International Journal of Public Health Science (IJPHS)
Vol.13, No.4, December 2024, pp. 2033~2043
ISSN: 2252-8806, DOI: 10.11591/ijphs.v13i4.23891

Comparative analysis of deep learning models for various nonalcoholic fatty liver disease datasets

Konakanchi Venkata Subrahmanya Srirama Murthy[1], Reddy Shiva Shankar[1], Samarendra Narayana Pradhan[1,2], Bhabodeepika Mohanty[1,2], Veeranki Venkata Rama Maheswara Rao[1,3]
[1]Department of Computer Science and Engineering, Sagi Ramakrishnam Raju Engineering College, Andhra Pradesh, India
[2]Dhaneswar Rath Institute of Engineering and Management Studies (DRIEMS) University, Odisha, India
[3]Shri Vishnu Engineering College for Women, Andhra Pradesh, India

Article Info

Article history:

Received Sep 7, 2023
Revised Mar 15, 2024
Accepted Apr 24, 2024

Keywords:

Artificial neural networks
Generalized feed forward
Modular neural network
Multi-layer perceptron
Nonalcoholic fatty liver disease

ABSTRACT

Fatty liver disease is caused by increased liver buildup or weight above 5-10%. This disorder is widespread in people with diabetes, overweight persons, and metabolic syndrome patients. Clinical decision support systems can improve liver failure diagnosis and prediction to reduce this situation. Many liver failure models have drawbacks, and liver failure prediction is still a problem. This work uses four large open-access critical care patient datasets to create and verify liver failure risk prediction models. This study aims to construct a clinically applicable diagnostic and predictive model that evaluates the probability or risk of liver failure in intensive care unit (ICU) patients using extreme gradient boosting (XGBoost), artificial neural networks (ANN), multi-layer perceptron (MLP), Modular Neural Network (MNN), and generalized feed forward (GFF). We evaluated performance metrics using these models: accuracy, sensitivity, specificity, and predictive accuracy.

Corresponding Author:

Reddy Shiva Shankar
Department of Computer Science and Engineering, Sagi Ramakrishnam Raju Engineering College
Bhimavaram, Andhra Pradesh, India
Email: shiva.csesrkr@gmail.com

1. INTRODUCTION

Nonalcoholic fatty liver disease (NAFLD) may develop several clinical and pathological symptoms similar to those of alcoholism. It is caused by fat buildup, primarily triglycerides in the hepatocytes, and individuals may develop superficial hepatic steatosis lesions, nonalcoholic steatohepatitis (NASH), and cirrhosis. Diabetes type 2 (DM-2) and obesity are the main risk factors for NAFLD, the hepatic manifestation of metabolic syndrome, which is rising alongside these two disorders. A multisystemic NAFLD is associated with metabolic diseases, including overweight and metabolic syndrome. The disorder may cause hepatic steatosis, NASH, severe fibrosis, cirrhosis, and hepatocellular carcinoma. The pathogenic process mirrors alcohol-induced liver damage in non-drinkers. NAFLD is also known as nonalcoholic steatohepatitis, diabetic hepatitis, nonalcoholic illness, and fatty liver.

It may be hepatic steatosis or NASH, marked by high necro-inflammatory activity. Steatohepatitis is merely one stage of NAFLD, and its primary clinical consequences are the tendency to proceed to cirrhosis and liver failure. Both illnesses have the potential to either regress or advance to hepatic cirrhosis, a condition that may lead to liver failure, portal hypertension, and hepatocellular cancer [1]. Although the prevalence of viral causes is decreasing, NAFLD remains the most common cause of chronic liver disease (CLD). This increasing trend is associated with changes in Western lifestyle, such as high rates of obesity and sedentary

behavior. NAFLD is a condition that affects the liver and is linked with metabolic syndrome (MS). It can occur in individuals who are slim and not diabetic, although it might have a genetic basis [2]. Insulin resistance (IR) is a medical condition that is linked to the development of multiple sclerosis (MS). IR also affects the liver, leading to the development of NAFLD. Due to this connection, it has been suggested that the term NAFLD be replaced with metabolic-associated fatty liver disease (MAFLD) [3]. Figure 1 shows NASH increasing liver fibrosis risk factors.

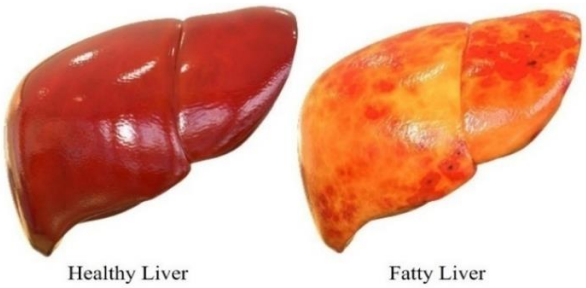

Figure 1. Representative image of healthy liver and NAFLD

Most individuals remain symptomatic for a long time and unaffected by their everyday lives, making it difficult to diagnose and manage persons who progressively proceed to NASH, NASH-cirrhosis, and hepatocellular cancer. Despite breakthroughs in pathogenic pathways and finding liver fibrosis as the best predictor of disease development, regulatory bodies have not authorized particular therapies. Lifestyle management for weight reduction is the only therapy outside controlled studies [4]. NAFLD is a condition that is diagnosed in individuals who do not have a history of alcohol consumption and are not affected by any specific conditions that cause fatty liver, such as viral hepatitis, drug-induced liver disease, total parenteral nutrition, hepatolenticular degeneration, and autoimmune liver disease. To confirm the diagnosis, liver imaging that fulfils the criteria for diffuse fatty liver is required. The patient in question has been officially diagnosed with fatty liver by two experts from the hepatology department of a lower-tier hospital in Urumqi after assessing the physical examination results [5]. This diagnosis is supported by the evidence of insulin resistance's involvement in hepatic steatosis, as seen in Figure 2.

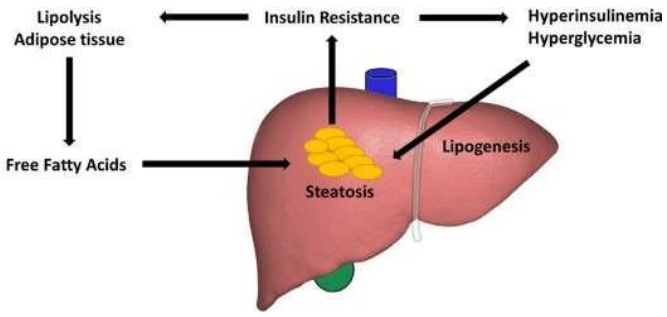

Figure 2. Role of insulin resistance in hepatic steatosis

2. RELATED WORK

NAFLD and metabolic syndrome (MetS) were suggested by Sookoian *et al.* [6]. An increasing amount of research means that genetic vulnerability and environmental exposure cause NAFLD. Typically co-occurs with MetS-associated traits, whether it causes or results from Reddy *et al.* [7] aimed to identify steatosis risk factors and measure their severity with fibrosis. The research comprised 195 fatty liver patients. The finding is based on hospital data assessing fatty liver risk factors. Steatosis was 15.8% at autopsy. Clinical Features and Natural History of Nonalcoholic Steatosis Syndromes by Falck-Ytter *et al.* [8] explained that liver illness may lead to cirrhosis and cause mortality. Current best estimates put NAFLD at 20% and NASH at 2-3% in the general population. NAFLD, steatohepatitis, and metabolic syndrome were postulated by Marchesini *et al.* [9]. NAFLD patients without diabetes were consecutively monitored.

The logistic regression analysis showed that metabolic syndrome increased NASH risk. Multiple metabolic abnormalities may cause severe liver damage. To diagnose nonalcoholic steatohepatitis, Saadeh *et al.* [10] presented the utility of radiological imaging in NAFLD. NAFLD was included, excluding those with additional liver disorders and heavy alcohol use. Each patient underwent a comprehensive abdominal examination, including ultrasonography, computed tomography (CT), and Magnetic resonance imaging (MRI). NASH patients showed higher ferritin, aspartate aminotransferase, hepatocyte ballooning, and fibrosis. According to the research, no radiological modality showed differences between NASH and nonprogressive NAFLD. These radiological modalities only showed steatosis severity.

Amarapurkaret *et al.* [11] reviewed the Asia–Pacific Working Party on NAFLD recommendations for assessing and managing NAFLD. This study shows that NAFLD is a severe Asia-Pacific public health issue. Asians with diabetes are less obese at a younger age than Caucasians but have more complications and untimely mortality. According to Reddy *et al.* [12], NAFLD from Steatosis to Cirrhosis is dangerous. For the illness rate in Americans, early intervention should enhance physical activity and apply nutritional and anti-obesity strategies. Ultrasonography may be utilized in patients with normal alanine transaminase (ALT) levels. After multiple investigations using American data, the researchers concluded that all patients should focus on decreasing belly girth rather than significantly lowering body weight, which requires lifelong maintenance of these favorable improvements. Gramlich *et al.* [13] describe the pathologic hallmarks of fibrosis in NAFLD. Two hematopathologists blindly examined liver biopsy specimens using a 19-item pathological procedure and a second methodology. Univariate and multivariate studies showed that 21.2% of 132 NAFLD patients had advanced fibrosis. Continuous hepatocyte damage is the most conspicuous pathological sign of hepatic fibrosis in NAFLD patients—liver disease phases. Zhu *et al.* [14] validated that NAFLD prevalence was 20.52%, demonstrating vital fatty liver index (FLI) predictive values for ultra-monographic diagnosis. Noninvasive indexes can help select potential people before imaging tests, reducing costs.

Zhu *et al.* [15] NAFLD prevalence and economy to determine whether adult NAFLD prevalence is related to national income. They analyzed PubMed to find suitable records published before September 2014. The global prevalence of NAFLD was 24.24%. The research revealed that males had a greater NAFLD prevalence than females, particularly in Europe. This research offers a unique epidemiologic viewpoint on NAFLD worldwide. Pan *et al.* [16] predicted NAFLD risk. Globally, it is a frequent liver condition with no good prediction or diagnostic tools. The cross-sectional research suggests a noninvasive NAFLD screening technique. There were 2,446 participants, 574 of whom had NAFLD. A risk prediction nomogram model was created using multivariable logistic regression analysis to detect NAFLD risk variables. To develop a model for predicting the risk of NAFLD, the researchers integrated various parameters such as demographic, clinical, and dietary factors identified in previous studies. They also included nutritional characteristics to establish a nomogram model to predict the risk of NAFLD. Zelber-Sagi *et al.* [17] suggested a population-based investigation on primary NAFLD prevalence and biochemical and anthropometric parameters. Its main clinical consequences are the rising frequency of NAFLD and the associated risk of cirrhosis and liver failure. Due to the difficulty of excluding excess alcohol intake, the epidemiological investigation of NAFLD is suited for the Israeli community, which has minimal alcohol consumption. NAFLD is common among Israelis and linked to metabolic syndrome. ALT significantly underestimates the frequency of NAFLD. The noise of the images was removed from the photos to identify the exact wound and get a correct diagnosis [18].

Karunasri *et al.* [19] proposed the epidemiology of NAFLD. Hepatocellular carcinoma (HCC) and NAFLD are the top two reasons for liver transplantation in the United States. NAFLD is caused by metabolic syndrome (MS), and it affects 25-30% of the population. The disease occurs in the liver and shares many risk factors with multiple sclerosis. Ismaiel *et al.* [20] presented NAFLD prediction using machine learning categorization and random forest analysis. AIP has been tested as a noninvasive NAFLD predictor in a few trials. Lallukka *et al.* [21] Jarvinen's idea links insulin resistance syndrome traits to NAFLD prediction. It includes dataset screening methods. Finally, ultrasonography-based longitudinal studies, especially in Asian populations, show that NAFLD predicts T2DM regardless of age or weight. Cho e*t al.* [22] used regression in machine learning with specific regression analysis implementations to portray. It gives many NAFLD diagnostic methods. It has various datasets. Gallstone disease and unexplained deep venous thrombosis may increase the risk of NAFLD.

3. METHOD

3.1. Objectives

This work aims to develop a clinically relevant diagnostic and predictive model that estimates the likelihood or risk of liver failure for a patient in an intensive care unit (ICU) using machine learning (ML) and deep learning (DL) models. The developed models are designed to output a finite 0-100 liver failure risk index (LFRI). The higher the value of the LFRI, the more likely the patient will experience liver failure.

3.2. Overview of liver functionality

It is vital to understand liver functionality to understand these realities of liver failure and the importance of identifying its failure well in advance. The liver is a solid organ and gland considered one of the most essential organs in the human body. It is comprised of two lobes. It is situated above the right of the stomach (right upper quadrant) and below the diaphragm. Detoxifying toxins and metabolizing medications in the liver is essential for overall health. Its primary functions include bile production, which helps the body absorb fats, proteins, carbohydrates, and vitamins. It also absorbs and metabolizes bilirubin, creates blood-clotting factors (coagulants), metabolizes fats, proteins, and carbohydrates, stores vitamins and minerals, filters blood, produces albumin, and removes aged and damaged red blood cells [23]. Detoxification of toxins and metabolism of medications are essential processes carried out by the liver to maintain overall body health. The hepatic artery supplies blood and oxygen from the heart and lungs to the liver, whereas veins provide blood-containing nutrients from the intestine [24].

3.3. Problem statement

In most cases, LT in liver failure patients could be avoided if the liver failure is detected in the early stages. However, detecting a failing liver is complicated in its early stages. Although liver function tests (LFTs) are commonly performed by healthcare providers, detecting liver disease at an early stage remains challenging. This is because abnormal LFTs may indicate other medical conditions besides liver-related issues. These conditions could include metastatic cancer, inflammatory or infectious diseases, and congestive heart failure. Hence, LFTs can be misleading and result in inappropriate treatments, leading to increased costs, morbidity, and death. Therefore, an accurate decision support system that can detect liver failure before its onset is necessary for the proper medication and medical treatment of patients.

3.4. Workflow

The four datasets were downloaded from the Kaggle repository. These datasets were discussed in the Discussion Section. The NAFLD detection process includes data pre-processing, model design, and model evaluation; the entire procedure is in Figure 3. Data pre-processing is the first and most crucial stage in machine learning. These took raw data and made it usable for the machine learning model. The suggested architecture allows two-phase data pre-processing. Null values are deleted during data cleaning. The picture dataset is downsized to train the model more effectively. The photos are scaled to 224x224. Two steps are described in the model design.

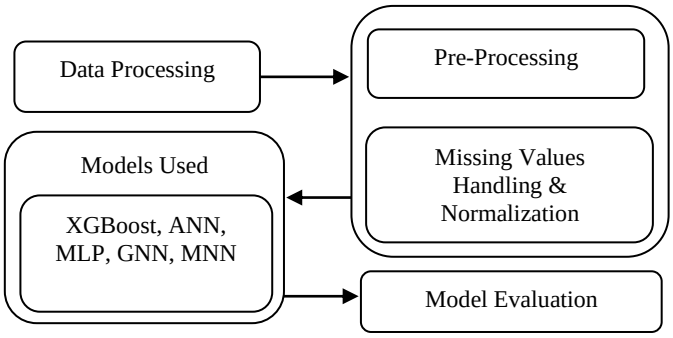

Figure 3. Model workflow

3.5. Models used

3.5.1. Extreme gradient boosting (XGBoost)

Open-source XGBoost functions and steps employ supervised machine learning to estimate or forecast results. Several XGBoost decision trees can anticipate the outcome. The machine learning system is trained using batch learning and generalized using a model-based method. Predictor-outcome models are created by using all relevant data. The test data is then generalized using these models. In computers, "extreme" means pushing processing power limits. Regression and classification employ "gradient boosting" to enhance poor prediction models. Good image classification performance has made XGBoost popular [25]. CNN extracts feature from the input, and XGBoost recognizes them for more accurate output, as shown in Figure 4.

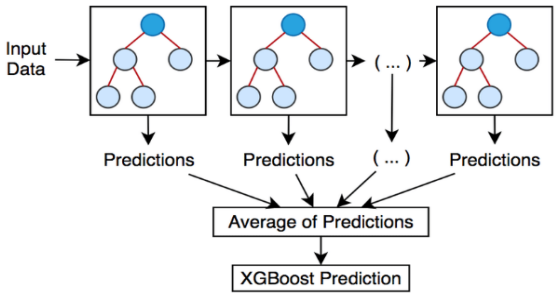

Figure 4. The process of XGBoost

3.5.2. Artificial neural networks (ANN)

ANNs are computer architectures that are modeled after brains. It is built by a series of "neurons" (or "nodes"), which are organized into layers [26]. These neurons exhibit global behavior determined by the established connections between the various processing elements and the related parameters in the neural network architecture. Neural networks consist of layers of interconnected neurons, with weights assigned to the connections between the i^{th} and thej^{th}neuronsin a layer. These weights determine the strength of the connection between the neurons in successive layers. The input layer receives data, which is then processed and translated through one or more hidden layers before being outputted by the output layer. The complexity of the network determines the number of hidden layers and neurons per layer. Typically, a two-hidden-layer ANN design is used. Each node mathematically processes incoming data in each layer before being passed on to the next layer. Figure 5 shows an example of a typical two-hidden-layer ANN design, while Figure 6 illustrates how the j^{th}node in a layer processes incoming data (x_i) from the previous layer. Steps involved:

i. Initially, a summation of weighted values is computed, followed by the addition of the bias term (θj) to this summation, as per (1).

$$Net_j = P_m\, x_{ij} * w_{ij} + \theta_j\ (j = 1,2, \ldots . n) \quad (1)$$

ii. A mathematical 'transfer function' transforms Netj. This function normalizes all network inputs and outputs to a range. It helps the neural network discover data patterns and trends better than 50–500 values. Many transfer functions work for this. For reasons explained under model development, a sigmoid function is employed in this attempt, as illustrated in (2).

$$f(x) = \frac{1}{(1+ e^{-x})} \quad (2)$$

iii. After creating a neural network for an application, it is necessary to use random starting weights for training. These models are popular because they learn from input data patterns. Network activity can be monitored or unsupervised. Supervised training uses the intended output and inputs to optimize network weights for the least output error. This type of training can be used for function approximation, regression analysis, time series prediction, and pattern and sequence recognition. On the other hand, unsupervised training is used to make sense of inputs and characterize "unlabeled" data. It is used for grouping and anomaly detection.

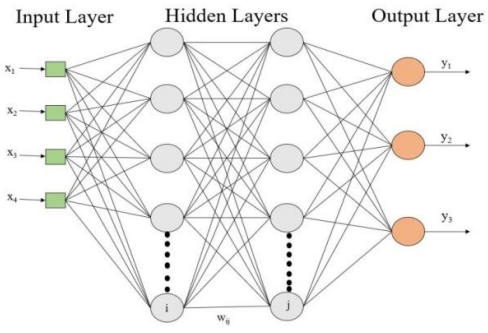

Figure 5. General structure of two hidden layer ANN

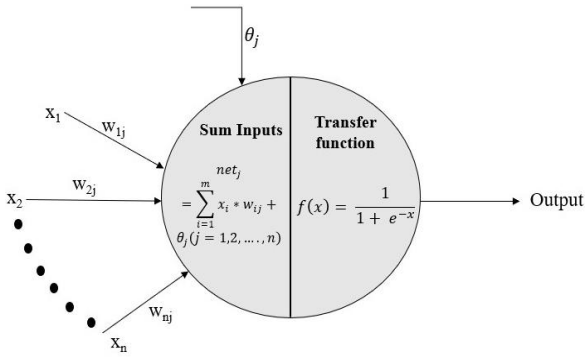

Figure 6. Data processing in a neuron

3.5.3. Multi-layer perceptron model (MLP)

Feedforward artificial neural networks with three or more layers are MLPs. Each hidden and output neuron has a non-linear activation function. Including multiple layers and non-linear activation, functions set MLP apart from standard linear Perceptron, enabling these networks to discuss data that isn't linearly separable. MLPs are universal function approximators and can be applied to develop mathematical models using regression analysis. These networks are well-suited for various modeling tasks, including pattern classification, prediction, and function approximation. Pattern classification is concerned with the data type divided into discrete classes. Prognosis is related to forecasting time series data when the current and previous trends are known. In contrast, function approximation involves modeling the relationship between the variables [27].

3.5.4. Generalized feed forward ANN model (GFF)

A GFF neural network is an ANN where the unit connections do not form a cycle, such as recurrent neural network models [28]. In this simple ANN, data moves from the input nodes to the output nodes via the hidden nodes. Unlike MLP, which is based on perceptron, this network uses a generalized shunting neuron (GSN) model as its computational unit. Shunting neurons can create complex, non-linear decision boundaries to help the GFF neural network design classify intricate patterns, predict time series, recognize patterns, and mine data dynamically.

3.5.5. Modular neural network model (MNN)

MNN is a particular class of MLP in which several parallel MLP sare used to process the inputs and then recombine the results, as shown in Figure 7. This process forms some structure within the topology, which helps develop a specialized function in each submodule. Divide and conquers corporate approach has many advantages to a neural network, such as scalability, robustness, and flexibility in design and implementation.

Moreover, these networks require fewer weights than an MLP to build a network of similar size because of partial interconnection between its layers. Hence, this reduces the necessary training exemplars and helps speed up the training times. However, this network can be segmented into modules in many ways, and it is unclear how to best design the modular topology based on the data [29]. Figure 7 shows a modular neural network architecture with' k' modules.

We developed and used XGBoost, ANN, MLP, GNN, and MNN model architectures in this research. We repeatedly modified model parameters for each model design. Typical parameters that changed iteratively include the number of hidden layers and their processing parts. Model development began with basic model construction. Neural network models and then repeatedly changed the parameters to enhance their complexity. We constructed models with two hidden layers, then three, and so on. We also tried increasing the number of processing components for each hidden layer from 5 to 50 to generate alternative models. Finally, the validation dataset was used to compare the performance of all these models to choose the best one. The first and second hidden layers with 10 and 2 processing components produced the best models in this investigation. Figure 8 explains the whole neural network model development process using different architectures and validation for analyzing its performance.

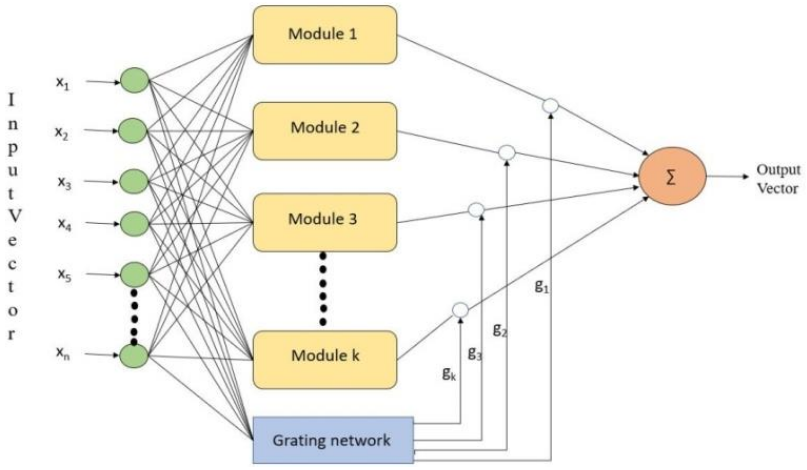

Figure 7. Modular neural network architecture

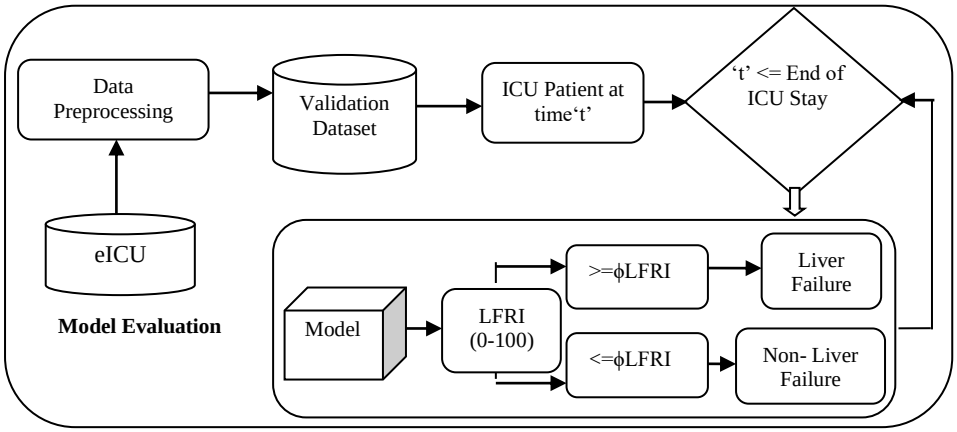

Figure 8. Process of model development and validation

4. RESULTS

This study validated the developed models against different performance metrics such as sensitivity, specificity, and predictive capacity. The description below shows a comparative analysis of various developed models regarding these performance metrics and LFRI plots obtained for the best model. As explained in the 'Model Development' section, we have created multiple models by iteratively changing different parameters of neural network architectures and validated these developed models with the ICU validation dataset. Table 1 shows the performance metrics–sensitivity (St) and specificity (Sp)-obtained for some of the developed models, which were configured with two, three, and four hidden layers and six, four, and three processing elements (PEs) in each of these hidden layers respectively. Table 2 includes the model performance metrics calculated for the best models developed by configuring the various neural network model architectures with two hidden layers and 10, 2 processing elements in the first and second hidden layers. It was evaluated for its predictive capacity.

Table 1. Performance of St and Sp metrics

Models	2 Hidden layers 4 PEs in each hidden layer		3 Hidden layers 4 PEs in each hidden layer		4 Hidden layers 3 PEs in each hidden layer	
	St	Sp	St	Sp	St	Sp
XGBoost	73.2%	68.5%	71.4%	64.3%	70.4%	61.7%
ANN	74.6%	69.9%	73.1%	67.4%	71.3%	63.6%
MLP	75.2%	71.5%	72.7%	65.3%	68.1%	64.2%
GFF	69.5%	70.8%	65.2%	66.7%	63.6%	66.3%
MNN	76.8%	35.6%	74.8%	33.2%	72.4%	29.7%

Table 2. Performance metrics for the best models

Models	2 Hidden layers 2 PEs in each hidden layer		Predictive capacity
	St	Sp	
XGBoost	81.2%	68.5%	81.8%(N=515)
ANN	84.6%	69.9%	85.5% (N=532)
MLP	83.3%	77.5%	83.5% (N=525)
GFF	76.1%	81.6%	80.1% (N=546)
MNN	91.8%	42.9%	90.3% (N=522)

4.1. Results obtained

Table 3 shows three evaluation metrics obtained for various methods as described. With the help of a confusion matrix, evaluation metrics have evolved: accuracy, sensitivity, and specificity values of different machine learning methods for NAFLD prediction from Lipids data. Table 4 represents the accuracy, sensitivity, and specificity values of other machine learning methods for NAFLD prediction from hormonal data. These metrics were evaluated by using the confusion matrix. Table 4 describes the three metrics with various models. Table 5 accuracy, sensitivity, and specificity values of different machine learning methods for NAFLD prediction from glycans data. Table 6 accuracy, sensitivity, and specificity values of other machine learning methods for NAFLD prediction from fatty acids data.

Table 3. Evaluation metrics for the lipids data

Models	Accuracy	Sensitivity	Specificity
XGBoost	0.80	0.84	0.91
ANN	0.67	0.79	0.77
MLP	0.87	0.90	0.92
GFF	0.70	0.80	0.84
MNN	0.89	0.93	0.94

Table 4. Evaluation metrics for the hormonal data

Models	Accuracy	Sensitivity	Specificity
XGBoost	0.53	0.69	0.77
ANN	0.49	0.69	0.66
MLP	0.55	0.68	0.81
GFF	0.53	0.66	0.76
MNN	0.56	0.67	0.84

Table 5. Evaluation metrics for the glycans data

Models	Accuracy	Sensitivity	Specificity
XGBoost	0.55	0.62	0.74
ANN	0.44	0.63	0.53
MLP	0.55	0.55	0.78
GFF	0.48	0.58	0.68
MNN	0.57	0.58	0.83

Table 6. Evaluation metrics for the fatty acids data

Models	Accuracy	Sensitivity	Specificity
XGBoost	0.50	0.60	0.74
ANN	0.43	0.60	0.59
MLP	0.54	0.58	0.78
GFF	0.44	0.52	0.72
MNN	0.59	0.66	0.81

Comparison of different datasets for the MNN model

Table 7 shows the MNN values obtained from lipids, hormones, glycans, and fatty acids datasets. Among XGBoost, ANN, MLP, GFF, and MNN, the MNN had the highest accuracy in all the various lipids, hormones, glycans, and fatty acids. Using Table 7 values, we plotted Figures 9 and 10 as a graph for accuracy, sensitivity, and specificity with various datasets for the MNN Model.

Table 7. Comparison of Datasets with MNN

MNN	Lipids	Hormones	Glycans	Fatty acids
Accuracy	0.89	0.56	0.57	0.59
Sensitivity	0.93	0.67	0.58	0.66
Specificity	0.94	0.84	0.83	0.81

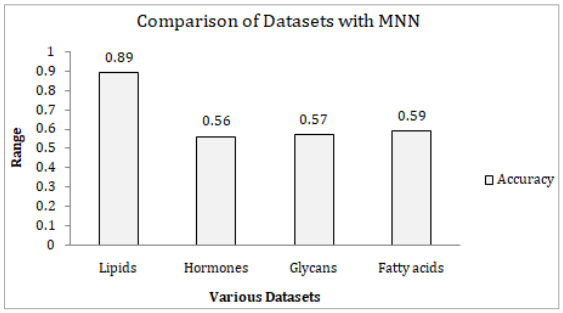

Figure 9. Graph for accuracy by comparing with various datasets for the MNN model

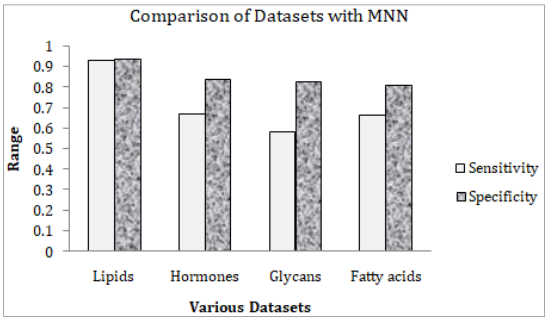

Figure 10. Graph for Sp and St comparison of different datasets for the MNN model

5. DISCUSSION

It is essential to detect liver failure and dysfunction early to provide timely medical intervention and improve patient outcomes. In our research, we have developed a neural network modeling technique to predict the likelihood of liver failure in ICU patients before it occurs. This approach is more effective than previous methods since it accurately evaluates the impact of various inputs or independent factors on the output or dependent variables [15], [16].

Some previous approaches tried to use organ failure scoring systems to evaluate the illness severity among liver failure patients. As the scoring systems are tended for broader organ systems, the high values of these scores indicate that the patient is at critical (without a section of the various dimensions of the scoring system that contributes to the final composite value). Therefore, these prior approaches potentially fail to identify the patients at the highest risk for a specific acute condition with high sensitivity and specificity [20]. Low-sensitivity diagnostic tools often fail to determine the correct outcome for any patient, leading to uncertainty in benefiting from early treatment. We have overcome this drawback by generating a

predictive model for identifying liver failure patients. We achieved promising results of 83.3% sensitivity at a specificity of 77.5% and correctly identified 83.3% (N=629) of patients with liver failure in the ICU validation set. The LFRI successfully predicted the onset of liver failure in 83.5% (N=525) of the 629 patients, with a median of 17.5 hours before its start [21], [24], [27].

Recently, various ML techniques have been used to develop prediction models for the early detection of liver failure. However, most of these models were not designed or customized for the intensive care unit (ICU) patient population [30]. The ICU is a crucial department in a hospital, with each unit having a unique atmosphere that reflects the specific surgical techniques medical professionals employ. ICU teams comprise highly skilled intensive care doctors, specialists, and nurses skilled in providing care to critically ill patients using specialized, technical, and monitoring equipment. Unlike general patients, daily monitoring of ICU patients is necessary because optimizing patient statuses, including but not limited to hemodynamics, ventilation, and nutrition, is critical to improving the survival of patients [31]. So, the surveillance and monitoring of ICU patients are essential, and the inability to detect or predict liver failure in these patients may lead to catastrophic consequences. As explained in the 'Model Training and Validation Set Generation' section, we have addressed this challenge by developing models from abroad ICU patient populations with and without liver failure. Systematic reviews evaluating this kind of approach all conclude that such studies have the characteristics of deficiencies in study design, adequate statistical methodology, and poor reporting [32].

Moreover, this validation makes the developed models only work effectively for a particular healthcare institution or a small subset of patients, impacting overall clinical utility. It is necessary to see how well a model performs with patients from a different but "plausibly related" population. Therefore, impact studies should not be considered until the robustness and generalizability of the developed model are verified with one or more external validation databases [33].

Further, all the previous approaches developed and validated the predictive models using datasets that included an almost equal number of liver and non-liver failure patients. Testing a developed predictive model with a validation dataset with an equal ratio of with and without liver failure conditions would artificially inflate the sensitivity and specificity values. For example, a predictive model can quickly achieve 50% sensitivity for such validation sets by putting '1' to the entire dataset without considering any input values/predictors [34]. However, in the real world, a very poor prevalence of liver failure patients, between 1.0 and 5.0%, can be seen in the ICU patient population. In this study, there were only 755 patients diagnosed with liver failure in the ICU out of 81,135 patient admissions. The ratio (approx .0.01) of patients with liver failure to those that do not have the condition is thus much less than the near 0.5 ratios implemented in most prior approaches. Hence, our study aims to address the previous limitation by validating the LFRI models with a representative real-world ICU patient dataset as described.

6. CONCLUSION

In conclusion, neural network models developed in this effort have been demonstrated to predict the likelihood of liver failure for a patient in ICU many hours before standard screening protocols. Data is accomplished by considering several data sources from the patient's electronic medical record (EMR), including but not limited to laboratory results and vital signs. The performance of this model has also been validated externally using data from critical care patients from a completely different database and achieved a high sensitivity of 83.3% at a specificity of 77.5%. Moreover, this model has identified 83.5% (N=525) of liver failure patients with a median of 17.5 hours before the onset of liver failure. Achieving such a high performance when validated with an external database of patient records substantiates that our approach has built a promising generalized model for predicting liver failure in the ICU population. Coordinating evidence-based remedies and performance enhancement measures with such models can significantly improve ICU patient outcomes and help learn healthcare systems.

REFERENCES

[1] D. E. Kleiner and H. R. Makhlouf, "Histology of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis in adults and children," *Clinics in Liver Disease*, vol. 20, no. 2, pp. 293–312, May 2016, doi: 10.1016/j.cld.2015.10.011.

[2] M. Eslam *et al.*, "MAFLD: A consensus-driven proposed nomenclature for metabolic associated fatty liver disease," *Gastroenterology*, vol. 158, no. 7, pp. 1999–2014, May 2020, doi: 10.1053/j.gastro.2019.11.312.

[3] A. Lonardo, F. Nascimbeni, M. Maurantonio, A. Marrazzo, L. Rinaldi, and L. E. Adinolfi, "Nonalcoholic fatty liver disease: Evolving paradigms," *World Journal of Gastroenterology*, vol. 23, no. 36, pp. 6571–6592, Sep. 2017, doi: 10.3748/wjg.v23.i36.6571.

[4] S. Reddy, N. Sethi, R. Rajender, and G. Mahesh, "Forecasting diabetes correlated non-alcoholic fatty liver disease by exploiting naïve bayes tree," *ICST Transactions on Scalable Information Systems*, p. 173975, Jul. 2018, doi: 10.4108/eai.29-4-2022.173975.

[5] C. Acierno, A. Caturano, P. C. Pafundi, R. Nevola, L. E. Adinolfi, and F. C. Sasso, "Nonalcoholic fatty liver disease and type 2 diabetes: pathophysiological mechanisms shared between the two faces of the same coin," *Exploration of Medicine*, vol. 1, no. 5, pp. 287–306, 2020, doi: 10.37349/emed.2020.00019.

[6] S. Sookoian and C. J. Pirola, "Nonalcoholic fatty liver disease and metabolic syndrome: Shared genetic basis of pathogenesis," *Hepatology*, vol. 64, no. 5, pp. 1417–1420, Oct. 2016, doi: 10.1002/hep.28746.

[7] S. S. Reddy, N. Sethi, and R. Rajender, "Safe prediction of diabetes mellitus using weighted conglomeration of mining schemes," in *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Nov. 2020, pp. 1213–1220. doi: 10.1109/ICECA49313.2020.9297390.

[8] Y. Falck-Ytter, Z. M. Younossi, G. Marchesini, and A. J. McCullough, "Clinical Features and natural history of nonalcoholic steatosis syndromes," *Seminars in Liver Disease*, vol. 21, no. 01, pp. 017–026, 2001, doi: 10.1055/s-2001-12926.

[9] G. Marchesini, "Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome," *Hepatology*, vol. 37, no. 4, pp. 917–923, Apr. 2003, doi: 10.1053/jhep.2003.50161.

[10] S. Saadeh *et al*., "The utility of radiological imaging in nonalcoholic fatty liver disease," *Gastroenterology*, vol. 123, no. 3, pp. 745–750, Sep. 2002, doi: 10.1053/gast.2002.35354.

[11] D. N. Amarapurkar, E. Hashimoto, L. A. Lesmana, J. D. Sollano, P. Chen, and K. Goh, "How common is non-alcoholic fatty liver disease in the Asia–Pacific region and are there local differences?," *Journal of Gastroenterology and Hepatology*, vol. 22, no. 6, pp. 788–793, Jun. 2007, doi: 10.1111/j.1440-1746.2007.05042.x.

[12] S. Reddy, G. Mahesh, and N. Preethi, "Exploiting machine learning algorithms to diagnose foot ulcers in diabetic patients," *EAI Endorsed Transactions on Pervasive Health and Technology*, p. 170752, Jul. 2018, doi: 10.4108/eai.24-8-2021.170752.

[13] T. Gramlich, D. E. Kleiner, A. J. McCullough, C. A. Matteoni, N. Boparai, and Z. M. Younossi, "Pathologic features associated with fibrosis in nonalcoholic fatty liver disease," *Human Pathology*, vol. 35, no. 2, pp. 196–199, Feb. 2004, doi: 10.1016/j.humpath.2003.09.018.

[14] J. Zhu *et al.*, "Validation of simple indexes for nonalcoholic fatty liver disease in western China: a retrospective cross-sectional study," *Endocrine Journal*, vol. 65, no. 3, pp. 373–381, 2018, doi: 10.1507/endocrj.EJ17-0466.

[15] J.-Z. Zhu, Y.-N. Dai, Y.-M. Wang, Q.-Y. Zhou, C.-H. Yu, and Y.-M. Li, "Prevalence of nonalcoholic fatty liver disease and economy," *Digestive Diseases and Sciences*, vol. 60, no. 11, pp. 3194–3202, Nov. 2015, doi: 10.1007/s10620-015-3728-3.

[16] X. Pan et al., "Risk prediction for non-alcoholic fatty liver disease based on biochemical and dietary variables in a Chinese Han population," *Frontiers in Public Health*, vol. 8, Jul. 2020, doi: 10.3389/fpubh.2020.00220.

[17] S. Zelber-Sagi, D. Nitzan-Kaluski, Z. Halpern, and R. Oren, "Prevalence of primary non-alcoholic fatty liver disease in a population-based study and its association with biochemical and anthropometric measures," *Liver International*, vol. 26, no. 7, pp. 856–863, 2006, doi: 10.1111/j.1478-3231.2006.01311.x.

[18] V. M. Gupta, K. Murthy, and R. S. Shankar, "A novel approach for image denoising and performance analysis using SGO and APSO," *Journal of Physics: Conference Series*, vol. 2070, no. 1, 2021, doi: 10.1088/1742-6596/2070/1/012139.

[19] K. Karunasri, G. Mahesh, and R. S. Shankar, "Medical Images Security in Cloud Computing Using Cp-Abe Algorithm," *ARPN Journal of Engineering and Applied Sciences*, vol. 17, no. 7, pp. 759–766, 2022..

[20] A. Ismaiel *et al.*, "Atherogenic index of plasma in non-alcoholic fatty liver disease: systematic review and meta-analysis," *Biomedicines*, vol. 10, no. 9, p. 2101, Aug. 2022, doi: 10.3390/biomedicines10092101.

[21] S. Lallukka and H. Yki-Järvinen, "Non-alcoholic fatty liver disease and risk of type 2 diabetes," *Best Practice & Research Clinical Endocrinology & Metabolism*, vol. 30, no. 3, pp. 385–395, Jun. 2016, doi: 10.1016/j.beem.2016.06.006.

[22] Y. Cho *et al*., "Skeletal muscle mass to visceral fat area ratio as a predictor of NAFLD in lean and overweight men and women with effect modification by sex," *Hepatology Communications*, vol. 6, no. 9, pp. 2238–2252, Sep. 2022, doi: 10.1002/hep4.1975.

[23] A. E. Rigamonti, A. Bondesan, E. Rondinelli, S. G. Cella, and A. Sartorio, "The Role of aspartate transaminase to platelet ratio index (apri) for the prediction of non-alcoholic fatty liver disease (NAFLD) in severely obese children and adolescents," *Metabolites*, vol. 12, no. 2, p. 155, Feb. 2022, doi: 10.3390/metabo12020155.

[24] E. X.-X. Tan *et al*., "Non-obese non-alcoholic fatty liver disease (NAFLD) in Asia: an international registry study," Metabolism, vol. 126, p. 154911, Jan. 2022, doi: 10.1016/j.metabol.2021.154911.

[25] A. Kotronen *et al.*, "Prediction of non-alcoholic fatty liver disease and liver fat using metabolic and genetic factors," *Gastroenterology*, vol. 137, no. 3, pp. 865–872, 2009, doi: 10.1053/j.gastro.2009.06.005.

[26] I. C. Efrem *et al.*, "A Study of Biomarkers Associated with Metabolic Dysfunction-Associated Fatty Liver Disease in Patients with Type 2 Diabetes," *Diagnostics*, vol. 12, no. 10, p. 2426, Oct. 2022, doi: 10.3390/diagnostics12102426.

[27] A. L. Fracanzani *et al.*, "Risk of nonalcoholic steatohepatitis and fibrosis in patients with nonalcoholic fatty liver disease and low visceral adiposity," *Journal of Hepatology*, vol. 54, no. 6, pp. 1244–1249, Jun. 2011, doi: 10.1016/j.jhep.2010.09.037

[28] V. M. Kumar, R. R. Gundepalli, G. M. Chavan, and G. Maddela, "Study of prevalence of NAFLD and its associated risk factors in non-obese and non-diabetic rural populatio," *European Journal of Molecular & Clinical Medicine*, vol. 9, no. 5, pp. 555–564, 2022.

[29] A. Nogami *et al.*, "Non-invasive imaging biomarkers for liver steatosis in non-alcoholic fatty liver disease: present and future," *Clinical and Molecular Hepatology*, vol. 29, no. Suppl, pp. S123–S135, Feb. 2023, doi: 10.3350/cmh.2022.0357.

[30] S. Reddy and G. Mahesh, "Risk assessment of type 2 diabetes mellitus prediction using an improved combination of NELM-PSO," *ICST Transactions on Scalable Information Systems*, p. 169579, Jul. 2018, doi: 10.4108/eai.3-5-2021.169579.

[31] D. Zhao, C. Huang, Y. Wei, F. Yu, M. Wang, and H. Chen, "An effective computational model for bankruptcy prediction using kernel extreme learning machine approach," *Computational Economics*, vol. 49, no. 2, pp. 325–341, Feb. 2017, doi: 10.1007/s10614-016-9562-7.

[32] P. Chen and C. Pan, "Diabetes classification model based on boosting algorithms," *BMC Bioinformatics*, vol. 19, no. 1, p. 109, Dec. 2018, doi: 10.1186/s12859-018-2090-9.

[33] S. Bashir, U. Qamar, F. H. Khan, and M. Y. Javed, "An efficient rule-based classification of diabetes using ID3, C4.5, & CART Ensembles," in *2014 12th International Conference on Frontiers of Information Technology*, Dec. 2014, pp. 226–231. doi: 10.1109/FIT.2014.50.

[34] A. Kulkarni, S. Shinde, and D. Kadam, "Automated prediction of non alcoholic fatty liver disease using machine learning algorithms," *International Research Journal of Engineering and Technology*, vol. 7, no. 20, pp. 488–491, 2020.

BIOGRAPHIES OF AUTHORS

Konakanchi Venkata Subrahmanya Srirama Murthy is an Assistant Professor in the Department of Computer Science and Engineering in Sagi Rama Krishnam Raju Engineering College, Bhimavaram, Andhra Pradesh, India. He Received B.Tech and M.Tech Degrees from Andhra University, Visakhapatnam. He is pursuing Ph.D in Computer Science and Engineering at KLU. His Research areas are Machine Learning, Edge Computing, Natural Language Processing and Image Processing. He can be contacted at email: kvssrmurthy75@gmail.com.

Reddy Shiva Shankar is an Assistant Professor at the Department of Computer Science and Engineering in Sagi Ramakrishnam Raju Engineering College, Bhimavaram, Andhrapradesh, India. He has completed his PhD degree in Computer Science and Engineering with a specialization in Medical Mining and Machine Learning from BPUT, Odisha. His research areas are Image Processing, Medical Mining, Machine Learning, Deep Learning and Pattern Recognition. He published more than 70+ papers in International Journals and Conferences. S.S. Reddy has filed 05 patents. His research interests include Image Processing, Medical Mining, Machine Learning, Deep Learning and Pattern Recognition. He can be contacted at email: shiva.csesrkr@gmail.com.

Samarendra Narayan Pradhan is a Data scientist, having more than 15+ yrs. of experience in wireless network planning, optimization, Maximizing ROI: Budgeting and CAPEX Optimization in Wireless Networks, Customer Experience, Sales and Marketing Analytics in the Telecom Industry, Leveraging Deep Learning for Customer Churn Prediction, Natural Language Processing Models for Customer Feedback Analysis. He is a MTech postgraduate in Computer Science and Engineering from Dhaneswar Rath Institute of Engineering and Management Studies (DRIEMS). He can be contacted at email: luckysamar@yahoo.com.

Bhabodeepika Mohanty is a part time guest faculty with Bhubanananda Orissa School of Engineering in Cuttack. Before that She was a M.Tech postgraduate in Computer Science and Engineering from Dhaneswar Rath Institute of Engineering and Management Studies (DRIEMS). She can be contacted with email: mohantybhabodeepika@gmail.com.

Veeranki Venkata Rama Maheswara Rao is a leading Researcher & Academician in Computer Science & Engineering and holds Ph.D. degree. He is currently working as a Professor in the Dept. of Computer Science & Engineering at Shri Vishnu Engineering College for Women (A), Andhra Pradesh, India. He is actively involved and successfully implemented three projects funded by DST. He has 45 research papers, 17 of which are Scopus-indexed and 7 of which are Web of Science-indexed. He has 23 years of experience that include six years of Industry experience, 19 years of Teaching experience and 15 years of Research experience. His Research interests include Data Mining, Web Mining, Cloud Computing, Big Data Analytics, Data Science, Artificial Intelligence and Machine Learning. He can be contacted at email: mahesh_vvr@yahoo.com.