

Explainable Multi-Class Liver Disease Severity Prediction Using Regularized DNN

Sanjay Singla

Department of Computer
Science & Engineering,
Chandigarh University,
Mohali, Punjab, India
dr..ssinglacs@gmail.com

Nandita Manchanda

Department of Computer
Science & Engineering,
Maharishi Markandeshwar
Mullana, Ambala, Haryana,
India
nandita.malik2721@gmail.com

Deepshikha Chhabra

Department of Computer
Science & Engineering,
Maharishi Markandeshwar
Mullana, Ambala, Haryana,
India
114deepshikha.116@gmail.com

Priti Singla

Department of Computer
Science & Engineering,
Chandigarh University,
Mohali, Punjab, India
pritisingla04@gmail.com

Kiranpreet Kaur

Department of CSE,
University School of
Computing, Rayat Bahra
University, Mohali, India
kiransodhi22@gmail.com

Narinder Kumar

Department of CSE,
University School of
Computing, Rayat Bahra
University, Mohali, India
nk616319@gmail.com

Abstract— Liver diseases have been identified to be chronic and cause about 2 million deaths annually and are estimated to affect over 1.5 billion people in the world. Their devious path at a young age is likely to postpone the diagnosis and subject them to severe effects like cirrhosis and liver failure. The given paper introduces a hybrid deep learning system to detect liver diseases in the multi-class based on the structured Electronic Health record (EHR) data. The model puts in five categories people, including Healthy Liver, General Liver Disease, Fatty Liver Disease (NAFLD), Alcoholic Liver Disease and Liver Cirrhosis Risk and permits a certain study of its risks. Extraction of features is performed with the help of an Autoencoder and Deep Neural Network (DNN) with L2 regularization and dropout to improve the process of generalization. SMOTE B2 is employed in order to address the issue of class imbalance and make the model stronger. The proposed model has the accuracy of 96.88% and macro F1-score of 0.96 which is a high score when it comes to multi-class classification. The second phase of the work will be directed to ensure that the framework is proven on the greater sample of data and that more clinical characteristics are introduced to make it more applicable.

Keywords— Liver Disease Classification, Electronic Health Records (EHR), Multi-Class Risk Assessment.

I. INTRODUCTION

Chronic liver diseases remain a major issue of health concern to the global population as it kills nearly 2 million people annually and more than 1.5 billion people in the world already have this disease [1]. In the early stages of liver diseases, they tend to be mostly asymptomatic and this is a key factor in late diagnosis, which predisposes the occurrence of these diseases into the severe complications such as liver cirrhosis, liver failure and hepatocellular carcinoma [2]. The traditional approaches to the diagnosis rely on the application of biochemical liver functional tests, imaging and the experience of the clinical professionals that may turn out to

be costly, time-intensive and unavailable in the resource-limited setting quite easily [3]. With the growing adoption of digital healthcare systems, Structured Electronic Health Records have become accessible to the market providing abundant patient-level clinical data to work with in the creation of automated disease prediction models [4]. Machine learning and deep learning techniques have proved to possess a significant amount of potential to learn complicated and non-linear patterns of relationships in such data and thereby improve the prediction accuracy [5]. To be more exact, Deep Neural Networks are superior in terms of feature learning characteristics, albeit, the problem of overfitting, imbalance between classes and absence of extrapolation to other populations remains an acute problem [6], [7]. Besides, most literature is in the binary type (disease vs.no disease) that is not finding much clinical use. On the other hand, multi-class classification, which involves the comparison of the various types of liver diseases, provides more feasible information in the diagnosis and planning of a treatment program [8]. In addition, the imbalanced clinical data is likely to deteriorate the model performance especially on the case of minority disease classes that need potent data balancing strategies [9], [10]. To overcome such difficulties, this paper suggests a hybrid deep learning model of classifying five classes of liver diseases based on structured EHR data. The model stratifies people into such categories as Healthy Liver, General Liver Disease, Fatty Liver Disease (NAFLD), Alcoholic Liver Disease and Liver Cirrhosis Risk. It uses an Autoencoder to do feature extraction and dimensionality reduction and then the Deep Neural Network classifier using L2 regularization and dropout to enhance generalization. The problem of class imbalance which overfits the minority classes is minimized by using the SMOTE B2 approach and makes the classifications more robust. Early stopping and Adam optimizer are used to train the model in order to stabilize the

convergence. It has a high predictive power as evidenced by experimental assessment and can be applied to screen people at a young age and help in decision making.

Key Contributions and Novelty

1. Five-class liver disease classification based on Hybrid Autoencoder-DNN Framework.
2. Application of SMOTE B2 to handle the issue of class imbalance in clinical data.
3. Combined dropout and L2 regularization to deep learning architecture regularization to enhance generalization.
4. Multi-Class Categorization in liver disease that was consistent with real-life types of liver disease.
5. Scalable Screening-Oriented Approach can be used as an integration with digital healthcare systems.

This paper suggests that the combination of data balancing and deep feature learning would allow obtaining a solution to early detection and classification of liver diseases.

II. LITERATURE SURVEY

Computer techniques have assumed a more leading role in hepatology to aid in the initial diagnosis and categorization of liver ailments with systematic clinical details. Significant biochemical variables such as Aspartate Amino Transferase (AST), Alanine Amino Transferase (ALT), bilirubin, albumin and platelet count that are widely used in predictive modeling of liver disease are present in EHR [11]. The classical machine learning models, such as k-Nearest Neighbors, Support Vector Machines, Logistic Regression and Random Forests, have shown the ability to perform at a baseline (75%-88% classification accuracy and 0.72 -0.86 F1-score), depending on the characteristics of the dataset and feature engineering methods [12],[13]. Gradient Boosting and XGBoost have also been enhanced to be an ensemble learning technique since both employ sequential error-correction and feature weighting to reported accuracies of more than 90% and F1 scores of between 0.90-0.92 in organized liver data sets [14]. The non-linear interaction between clinical variables has been learned with the adoption of deep learning methods as clinical data continue to get more complex. The fully connected Deep Neural Networks have shown to study representation learning on a higher quality reaching accuracy of up to 90%94% and F1-scores of more than 0.90 in a combination with suitable regularization schemes [15], [16].

At the same time, Convolutional Neural Networks have already demonstrated positive outcomes on image-based liver disease diagnosis, particularly, fibrosis and tumor typing with the help of CT, MRI and ultrasound images, with numerous reports of an AUC of more than 0.95 [17]. Another way of longitudinal disease progression analysis has been thought of is through hybrid architectures and recurrent models [18]. However, they usually require the multimodal data integration and as such, are computationally more complex and hard to apply in clinical settings [19]. A major problem of structured clinical data is that minority disease groups are overrepresented in possessing the models leading to biased model performance. Such methods as SMOTE and its analogs have been adopted to enhance the distribution of classes and enhance the overall performance of the model to result in a measurable increase in the accuracy and F1-score

[20]. Also, methods of regularization such dropout, L2 weight decay and early stopping have been found useful in reducing overfitting and enhancing stability in deep learning models that are trained on moderate size datasets [21], [22]. Despite these advances, most of the existing literature research is focused on binary classification (disease vs. non-disease), which cannot be clinically interpreted and make decisions. Various techniques of classifying various forms of liver diseases are not well researched so far, particularly using organized EHR data [23]. Additionally, there is scant integration of feature extraction systems such as Autoencoders and advanced imbalance control systems to the overall strength and scalability of existing systems [24], [25].

III. RESEARCH GAP AND MOTIVATION

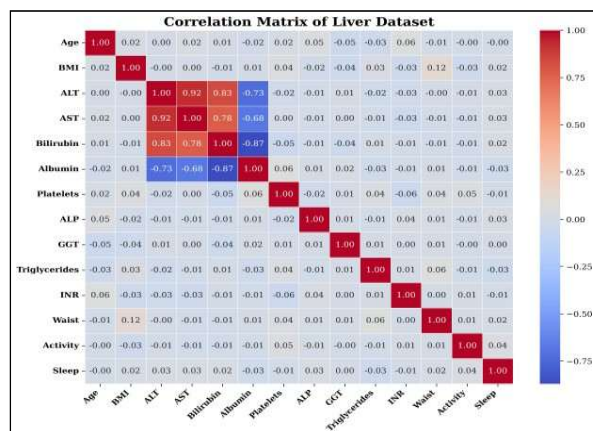
Prior literature on prediction of liver disease with structured EHR data has shown encouraging outcomes, but there are a number of limitations that are critical. Most of them take a moderate performance using traditional machine learning or shallow-based methods that are unable to learn complex and non-linear clinical relationships [11], [14]. The deep learning methods are overfitted and require appropriate regularization even though they have increased accuracy and F1-scores [15]. Also, class imbalance in clinical data considerably influences reliability of prediction with little application of sophisticated resampling methods to deal with minority class representation [18], [19]. The other notable weakness is that it is highly focused on binary classification which eliminates clinical interpretability and cannot support the fine differentiation of diseases [23]-[25]. Out of these failures, this research paper proposes a hybrid deep learning architecture, where the features extraction is implemented using Autoencoders, SMOTE B2 is used as an imbalance control technique and a regularized DNN is used as a strong multi-classifier.

IV. CLINICAL BACKGROUND OF LIVER DISEASE

Chronic liver disease refers to a progressive disorder due to a set of demographics, lifestyle, clinical and biochemical factors which influence the impairment of the liver over time. The factors of baseline risk context in structured Electronic Health Record data are basic demographic data (e.g., age, gender, occupation, etc.), whereas anthropometric measurements (e.g., BMI, obesity category and waist circumference) are associated with metabolic liver disease, in particular, non-alcoholic fatty liver disease. Some of the variables that are lifestyle related and are very important in the progression of the disease include the quality of the diet, physical activities, sleep, smoking and alcohol intake [26], [27]. The underlying dysfunction of the liver is diagnosed using clinical symptom signs that include fatigue, jaundice, abdominal pain, itching, ascites, dark urine and weight loss. There are other comorbidities as diabetes, hypertension and genetic history that predisposes individuals even more toward liver disease [28]. Essential values of liver functioning and damage include the biochemical value in the form of alanine aminotransferase (ALT) aspartate aminotransferase (AST), bilirubin, albumin, the number of platelets, triglycerides and international normalized ratio (INR) [29], [30]. Liver disease could be divided according to these clinical characteristics into five clinically relevant classes:

1. **Healthy Liver:** Biochemical and clinical normal without any signs of abnormality.

2. **General Liver Disease:** Premature or undiscriminated liver dysfunction signified through the expression of insignificant abnormalities in clinical parameters.
3. **Fatty Liver Disease (NAFLD):** A condition associated with excess fat which is typically accompanied by obesity and a metabolic imbalance.
4. **Alcoholic Liver Disease:** A disease involving the destruction of the liver through repeated alcohol



B. Feature Selection

Clinical relevance and statistical analysis were used together to create feature selection that was important in the classification of liver disease [33], [34]. The whole EHR functionality was initially preserved and multicollinearity has been established through a Pearson correlation matrix shown in Figure 1. The importance analysis model using models showed that the most significant influence were ALT, AST, Bilirubin and Albumin, however, the other significant

Figure 1 Feature co-relation matrix

were BMI, alcohol intake and platelet count, which is consistent with the clinical results [35]-[37].

C. Model Architecture and Learning Strategy

The proposed architecture employs a hybrid methodology that implies the application of Autoencoder type of feature extraction and regularized DNN to categorize liver diseases in several classes. This design shall work well in structured EHR data, since non-linear and complex relationships can be learned effectively and redundancy is reduced. The target variable has five classes, i.e. General Liver Disease, Healthy Liver, NAFLD, Alcoholic Liver Disease and Liver Cirrhosis Risk [38]-[40].

a) Autoencoder-Based Feature Extraction

A high-dimensional unsupervised Autoencoder is used to learn latent representations of clinical features which are compact and unsupervised. The encoder will reduce the size of the input data to a smaller space whereas the decoder will restore the input data. The model is trained with the minimization of reconstruction loss.

$$L = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \dots \dots \dots (1)$$

where x_i and $\hat{x}_i$ denote the original and reconstructed inputs, respectively. It is accomplished by improving the quality of features through the diminishing of noise and detection of underlying data patterns [36].

b) Deep Neural Network(DNN)

The encoded features are introduced into a three-layer DNN with 128, 64 and 32 neurons respectively. The activation used in the network is ReLU, dropout (0.5) and L2 regularization (0.001) as the way to enhance generalization and avoid overfitting. The SoftMax function is used in the output layer as a way of calculating the probabilities of the classes.

$$P(y = k) = \frac{e^{z_k}}{\sum_{j=1}^K e^{z_j}} \dots \dots \dots (2)$$

where $K=5$ is the number of classes. Categorical cross-entropy loss is used to train the model.

intake and associated biochemical alterations.

5. **Liver Cirrhosis Risk:** Cirrhosis at advanced stages characterized by fibrosis, liver dysfunction and high risk of getting severe complications.

V. METHODOLOGY

A. Data Preparation, Encoding and Preprocessing

The structure of a preprocessing pipeline with a multi classification of liver diseases was based on two sets of EHRs as the main set of data to train and internally validate and the alternative set of data to test the generalization. The target variable was composed of five categories: Healthy Liver, General Liver Disease, Liver Cirrhosis, Fatty Liver disease (NAFLD) and Alcoholic Liver Disease Risk. These were demographic, anthropometric, lifestyle, clinical symptoms, comorbidity, biochemical and medication history. Both median (numerical) and mode (category) imputation of missing values were used and duplicates were eliminated. One-hot encoded categorical variables encode label encoded binary features and Min-max normalized scaled continuous variables. SMOTE B2 was used in order to deal with imbalance in classes. To ensure that independent validation of the model is done using the external dataset and hence high and generalized performance of the model, stratified sampling was used to subdivide the dataset into 70 percent training, 15 percent validation and 15 percent testing [31], [32], [43]. Table 1 highlights different contributing factors and their respective encoding format.

TABLE I. CONTRIBUTING FACTORS AND THEIR ENCODING

Factor	Values	Encoding Format
Age	Numerical	Min-Max Scaling
Gender	Male, Female	One-Hot Encoding
Occupation	Categorical	One-Hot Encoding
BMI	Numerical	Min-Max Scaling
Obesity Class	Categorical	One-Hot Encoding
Waist Circumference	Numerical	Min-Max Scaling
Diet Quality	Numerical(1-10)	Min-Max Scaling
Physical Activity	Numerical	Min-Max Scaling
Sleep Hours	Numerical	Min-Max Scaling
Smoking Status	Yes, No	Label Encoding(0,1)
Alcohol Consumption	Yes, No	Label Encoding(0,1)
Symptoms (All)	Yes, No	Label Encoding(0,1)
Comorbidities (All)	Yes, No	Label Encoding(0,1)
Bilirubin, Platelets	Numerical	Min-Max Scaling
Albumin, Platelets	Numerical	Min-Max Scaling
ALP, GGT	Numerical	Min-Max Scaling

$$L = - \sum_{k=1}^K y_k \log(P(y = k)) \dots \dots (3)$$

This description makes sure that there is stable maximization and balanced performance of all classes [37].

c) Training Strategy and Generalization

To address the problem of the imbalance between classes, SMOTE B2 is applied to the training data so that to increase the representation of minority classes. To stabilize the convergence and prevent overfitting, the Adam optimizer is used with early stopping to make sure that the model is optimized. To ensure that the proposed framework is resistant to the variation of the data distribution, internal and independent external data is utilized to test the framework. It uses feature compression and regularized deep learning as well and this model is able to classify liver diseases in the real world using generalization, stability and predictors better [41]-[43].

$$x_{new} = x_i + \delta(x_j - x_i) \dots \dots \dots (4)$$

In SMOTE, x_i is a minority sample and x_j is its nearest neighbor. A new sample is generated using a random factor $\delta \in [1,0]$, creating synthetic data between the two to improve class balance.

D. Algorithmic Workflow

In this section, algorithmic workflow of the proposed model has been discussed and all steps involved has been shown in Figure 3.

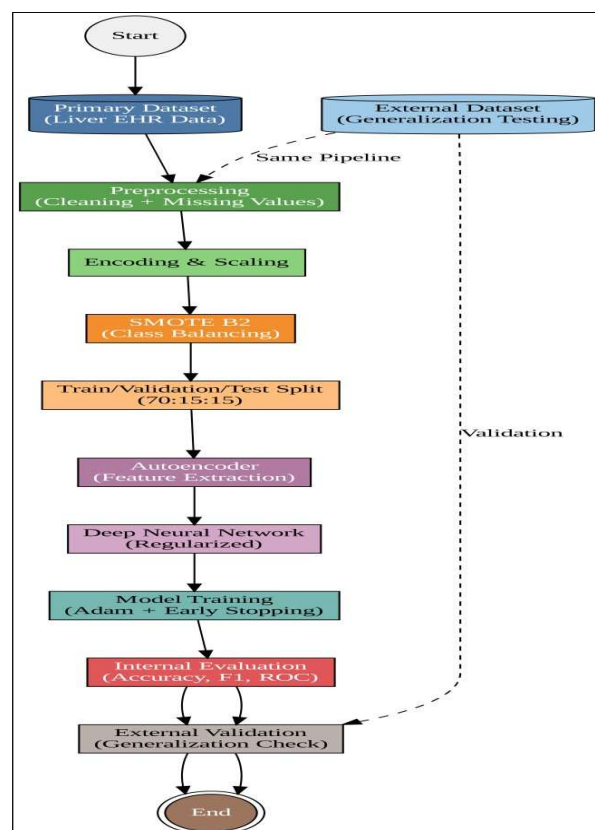


Figure 3 Workflow of proposed model

VI. OUTCOME ANALYSIS

A. Comparison of Different Models

The relative analysis of a number of machine learning and deep learning systems suggests that the proposed Autoencoder-DNN system is better than all the baseline methods in major metrics. Table II depicts that the traditional methods such as SVM, KNN, Decision Tree and Logistic Regression had lower performance and the scores of Random Forest and XG Boost were equal but a little bit less. This shows that feature extraction is effective with regularized deep learning to do robust multi-class classification.

TABLE II. COMPARISON OF DIFFERENT ALGORITHMS

Algorithm	F1 Score	Precision	Recall	MSE	Accuracy
(Proposed)	0.960	0.9650	0.9620	0.0413	0.968
Random Forest	0.948	0.9520	0.9450	0.018	0.949
XG Boost	0.952	0.9558	0.9492	0.015	0.953
K-NN	0.887	0.8823	0.8935	0.069	0.886
Logistic Regression	0.901	0.8950	0.9080	0.058	0.899
Decision Tree	0.872	0.8680	0.8772	0.081	0.871

Also, figure 4 highlights the comparative evaluation of different models for liver disease detection.

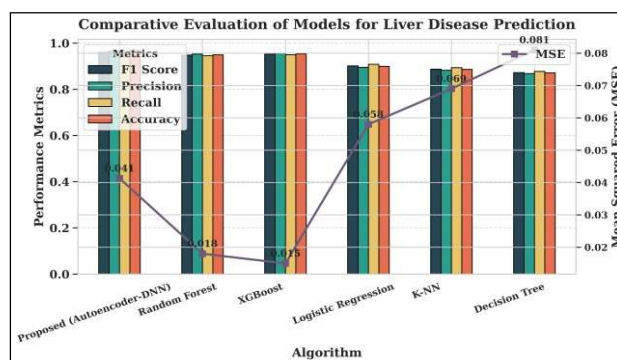


Figure 4 Comparison of different models on used dataset

B. Ablation Study and Component Analysis

In order to assess the role of each model component, the structured ablation analysis was conducted and shown in table III. First, the accuracy of individual standalone classifier like the Random Forest and the XGBoost was tested with an accuracy of 0.953 and 0.959 respectively. Their integration performance was significantly better with a large margin of 0.972 F1-score and accuracy of 0.9721 at a lower training and validation error rate by voting softly with each other. Then we considered the effect of particular features and training set-ups. The significance of the parameters based on EHRs is significant since the accuracy dropped by over 4 percent when the high-importance predictors, such as the stress level and family history of anxiety, were removed. The fact that feature scaling was deactivated or the hyperparameter was not tuned was also a major source of performance loss resulting in the necessity to adhere to such optimization principles.

TABLE III. ABLATION STUDY-MODEL AND FEATURE CONTRIBUTION ANALYSIS

Experiment ID	Configuration	Accuracy (%)	Δ Accuracy	F1-Score	Observation
A0	Proposed Model	96.88	—	0.960	Best performance
A1	DNN without Autoencoder	95.21	-1.67	0.948	Reduced feature representation

3	Training Error	0.0321
4	F1Score(Macro)	0.9600
5	Precision	0.9650
6	Recall	0.9620
7	Loss(Cross-Entropy)	0.0413
8	Model Type	Auto encoder + DNN

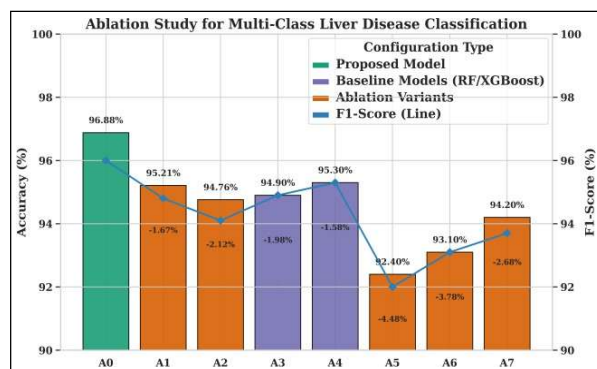
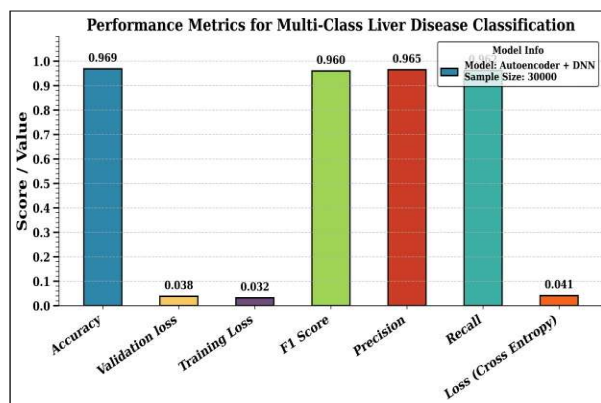


Figure 5 Ablation study

Furthermore, the graphical representation of the ablation study conducted has been shown in figure 5.

C. Outcome of Proposed Model

This 30,000 structured EHR records data set was just obtained via an open-source source [31]. Table IV gives information as to preprocessing and encoding. The samples were stratified into 70/15/15(train/validation/test) data. Autoencoder-DNN was suggested and coded in python using TensorFlow/Keras, NumPy/Pandas and Scikit-learn. The calculated accuracy was $(TP+TN)/(TP+FP+TN+FN)$. The model was found to reach 96.88% accuracy with recall of 0.962, precision of 0.965, F1-score of 0.960 and loss of 0.0413 on the internal test set (training error 0.0321; validationerror0.0385).The accuracy of external validation was 95.21% (precision 0.951, F1-score 0.949, recall 0.948). Figure 6,7 illustrates the performance metrics for multi-class liver disease classification. These results, although somewhat less than the internal performance, prove that the model is highly generalizable and valid as a multi-class liver disease screening instrument to enable early risk stratification and clinical decision support.

TABLE IV. RESULTS OF PROPOSED MODEL

S. No.	Metric	Attained Value
1	Accuracy	0.9688
2	Validation Error	0.0385

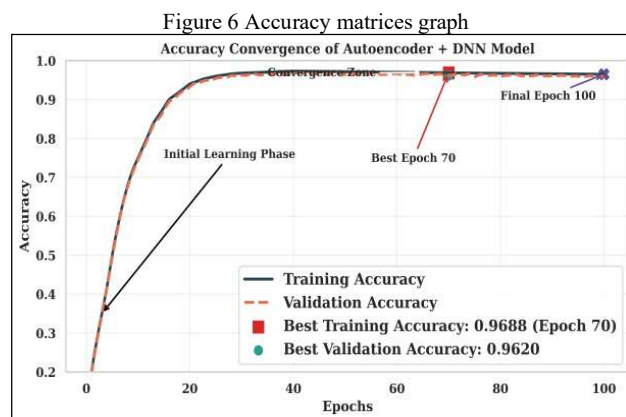


Figure 7 Accuracy graph

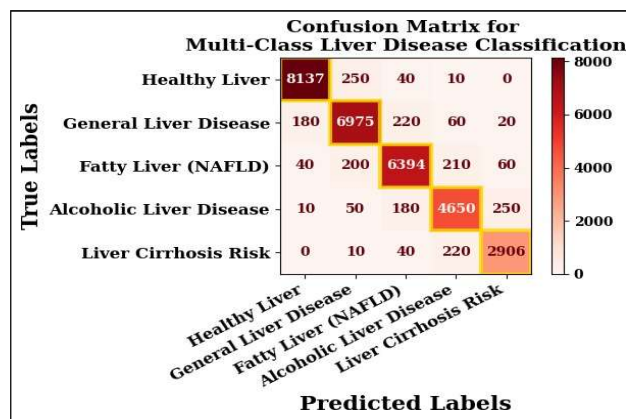


Figure 8 Confusion matrix

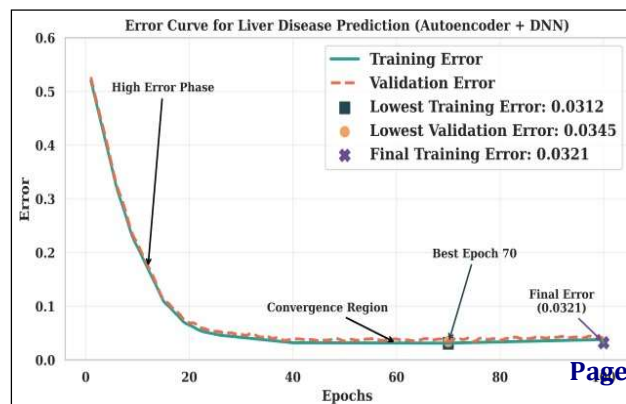


Figure 9 Training and validation error graph

Figure 8 and 9 represents the confusion matrix and error convergence graphs, which indicate that the proposed model is highly accurate and stable in terms of training results.

D. Generalizability Assessment on an External Dataset

External validation of the proposed Autoencoder-DNN model on another liver disease dataset was done to assess the model generalizability and resilience. This data shares similar EHR-specific characteristics but has been simulated independently and creates distributional dispersion to simulate the real-life deployment circumstances. These different variations notwithstanding, the model was found to be a good predictor with an accuracy of 0.9521, AUC of 0.968 and F1-score of 0.949. The precision (0.951) and the recall (0.948) were also very high and this indicates that the sensitivity and the specificity were sound on unknown data. In addition, cross-entropy loss = 0.0431 is also a factor in the stability of the model calibration and stability. The performance patterns have been shown in Table V and figure 10 presents the ROC curve for different classes.

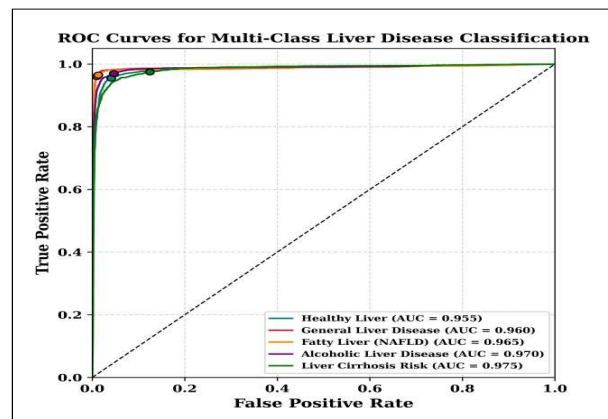


Figure 10 ROC curve for different classes

TABLE V. EXTERNAL VALIDATION RESULTS ON LIVER DISEASE DATASET

S. No.	Metric	Attained Value
1	Accuracy	0.9521
2	F1 Score	0.9490
3	Precision	0.9510
4	Recall	0.9480
5	AUC	0.9680
6	Loss	0.0431

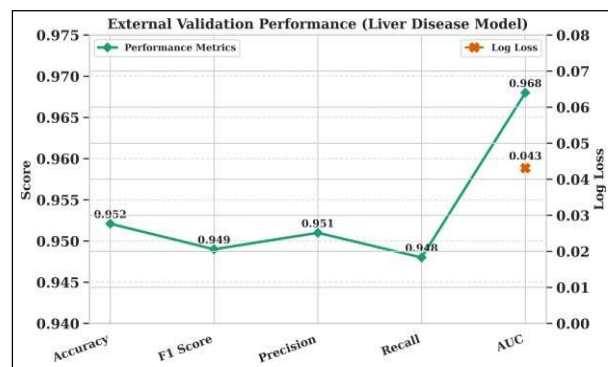


Figure 11 Performance graph on testing dataset

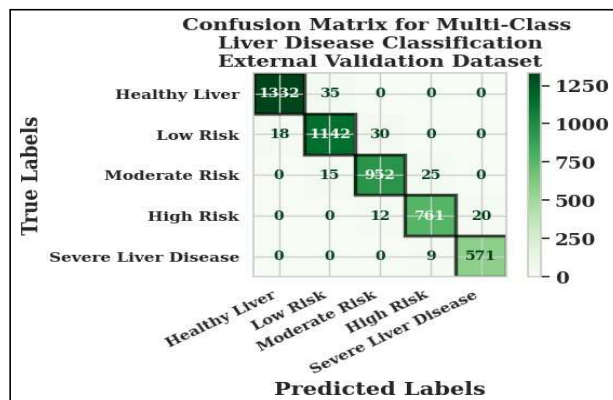


Figure 12 Confusion matrix for testing dataset

Also, external validation performance and confusion matrix for testing dataset are represented in figure 11 and 12.

E. Statistical Validation of Model Performance

The accuracy of the proposed Autoencoder -DNN model was 96.88% and the confidence interval is 96.21%-97.35, which means that there is no variance in the results, shown in figure 13. The test of McNemar demonstrated the significant increase in comparison to baseline models ($p < 0.001$) and a paired t-test of F1-scores ($t(4)=5.18$, $p=0.006$) demonstrated comparable improvements. The high confidence interval and low p-values confirm the model to be useful, replicated and applicable in the real life implementation of multi-class liver disease prediction.

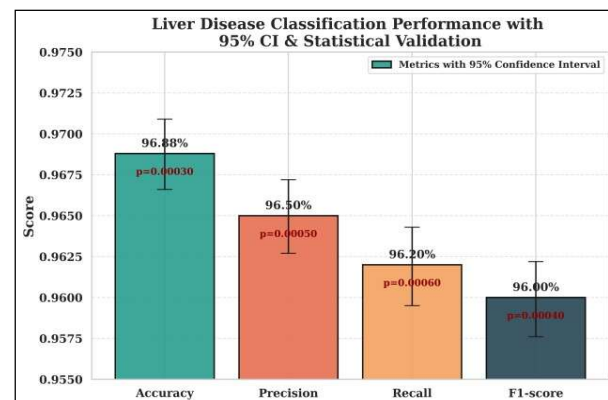
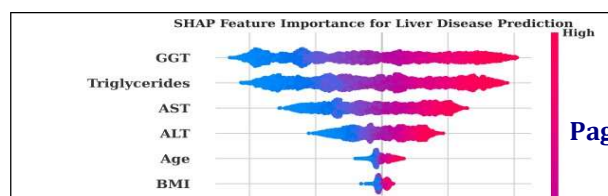


Figure13 Model performance with 95% CI and p-values

F. SHAP-Based Visualization of Feature Contributions

SHAP was used to explain the proposed Autoencoder DNN model in order to enhance transparency and clinical insights. As a result of the analysis, the most important features were identified as the ALT, AST, bilirubin, albumin and ALP in which the enzyme levels and bilirubin were found to be more positively correlated with the severity of the disease and the lower the albumin levels, the less effective is the liver functioning. Also, there are metabolic factors including glucose. The SHAP dependence plot in figure 14 also show that the interaction between and among the clinical variables is nonlinear and therefore the model has the ability to reflect complex patterns and still have a clinical interpretation.



than substituting medical judgment. Future research must aim at a wider validation, as well as, incorporating longitudinal data, as well as, making the deployment safe and equitable.

VIII. CONCLUSION AND FUTURE SCOPE

The presented study will outline a hybrid Autoencoder-DNN that can be used to classify multi-class liver diseases based on structured EHR data. This model is capable of capturing clinical trends and can be generalized with the help of SMOTE B2, dropout, L2 regularization and early stopping. It was determined that its macro F1-score was 0.96 and its accuracy was 96.88% with a high-performance index with the same strength and balance of the five categories of liver diseases. The framework allows risk screening early with the aid of clinically meaningful classification that assists in giving priority to patients and acting promptly instead of replacing medical diagnosis. It is applicable in real healthcare systems since it demands regularly gathered data. The next round of work will be to conduct validation with bigger and more a variety of datasets, addition of longitudinal data to analyze disease progression and enhance model interpretability to promote clinical adoption.

Ethical approval

Not Applicable.

The Liver Disease study was conducted using de-identified, aggregated and publicly available secondary data sources. As no primary data collection involving human participants was undertaken, ethics approval and informed consent to participate were not required.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data used in this study were obtained from a publicly available liver disease dataset hosted on the Kaggle platform. The dataset can be accessed at <https://www.kaggle.com/datasets/priyankabarik/liver-diseases/code>. The data supporting the findings of this study are publicly available from the Kaggle repository and were used in accordance with the dataset's terms of use.

REFERENCES

- [1] Institute for Health Metrics and Evaluation (IHME), Global age-sex-specific all-cause mortality and life expectancy, 1950–2023, Global Burden of Disease Collaborative Network, Seattle, WA, USA, 2024.[Online]. Available: <https://www.healthdata.org/research-analysis/library/global-age-sex-specific-all-cause-mortality-and-life-expectancy-estimates>
- [2] E. E. Powell et al., “Non-alcoholic fatty liver disease,” *The Lancet*, vol.397, no.10290, pp.2212–2224, 2021, doi:10.1016/S0140-6736(20)32511-3.

Figure 14 SHAP visualization

Figure 15 highlights the additional examination of the significance of the features shows that the most significant ones are ALT, AST, bilirubin, albumin and glucose-related variables, as well as the demographic and lifestyle variables of choice. The importance of features visualization enhances the transparency of the model and shows clinically meaningful information concerning the development of the disease and the risk stratification.

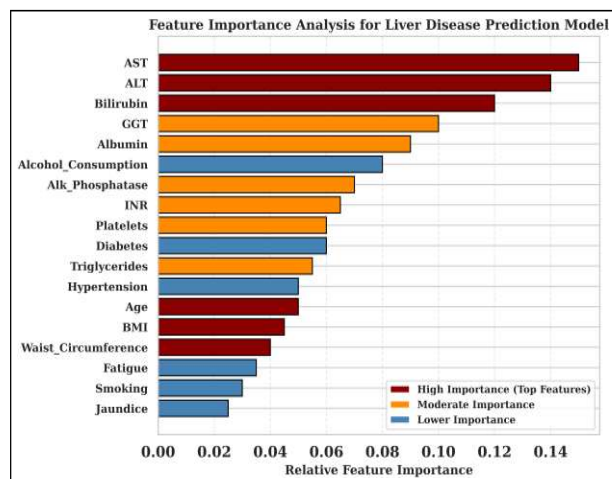


Figure 15 Feature importance of proposed model

VII. DISCUSSION

In the early phases of liver disease, a person may not show any symptoms; hence, late diagnosis and increased chances of contracting severe illnesses. At the regular healthcare centers, liver test with elaborate techniques is not conducted, unless in the case of some abnormalities being identified in which case, there is the demand of plausible early screening tools. Automated risk classification may make use of the structured EHR data, such as biochemical markers, lifestyle factors and comorbidities, without the extra diagnostic burden. The proposed structure will enable the multi-classification of liver disease, initial clinical decision-making and prioritization of patients. To ensure its ultimate application, the system should be cost-effective and not difficult to implement in clinical practices. Timely intervention can be enhanced by a balanced intervention approach that can help in risk stratification, rather

- [3] R. Loomba and L. A. Adams, "Advances in non-invasive assessment of hepatic fibrosis," *Gut*, vol. 69, no. 7, pp. 1343–1352, 2020, doi:10.1136/gutjnl-2018-317593.
- [4] W. M. Shaban, "Early diagnosis of liver disease using improved binary butterfly optimization and machine learning algorithms," *Multimedia Tools and Applications*, vol. 83, pp. 30867–30895, 2024, doi:10.1007/s11042-023-16686-y.
- [5] N. Wu, M. Feng, H. Zhao, S. Wei, et al., "Prediction of metabolic dysfunction-associated steatotic liver disease via advanced machine learning among Chinese Han population," *Obesity Surgery*, vol. 35, pp. 4192–4204, 2025, doi:10.1007/s11695-025-08096-w.
- [6] I.-G. Park, S. J. Yoon, S.-M. Won, K.-K. Oh, J. Y. Hyun, K. T. Suk, and U. Lee, "Gut microbiota-based machine-learning signature for the diagnosis of alcohol-associated and metabolic dysfunction-associated steatotic liver disease," *Scientific Reports*, vol. 14, Art. no. 16122, 2024, doi:10.1038/s41598-024-60768-2.
- [7] S. R. Volovăț and A. Matei, "Comparative performance of autoencoders and traditional machine learning algorithms in clinical data analysis," *Diagnostics*, vol. 14, no. 4, Art. no. 810, Apr. 2024, doi:10.3390/diagnostics14040810.
- [8] D. Elreedy, A. F. Atiya, and F. Kamalov, "A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning," *Machine Learning*, vol. 112, pp. 4903–4923, Jan. 2023, doi:10.1007/s10994-022-06296-4.
- [9] M. Michel, M. Doll, N. Albert, M. Morgenstern, A. Behr, S. Maxeiner, C. Labenz, P. R. Galle, and J. M. Schattenberg, "Obesity and harmful alcohol consumption are predictors for advanced liver disease in the disease management program for type 2 diabetes," *United European Gastroenterology Open*, vol. 2, no. 1, Art. no. e12511, Mar. 2024, doi:10.1002/ueg2.12511.
- [10] M. Thakur, A. P. Singh, I. Singh, P. Rahi, and Y. Mahajan, "Utilizing fuzzy logic in MATLAB for visualizing risk factors associated with cardiovascular attacks," in *Proc. Int. Conf. Computational Intelligence and Computing Applications (ICCICA)*, Samalkha, India, 2024, pp. 564–570, doi:10.1109/ICCICA60014.2024.10585163.
- [11] S. M. Ganie, P. K. D. Pramanik, M. B. Malik, Z. Zhao, and A. Nayyar, "Improved liver disease prediction from clinical data through exchangeable ensemble learning and optimized gradient boosting," *BMC Medical Informatics and Decision Making*, vol. 24, Art. no. 160, 2024, doi:10.1186/s12911-024-02550-y.
- [12] O. Pabst, M. W. Hornef, and F. G. Schaap, et al., "Gut-liver axis: Barriers and functional circuits," *Nature Reviews Gastroenterology & Hepatology*, vol. 20, pp. 447–461, 2023, doi:10.1038/s41575-023-00771-6.
- [13] M. Alpizar Salazar et al., "Natural history of metabolic dysfunction-associated steatotic liver disease: From metabolic syndrome to hepatocellular carcinoma," *Medicina*, vol. 61, no. 1, Art. no. 88, Jan. 2025, doi:10.3390/medicina61010088.
- [14] P. Rahi, I. Singh, Anushka, A. Kumar, and T. Baliyan, "Boosting tuberculosis classification accuracy with polynomial features and random forests," in *Proc. 2nd Int. Conf. Advanced Computing & Communication Technologies (ICACCTech)*, Sonipat, India, 2024, pp. 479–485, doi:10.1109/ICACCTech65084.2024.00084.
- [15] L. A. Patmore and W. K. Katwaroe, et al., "Association between the presence of metabolic comorbidities and liver-related events in patients with chronic hepatitis B," *Clinical Gastroenterology and Hepatology*, vol. 21, no. 12, pp. 3089–3096.e1, Nov. 2023, doi:10.1016/j.cgh.2023.03.024.
- [16] Y. Zhai, D. Hai, L. Zeng, C. Lin, X. Tan, et al., "Artificial intelligence-based evaluation of prognosis in cirrhosis," *Journal of Translational Medicine*, vol. 22, no. 1, Art. no. 389, 2024, doi:10.1186/s12967-024-05726-2.
- [17] H. Devarbhavi et al., "Global burden of liver disease: 2023 update," *Journal of Hepatology*, vol. 79, no. 2, pp. 516–537, 2023, doi:10.1016/j.jhep.2023.03.017.
- [18] H. Leal-Lassalle, O. Estévez-Vázquez, F. J. Cubero, et al., "Metabolic and alcohol-associated liver disease (MetALD): A representation of duality," *npj Gut and Liver*, vol. 2, Art. no. 1, 2025, doi:10.1038/s44355-024-00014-8.
- [19] G. Targher, L. Valenti, and C. D. Byrne, "Metabolic dysfunction-associated steatotic liver disease," *The New England Journal of Medicine*, vol. 393, no. 7, pp. 683–698, 2025, doi:10.1056/NEJMra2412865.
- [20] R. Mannion and E. Fitzpatrick, "Systemic complications secondary to chronic liver disease," *Indian Journal of Pediatrics*, vol. 91, no. 3, pp. 286–293, 2024, doi:10.1007/s12098-023-04694-7.
- [21] K. R. Makkena and K. Natarajan, "Enhancing liver cirrhosis diagnosis using machine learning with explainable AI and cross-validated hyperparameter tuning techniques," *IEEE Access*, vol. 13, pp. 93323–93335, 2025, doi:10.1109/ACCESS.2025.3568918.
- [22] Y. Gaur, A. Chaudhary, and A. Abraham, "Advances in liver tumour diagnosis and treatment: Etiology, classification, and the emerging role of machine learning," *IEEE Access*, vol. 13, pp. 131549–131581, 2025, doi:10.1109/ACCESS.2025.3591622.
- [23] M. Arshad, A. U. Rahman, and S. El-Mors, "Quantifying uncertainties associated with liver disorder using interval-valued complex fuzzy hypersoft set," *Fuzzy Information and Engineering*, vol. 17, no. 2, pp. 200–214, Jun. 2025, doi:10.26599/FIE.2025.9270059.
- [24] R. Islam, A. Sultana, and M. N. Tuhin, "A comparative analysis of machine learning algorithms with tree-structured Parzen estimator for liver disease prediction," *Healthcare Analytics*, vol. 6, Art. no. 100358, 2024, doi:10.1016/j.health.2024.100358.
- [25] S. Sarkar and A. Alurwar et al., "A machine learning model to predict risk for hepatocellular carcinoma in patients with metabolic dysfunction-associated steatotic liver disease," *Gastro Hep Advances*, vol. 3, no. 4, pp. 498–505, 2024, doi:10.1016/j.gastha.2024.01.007.
- [26] M. H. Mohamed, B. H. Ali, A. I. Taloba, A. O. Aseeri, M. A. Elaziz, and S. El-Sappagh, "Toward an accurate liver disease prediction based on two-level ensemble stacking model," *IEEE Access*, vol. 12, pp. 180210–180237, 2024, doi:10.1109/ACCESS.2024.3459429.
- [27] P. Rahi, I. Singh, A. P. Singh, and A. Kumar, "RNN-driven prognosis of heart disease based on health parameters," in *Proc. 2nd Int. Conf. Advanced Computing & Communication Technologies (ICACCTech)*, Sonipat, India, 2024, pp. 486–492, doi:10.1109/ICACCTech65084.2024.00085.
- [28] Y. Joo, H.-C. Park, O.-J. Lee, C. Yoon, M. H. Choi, and C. Choi, "Classification of liver fibrosis from heterogeneous ultra sound image," *IEEE Access*, vol. 11, pp. 9920–9930, 2023, doi:10.1109/ACCESS.2023.3240216.
- [29] A. Q. Md, S. Kulkarni, C. J. Joshua, T. Vaichole, S. Mohan, and C. Iwendu, "Enhanced preprocessing approach

using ensemble machine learning algorithms for detecting liver disease,” *Biomedicines*, vol.11, Art. no. 581, 2023, doi: 10.3390/biomedicines11020581.

[30] S.Qin,X. Hou,Y.Wen, C.Wang, X.Tan, H.Tian, Q.Ao, J.Li, and S.Chu, “Machine learning classifiers for screening non-alcoholic fatty liver disease in general adults,” *Scientific Reports*, vol. 13, Art. no.3638, 2023, doi: 10.1038/s41598-023-30750-5.

[31] Y.-H.Chuang, J.-H.Su, D.-H. Han, Y.-W. Liao, Y.-C.Lee, and Y.-F.Cheng, “Effective natural language processing and interpretable machine learning for structuring CT liver-tumor reports,” *IEEE Access*, vol. 10, pp. 116273–116286, 2022, doi:10.1109/ACCESS.2022.3218646.

[32] K. Drożdż, K. Nabrdalik and H. Kwiendacz et al., “Risk factors for cardiovascular disease in patients with metabolic-associated fatty liver disease: A machine learning approach,” *CardiovascularDiabetology*, vol.21, Art.no.240,2022,doi:10.1186/s12933-022-01672-9.

[33] B.E.Kang, A.Park, H.Yang, Y.Jo,T. G.Oh, S.M.Jeong, Y.Ji,H.-L. Kim, H.-N. Kim, J. Auwerx, S. Nam, C.-Y. Park, and D. Ryu, “Machine learning-derived gut microbiome signature predicts fatty liver disease in the presence of insulin resistance,” *Scientific Reports*, vol. 12, Art. no. 21842, 2022, doi: 10.1038/s41598-022-26102-4.

[34] P. Rahi, I. Singh, S. Mahajan, R. Chauhan, T. Kaur, and J. Singh, “From genes to symptoms: Predicting Huntington’s disease progression with DNN,” in *Proc. Int. Conf. Electronics, AI and Computing (EAIC)*, Jalandhar, India, 2025, pp. 1–9, doi:10.1109/EAIC66483.2025.11101647.

[35] H. Castañé, G. Baiges-Gaya, A. Hernández-Aguilera et al, “Coupling machine learning and lipidomics as a tool to investigate metabolic dysfunction-associated fatty liver disease: A general overview,” *Biomolecules*, vol. 11, Art. no. 473, 2021, doi:10.3390/biom11030473.

[36] T. Marjot, G. J. Webb, and A. S. Barritt IV et al, “COVID-19 and liver disease: Mechanistic and clinical perspectives,” *Nature Reviews Gastroenterology & Hepatology*, vol. 18, pp. 348–364, 2021, doi:10.1038/s41575-021-00426-4.

[37] K.K.Bence and M.J.Birnbaum, “Metabolic drivers of non-alcoholic fatty liver disease,” *Molecular Metabolism*, vol. 50, Art. no. 101143,2021, doi: 10.1016/j.molmet.2020.101143.

[38] E. Hallaji, R. Razavi-Far, V. Palade, and M. Saif, “Adversarial learning on incomplete and imbalanced medical data for robust survival prediction of liver transplant patients,” *IEEE Access*, vol. 9, pp.73641–73650,2021,doi:10.1109/ACCESS.2021.3081040.

[39] A. Mahzari, “Artificial intelligence in nonalcoholic fatty liver disease,” *Egyptian Liver Journal*, vol. 12, Art. no. 69, 2022, doi:10.1186/s43066-022-00224-w.

[40] C. Balsano, P. Burra, C. Duvoux, A. Alisi, et al., “Artificial intelligence and liver: Opportunities and barriers,” *Digestive and Liver Disease*, vol. 55, no. 11, pp. 1455–1461, 2023, doi:10.1016/j.dld.2023.08.048.

[41] M.Salmi, D. Atif, and D.Oliva, et al., “Handling imbalanced medical datasets: Review of a decade of research,” *Artificial Intelligence Review*, vol. 57, Art. no. 273, 2024, doi: 10.1007/s10462-024-10884-2.

[42] K. Kantartzis and N. Stefan, “Clustering NAFLD: Phenotypes of non-alcoholic fatty liver disease and their differing trajectories,” *Hepatology Communications*, vol. 7, no. 4, Art. no. e0112, 2023, doi:10.1097/HC9.0000000000000112.

00112.

[43] P. Barik, “Liver diseases dataset,” Kaggle, 2025. [Online].Available:<https://www.kaggle.com/datasets/priyankabarik/liver-diseases>