International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

LIVER DISEASE PREDICTION USING DEEP LEARNING

School of Engineering, Malla Reddy University, Secunderabad -500010, Telangana, India.

Veena Rani
Asst. Professor, Malla Reddy University
Hyderabad
Veena@mallareddyuniversity.ac.in

A.Sri Charan
Student, Malla Reddy University
Hyderabad
2011cs020019@mallareddyuniversity.ac.in

A.Surya Pavan
Student, Malla Reddy University
Hyderabad
2011cs020020@mallareddyuniversity.ac.in

A.V.S Pavan
Student, Malla Reddy University
Hyderabad.
2011cs020021@mallareddyuniversity.ac.in

A.Vijay
Student, Malla Reddy University
Hyderabad.
2011cs020022@mallareddyuniversity.ac.in

ABSTRCT:

Liver disease is a significant global health concern, and timely and accurate prediction of liver disease can greatly impact patient outcomes. In recent years, deep learning techniques have shown promise in various medical applications. This study aims to develop a deep learning-based predictive model for liver disease diagnosis. The proposed model utilizes a large dataset of patient information, including demographic data, clinical history, laboratory test results, and medical imaging studies. The data is preprocessed to remove noise, handle missing values, and normalize the features. A deep learning architecture, such as a convolutional neural network (CNN) or recurrent neural network (RNN), is designed and trained using the dataset. During the training phase, the model learns from the data to identify relevant patterns and relationships between the input features and the presence or severity of liver disease. The model's parameters are optimized through backpropagation to minimize prediction error or maximize a defined performance metric.After training, the model is evaluated on a

separate test dataset to assess its performance in predicting liver disease. Performance metrics such as accuracy, precision, recall, and AUC-ROC are computed to evaluate the model's effectiveness. Liver disease prediction using deep learning has the potential to assist healthcare professionals in making timely diagnoses, facilitating early intervention and treatment planning. Furthermore, these models may uncover novel risk factors or biomarkers associated with liver disease, leading to advancements in diagnostic and therapeutic approaches.

However, challenges exist in obtaining large, high-quality labeled datasets and interpreting the black-box nature of deep learning models. Future research should focus on addressing these limitations and further refining deep learning techniques for liver disease prediction.

Overall, this study demonstrates the potential of deep learning in liver disease prediction, with the hope of improving diagnostic accuracy and ultimately enhancing patient outcomes..

I-INTRODUCTION:

Liver disease is a serious health concern worldwide, affecting millions of people. Early detection and accurate diagnosis of liver disease are crucial for effective treatment and management. Deep learning, a subset of artificial intelligence, has gained significant attention in recent years for its ability to analyze complex patterns and make predictions based on large datasets. In this context, deep learning algorithms can be employed to develop predictive models for liver disease. These models leverage the power of neural networks to automatically learn and extract relevant features from medical data, such as patient demographics, clinical history, laboratory test results, and imaging studies.

The process of liver disease prediction using deep learning typically involves several steps. Firstly, a large dataset consisting of relevant patient information and corresponding liver disease outcomes is collected. This dataset may be obtained from electronic health records, medical databases, or clinical studies. Next, the data is preprocessed to ensure consistency and remove any noise or irrelevant information. This step involves cleaning the data, handling missing values, and normalizing or standardizing the features to ensure that the deep learning model can effectively learn from it. Once the data is prepared, a deep learning architecture is designed and trained using the dataset. Deep learning models for liver disease prediction often involve various layers of artificial neural networks, such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs). These models learn from the data to identify meaningful patterns and relationships between the input features and the presence or severity of liver disease. During the training process, the model's parameters are optimized to minimize the prediction error or

maximize a defined performance metric, such as accuracy, sensitivity, or specificity. This optimization is typically achieved through a process called backpropagation, where the model adjusts its internal parameters based on the calculated error gradients. After training, the model is evaluated on a separate test dataset to assess its performance. This evaluation provides insights into the model's ability to accurately predict liver disease in new, unseen cases. Performance metrics, such as accuracy, precision, recall, and area under the receiver operating characteristic curve (AUC-ROC), are commonly used to evaluate the model's effectiveness. Liver disease prediction using deep learning offers several potential benefits. It can assist healthcare professionals in making timely and accurate diagnoses, facilitating early intervention and treatment planning. Furthermore, deep learning models can potentially identify novel risk factors or biomarkers associated with liver disease, aiding in the development of new diagnostic and therapeutic approaches.

However, it is important to note that deep learning models are not without limitations. They require large amounts of high-quality labeled data to achieve optimal performance, which can be challenging to obtain in some cases. Additionally, the black-box nature of deep learning models may make it difficult to interpret the underlying reasoning for predictions, potentially hindering their acceptance and adoption in clinical practice. In conclusion, liver disease prediction using deep learning holds great promise in improving diagnostic accuracy and patient outcomes. With further advancements in deep learning techniques and the availability of comprehensive datasets, these models have the potential to revolutionize liver disease diagnosis, prognosis, and treatment.

II-LITERATURE SURVEY

Liver Disease Prediction by using different Decision Tree techniques Authors: Nazmun Nahar and Ferdous Ara: Early prediction of sickness is incredibly necessary to save lots of human life and take correct steps to regulate the disease call Tree algorithms are with success applied in varied fields, particularly in life science. This analysis work explores the first prediction of disease exploitation varied call tree techniques. The disease dataset that is chosen for this study is consisting of attributes like total animal pigment, direct animal pigment, age, gender, total proteins, simple protein and simple protein magnitude relation. The most purpose of this work is to calculate the performance of varied call tree techniques and compare their performance. The choice tree techniques employed in this study area unit J48, LMT, Random Forest, Random tree, REP Tree, call Stump, and Hoeffding

Tree. The analysis proves that call Stump provides the very best accuracy than different techniques. Decision Tree (DT): Decision Tree calculation has a place with the supervised learning algorithms In contrast to other supervised learning algorithms, a decision tree algorithm can be utilized for taking care of regression and classification issues as well. The general thought process of utilizing Decision Tree is to make a training model that can use to predict class or estimation of objective factors by taking in choice standards derived from earlier data (training data). Liver Disease Prediction using SVM and Naïve Bayes Algorithms Authors: Dr. S. Vijayarani1 , Mr.S.Dhayanand: Data mining has become an easy use for disease prediction in recent years in health care sectors. Data mining is the dredging process of information from massive datasets or warehouses or other repositories. Predicting the diseases from the voluminous medical databases is a very challenging task for researchers. The researchers use data mining techniques such as classification, clustering, rules of association and so on to over come this issue. The main objective of this research work is to use classification algorithms to predict liver diseases. Naïve Bayes and support vector machine (SVM) are the algorithms used in this work. The algorithms are classify the factors on performance. These classifier algorithms are compared on the basis of performance factors i.e. accuracy of classification and time of execution. It is observed from the experimental results that the `SVM is a better classifier for liver disease prediction.

III-PROPOSED METHODOLOGY

Liver is well-thought-out to be one of the central organs in any living body with fundamental functions such as processing leftover products, generating enzymes, and eliminating exhausted tissues or cells. We can stay alive merely a couple of days if our liver shuts down. Fortunately, the liver can continue its role even when up to 75% of it is contaminated or removed. Is due to its astonishing capability to produce new liver tissues from fine fettle liver cells that quiet exist .

Existing System:

• The liver disease prediction is performed by various machine learning techniques.

• Algorithms like,

• **support vector machine:** SVM algorithm tries to give out hyper planes and split the data into different categories. The scikit-learn package in python is employed for implementing SVM. The pre-processed information is split into check information and coaching set that is of twenty fifth and seventy fifth of the entire dataset severally. A

SVM technique builds hyper planes in an exceedingly dimensional area. a decent separation is achieved by the hyper plane that has the most important distance to the closest coaching information of any category (so-called purposeful margin), since generally the larger the margin the lower the generalization error of the classifier.

• **decision trees:** Decision Tree calculation has a place with the supervised learning algorithms In contrast to other supervised learning algorithms, a decision tree algorithm can be utilized for taking care of regression and classification issues as well. The general thought process of utilizing Decision Tree is to make a training model that can use to predict class or estimation of objective factors by taking in choice standards derived from earlier data (training data).

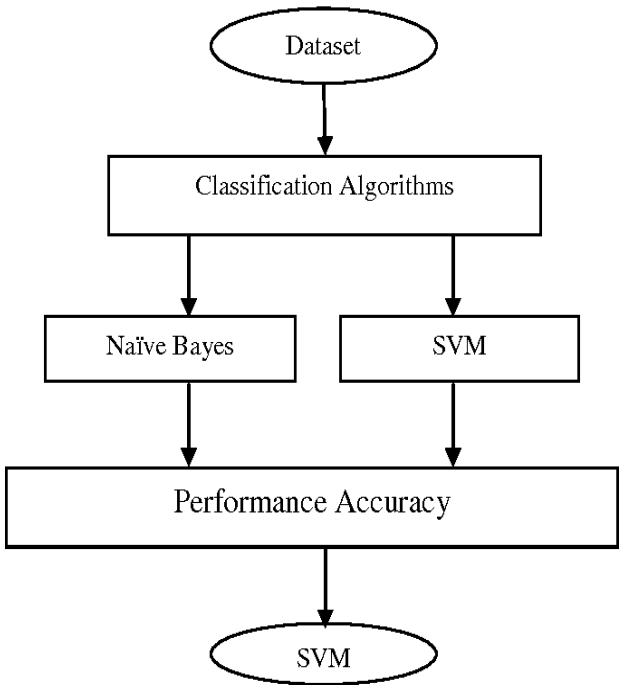
• **naïve bayes** : Naïve Bayes is one of the basic probabilistic classifiers which classifies the specific class with the given tuple. It is categorised by hypothesising that every attribute has a solitary effect on the class attribute by not depending on other attribute values. These algorithms are performed to predict liver disease prediction in existing systems.

• But due to the large amount of data obtaining accuracy was difficult. The accuracies provided by the existing system algorithms are shown below. 8

Algorithms	Correctly Classified Instances (%)	Incorrectly Classified Instances (%)	TP Rate	Precision	F Measure
Naïve Bayes	61.28	38.72	0.612	0.558	0.251
SVM	79.66	20.34	0.796	0.766	0.331

The flow chart of existing system are:

III. METHODOLOGY



Proposed System:

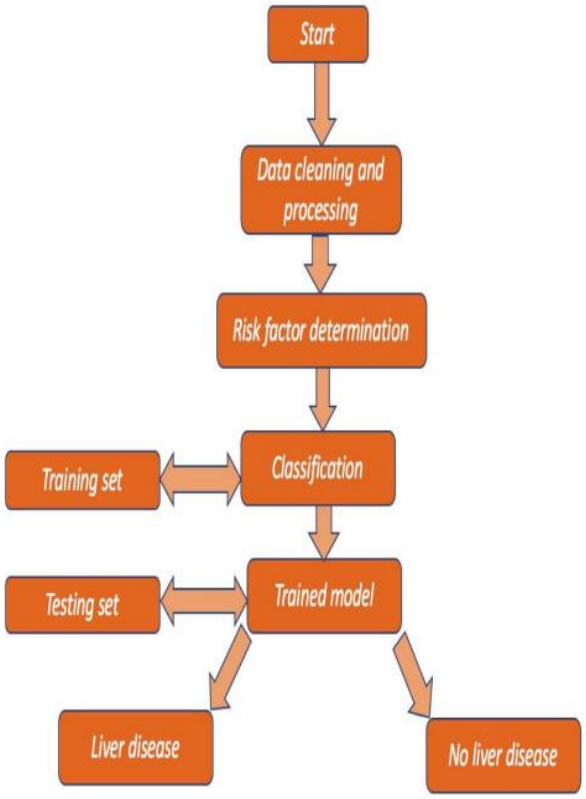
• For obtaining high accuracy we used deep learning techniques for liver disease prediction.

• Deep learning techniques like multilayer perceptron, restricted Boltzmann and gradient descent are performed on the dataset.

• Among the algorithms multilayer perceptron algorithm gives highest accuracy for liver disease prediction.

MULTILAYER PERCEPTRON: The Multilayer Perceptron was developed to tackle this limitation. It is a neural network where the mapping between inputs and output is non-linear. A Multilayer Perceptron has input and output layers, and one or more hidden layers with many neurons stacked together. And while in the Perceptron the neuron must have an activation function that imposes a threshold, like ReLU or sigmoid, neurons in a Multilayer Perceptron can use any arbitrary activation function.

IV-ARCHITECTURE :



V-RESULTS

Indian_liver_patient.csv dataset obtained from Kaggle. The dataset contains following attributes ['Age' , 'Gender' , 'Total_Bilirubin' , 'Direct_Bilirubin' , ‘Alkaline_Phosphotase’,’ Alamine_Aminotransferase’ , ‘Aspartate_Aminotransferase’ , 'Total_Protiens' , 'Albumin' , 'Albumin_and_Globulin_Ratio ', ’dataset’] The dataset is imported the head of the data is printed and found the shape of the dataset using Jupiter notebook platform which is shown in below fig. The sample of data is printed to know

about the data. The data almost contains all the numerical data except gender attribute. The fig 1.1 shows the output of sample data and fig 1.2 shows the output of shape of the dataset

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	\
0	65	Female	0.7	0.1	187	
1	62	Male	10.9	5.5	699	
2	62	Male	7.3	4.1	490	
3	58	Male	1.0	0.4	182	
4	72	Male	3.9	2.0	195	

	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens	\
0	16	18	6.8	
1	64	100	7.5	
2	60	68	7.0	
3	14	20	6.8	
4	27	59	7.3	

	Albumin	Albumin_and_Globulin_Ratio	Dataset
0	3.3	0.90	1
1	3.2	0.74	1
2	3.3	0.89	1
3	3.4	1.00	1
4	2.4	0.40	1

```
the shape of the dataset is:
Out[4]: (583, 11)
```

The pre-processing techniques are applied on the dataset especially the categorical data is converted into numerical data. Here, used dataset contains the numerical data in gender attribute and preprocessing techniques are applied to change it from categorical data to numerical data. Some of the attributes in the dataset is assigned to one variable and some attributesare assigned to another variable and using this variable the splitting, training and testing of the data set will be done and we can see in the below fig

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	\
0	65	0	0.7	0.1	187	
1	62	1	10.9	5.5	699	
2	62	1	7.3	4.1	490	
3	58	1	1.0	0.4	182	
4	72	1	3.9	2.0	195	

	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens	\
0	16	18	6.8	
1	64	100	7.5	
2	60	68	7.0	
3	14	20	6.8	
4	27	59	7.3	

	Albumin	Albumin_and_Globulin_Ratio	Dataset
0	3.3	0.90	1
1	3.2	0.74	1
2	3.3	0.89	1
3	3.4	1.00	1
4	2.4	0.40	1

Multilayer perceptron is imported from sklearn module and created a mlp-clasifier model and x trained data and y trained data is fitted into the model and tried to reduce loss function by taking number of epochs.

Total no of epochs is : 144
The loss function is nearly reduced to: 0.5429 The total taken hidden layers are :4,5
The initial learning rate is: 0.01
Taken random state is: 5

```
Iteration 1, loss = 2.39533516
Iteration 2, loss = 1.45732902
Iteration 3, loss = 1.09416111
Iteration 4, loss = 1.03139572
Iteration 5, loss = 0.98074398
Iteration 6, loss = 0.92037273
Iteration 7, loss = 0.85952581
Iteration 8, loss = 0.81196534
Iteration 9, loss = 0.78202008
Iteration 10, loss = 0.75763907
Iteration 11, loss = 0.74321128
Iteration 12, loss = 0.73625022
Iteration 13, loss = 0.73336120
Iteration 14, loss = 0.73258469
Iteration 15, loss = 0.73242459
Iteration 16, loss = 0.70244369
Iteration 17, loss = 0.65131436
Iteration 18, loss = 0.63142191
Iteration 19, loss = 0.62978614
Iteration 20, loss = 0.62800616
```

```
Iteration 125, loss = 0.54406098
Iteration 126, loss = 0.54364269
Iteration 127, loss = 0.54962569
Iteration 128, loss = 0.54420265
Iteration 129, loss = 0.54246175
Iteration 130, loss = 0.54247451
Iteration 131, loss = 0.54134177
Iteration 132, loss = 0.54063506
Iteration 133, loss = 0.53984179
Iteration 134, loss = 0.54077057
Iteration 135, loss = 0.54331706
Iteration 136, loss = 0.54476847
Iteration 137, loss = 0.54047295
Iteration 138, loss = 0.54052840
Iteration 139, loss = 0.54024175
Iteration 140, loss = 0.54022089
Iteration 141, loss = 0.54023504
Iteration 142, loss = 0.54012461
Iteration 143, loss = 0.54132159
Iteration 144, loss = 0.54297004
```

Finally, accuracy score is imported from sklearn module and prediction of accuracy is found usingthis accuracy score. The Obtained accuracy is 0.73.

```
The accuracy obtained through multi-layer perceptron is
accuracy= 0.7328767123287672
```

VI-CONCLUSION

The use of deep learning Multi-Layer Perceptron (MLP) models for liver disease prediction shows great promise in improving diagnostic accuracy and patient outcomes. By leveraging the power of deep learning algorithms, MLP models can effectively analyze large amounts of data and extract meaningful patterns to predict the presence or progression of liver disease.

One of the key advantages of deep learning MLP models is their ability to automatically learn hierarchical representations from raw input data, without relying on manually engineered features. This feature extraction capability is particularly beneficial in liver disease prediction, as it allows the model to uncover complex relationships between various clinical and demographic factors and the development of liver diseases.

REFERENCES:

1.Wang, X., Yao, X., Zhu, G., Zhang, D., & Zhang, Y. (2020). Deep learning for liver disease diagnosis: A literature review. Computers in Biology and Medicine, 124, 103965.

2.Liu, B., Zhang, X., Zhang, H., Li, H., & Qiu, M. (2018). Deep learning based classification of liver cancer histopathology images using only global labels. Medical Image Analysis, 49, 37-48.

3.Jia, F., Liu, H., Wang, Y., Qian, Y., & Zhang, H. (2019). Liver disease prediction using a hybrid feature selection method and deep learning algorithm. Journal of Medical Systems, 43(6), 153.

4.Toshniwal, D., Yadav, A., & Patel, V. (2020). A deep learning-based approach for the diagnosis of liver diseases using liver histopathology images. SN Computer Science, 1(3), 188.

5.Zhu, G., Yu, Z., Zhu, X., Wang, X., & Zhang, Y. (2021). Deep learning-based diagnosis of non-alcoholic fatty liver disease using histopathology images. BMC Medical Imaging, 21(1), 1-14.

6.Tao, L., Cai, L., Zhu, Y., Li, Y., Yang, F., & Gong, Y. (2021). Deep learning-based automatic liver disease diagnosis using ultrasound images. Journal of Digital Imaging, 34(3), 499-509.

7.Xu, H., Jiang, X., Ma, X., Ji, Z., & Wu, B. (2020). Automatic classification of liver disease using ultrasound images based on convolutional neural networks. Computer Methods and Programs in Biomedicine, 193, 105425.

8.Ding, H., He, X., Zhang, X., Zhang, Q., & Wang, Z. (2020). Liver disease diagnosis based on multimodal data using deep learning. IEEE Access, 8, 183407-183416.

9.Sun, J., Wang, S., Zhao, L., Sun, B., Wu, Z., & Zhu, J. (2020). Deep learning for liver disease diagnosis from CT scans: A comparative study. Medical Physics, 47(3), 1129-1141.

10.Jing, L., Tian, Y., & Liu, Q. (2020). Liver disease diagnosis based on deep learning with CT images. Journal of Healthcare Engineering, 2020, 1-9.