R. S., & Munigala, V. (2021). *Data Quality for Machine Learning Tasks*. https://doi.org/10.1145/3447548.3470817

Huang, C., Liu, L., Wang, H., Fang, M., Feng, H., Li, Y., Wang, M., Lin, T., Xiao, X., Wang, Z., Xu, X., He, Y., & Gao, C. (2021). Serum N-Glycan Fingerprint Nomogram Predicts Liver Fibrosis: A Multicenter Study. *Clinical Chemistry and Laboratory Medicine (Cclm)*, *59*(6), 1087–1097. https://doi.org/10.1515/cclm-2020-1588

Iyer, A., Loh, Z., Fitzsimmons, R. L., Reid, R. C., Ramnath, D., Clouston, A. D., Irvine, K. M., Powell, E. E., Schroder, K., Stow, J. L., Sweet, M. J., & Fairlie, D. P. (2019). Histone Deacetylase Inhibitors Attenuate Hepatic Fibrosis Through Suppression of Group 2 Innate Lymphoid Cells and Type 2 Inflammation. *The Faseb Journal*, *33*(S1). https://doi.org/10.1096/fasebj.2019.33.1_supplement.505.19

Kim, S., Park, S., & Lee, H. (2023). Machine Learning for Predicting Hepatitis B or C Virus Infection in Diabetic Patients. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-49046-9

Kirubakaran, J., Venkatesan, G. K. D. P., Baskar, S., Kumaresan, M., & Annamalai, S. (2019). RETRACTED ARTICLE: Prediction of Cirrhosis Disease From Radiologist Liver Medical Image Using Hybrid Coupled Dictionary Pairs on Longitudinal Domain Approach. *Multimedia Tools and Applications*, *79*(15–16), 9901–9919. https://doi.org/10.1007/s11042-019-7259-3

Kuo, Y.-H., Kee, K., Hsu, N.-T., Wang, J., Hsiao, C., Chen, Y., & Lu, S. (2019). Using AST-platelet Ratio Index and Fibrosis 4 Index for Detecting Chronic Hepatitis C in a Large-Scale Community Screening. *Plos One*, *14*(10), e0222196. https://doi.org/10.1371/journal.pone.0222196

Lee, J. H., Joo, I., Kang, T. W., Paik, Y. H., Sinn, D. H., Ha, S. Y., Kim, K., Choi, C.-H., Lee, G., Yi, J., & Bang, W. C. (2019). Deep Learning With Ultrasonography: Automated Classification of Liver Fibrosis Using a Deep convolutional Neural Network. *European Radiology*, *30*(2), 1264–1273. https://doi.org/10.1007/s00330-019-06407-1

Madsen, B. S., Thiele, M., Detlefsen, S., Sørensen, G. L., Kjærgaard, M., Møller, L. S., Rasmussen, D. N.,

Schlosser, A., Holmskov, U., Trebicka, J., Sørensen, G. L., & Krag, A. (2020). Prediction of Liver Fibrosis Severity in Alcoholic Liver Disease by Human Microfibrillar-associated Protein 4. *Liver International*, *40*(7), 1701–1712. https://doi.org/10.1111/liv.14491

Menegotto, A. B., Becker, C. D. L., & Cazella, S. C. (2021). Computer-Aided Diagnosis of Hepatocellular Carcinoma Fusing Imaging and Structured Health Data. *Health Information Science and Systems*, *9*(1). https://doi.org/10.1007/s13755-021-00151-x

Mostafa, F., Hasan, E., Williamson, M., & Khan, H. (2021). Statistical Machine Learning Approaches to Liver Disease Prediction. *Livers*, *1*(4), 294–312. https://doi.org/10.3390/livers1040023

Nam, J. Y., Sinn, D. H., Bae, J. H., Jang, E. S., Kim, J. W., & Jeong, S. H. (2020). Deep Learning Model for Prediction of Hepatocellular Carcinoma in Patients With HBV-related Cirrhosis on Antiviral Therapy. *Jhep Reports*, *2*(6), 100175. https://doi.org/10.1016/j.jhepr.2020.100175

Naseem, R., Khan, B., Shah, M. A., Wakil, K., Khan, A., Alosaimi, W., Uddin, I., & Alouffi, B. (2020). Performance Assessment of Classification Algorithms on Early Detection of Liver Syndrome. *Journal of Healthcare Engineering*, *2020*, 1–13. https://doi.org/10.1155/2020/6680002

Nguyen, D.-K., Lan, C.-H., & Chan, C.-L. (2021). Deep Ensemble Learning Approaches in Healthcare to Enhance the Prediction and Diagnosing Performance: The Workflows, Deployments, and Surveys on the Statistical, Image-Based, and Sequential Datasets. *International Journal of Environmental Research and Public Health*, *18*(20), 10811. https://doi.org/10.3390/ijerph182010811

Nia, N. G., Kaplanoğlu, E., & Nasab, A. (2023). Evaluation of Artificial Intelligence Techniques in Disease Diagnosis and Prediction. *Discover Artificial Intelligence*, *3*(1). https://doi.org/10.1007/s44163-023-00049-5

Philip, G., Djerboua, M., Carlone, D., & Flemming, J. A. (2020). Validation of a Hierarchical Algorithm to Define Chronic Liver Disease and Cirrhosis Etiology in Administrative Healthcare Data. *Plos One*, *15*(2), e0229218. https://doi.org/10.1371/journal.pone.0229218

R. Haluška, J. Brabec, & T. Komárek. (2022). Benchmark of Data Preprocessing Methods for Imbalanced Classification. *2022 IEEE International Conference on Big Data (Big Data)*, 2970–2979. https://doi.org/10.1109/BigData55660.2022.10021118

Rajesh, S., George, T., Philips, C. A., Ahamed, R., Kumbar, S., Mohan, N., Mohanan, M., & Augustine, P. (2020). Transjugular Intrahepatic Portosystemic Shunt in Cirrhosis: An Exhaustive Critical Update. *World Journal of Gastroenterology*, *26*(37), 5561–5596. https://doi.org/10.3748/wjg.v26.i37.5561

Sarker, I. H. (2021). Data Science and Analytics: An Overview From Data-Driven Smart Computing, Decision-Making and Applications Perspective. *Sn Computer Science*, *2*(5). https://doi.org/10.1007/s42979-021-00765-8

Schawkat, K., Ciritsis, A., Ulmenstein, S. von, Honcharova-Biletska, H., Jüngst, C., Weber, A., Gubler, C., Mertens, J. C., & Reiner, C. S. (2020). Diagnostic Accuracy of Texture Analysis and Machine Learning for Quantification of Liver Fibrosis in MRI: Correlation With MR Elastography and Histopathology. *European Radiology*, *30*(8), 4675–4685. https://doi.org/10.1007/s00330-020-06831-8

Shahhosseini, M., & Pham, H. (2022). Optimizing Ensemble Weights and Hyperparameters of Machine Learning Models for Regression Problems.

Machine Learning With Applications, *7*, 100251. https://doi.org/10.1016/j.mlwa.2022.100251

Shekhar, S., Fields, G., Ghavamzadeh, M., & Javidi, T. (2021). *Adaptive Sampling for Minimax Fair Classification* (arXiv:2103.00755). arXiv. http://arxiv.org/abs/2103.00755

Shi, P., Qiu, J., Abaxi, S. M. D., Wu, H., Lo, F. P.-W., & Yuan, W. (2023). Generalist Vision Foundation Models for Medical Imaging: A Case Study of Segment Anything Model on Zero-Shot Medical Segmentation. *Diagnostics*, *13*(11), 1947. https://doi.org/10.3390/diagnostics13111947

Shu, Z., Liu, W., Wu, H., Xiao, M., Wu, D. C., Cao, T., Ren, M., Jin-hua, T., Zhang, C., He, T., Li, X., Zhang, R., & Zhou, X. (2019). Symptom-Based Network Classification Identifies Distinct Clinical Subgroups of Liver Diseases With Common Molecular Pathways. *Computer Methods and Programs in Biomedicine*, *174*, 41–50. https://doi.org/10.1016/j.cmpb.2018.02.014

Singh, S., Hoque, S., Zekry, A., & Sowmya, A. (2023). Radiological Diagnosis of Chronic Liver Disease and Hepatocellular Carcinoma: A Review. *Journal of Medical Systems*, *47*(1). https://doi.org/10.1007/s10916-023-01968-7

Tanwar, N., & Rahman, K. F. (2021). Machine Learning in Liver Disease Diagnosis: Current Progress and Future Opportunities. *Iop Conference Series Materials Science and Engineering*, *1022*(1), 012029. https://doi.org/10.1088/1757-899x/1022/1/012029

Taylor-Weiner, A., Pokkalla, H., Han, L., Jia, C., Huss, R., Chung, C., Elliott, H., Glass, B., Pethia, K., Carrasco-Zevallos, O., Shukla, C., Khettry, U., Najarían, R. M., Taliano, R., Subramanian, G. M., Myers, R. P., Wapinski, I., Khosla, A., Resnick, M. B., … Younossi, Z. M. (2021). A Machine Learning Approach Enables Quantitative Measurement of Liver Histology and Disease Monitoring in NASH. *Hepatology*, *74*(1), 133–147. https://doi.org/10.1002/hep.31750

Wang, Q., & Sun, D. (2021). The Improved AdaBoost Algorithms for Imbalanced Data Classification. *Information Sciences*, *563*, 358–374. https://doi.org/10.1016/j.ins.2021.03.042

Wang, Y., Li, X., Konanur, M., Konkel, B., Seyferth, E. R., Brajer, N., Liu, J., Bashir, M. R., & Lafata, K. (2023). Towards Optimal Deep Fusion of Imaging and Clinical Data via a Model-based Description of Fusion Quality. *Medical Physics*, *50*(6), 3526–3537. https://doi.org/10.1002/mp.16181

Zhang, Z.-M., Tan, J., Wang, F., Dao, F., Zhang, Z.-Y., & Lin, H. (2020). Early Diagnosis of Hepatocellular Carcinoma Using Machine Learning Method. *Frontiers in Bioengineering and Biotechnology*, *8*. https://doi.org/10.3389/fbioe.2020.00254