Intelligent Predictive Model for Hepatitis C

1st Mehreen Shahzadi
Department of Data Science
Faculty of Computing &
Information Technology,
University of the Punjab
Lahore, Pakistan
Phddsf22m006@pucit.edu.pk

2nd Faisal Bukhari
Department of Data Science
Faculty of Computing &
Information Technology,
University of the Punjab
Lahore, Pakistan
faisal.bukhari@pucit.edu.pk

3rd Numan Shafi
Department of Computer Science
University of Engineering and
Technology Lahore, Pakistan
numan.shafi@uet.edu.pk

2023 3rd International Conference on Artificial Intelligence (ICAI) | 979-8-3503-2212-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICAI58407.2023.10136685

***Abstract*—Hepatitis C is the liver's festering that can lead to severe liver damage, usually caused by the hepatitis C virus. Hepatitis C has different stages. It is tough to cure in it's last stages; at the same time, it is expensive and painful process. The current research, however, is an alternative precaution to this issue. Hepatitis C can be predicted early by using multiple factors. The dataset related to hepatitis C was not publicly available. To overcome this challenge, the healthy and HCV effected samples were collected from different hospitals in Punjab. A questionnaire based survey was taken including different HCV related factor i.e. gender, weight loss, hives/ rashes, swelling, jaundice, drug addiction history, hepatic encephalopathy (drowsiness, slurred speech), Ascites (fluid buildup in belly/ abdomen), spider angiomas (Spiderlike blood vessels), shared syringe usage, medical history, and severeness. Different cleaning, scaling, and feature selection techniques were applied to collect the best feature data. After selection, various machine learning algorithms were applied. Random forest, KNN, Decision Tree, SVC, and MLP were used, but MLP yielded optimal results in all classification algorithms. We have gained 95.9 % accuracy when tested on unknown data based on the MLP model. As the predictions' results were satisfactory, it would be helpful for the people and act as a critical awareness.**

***Index Terms*—hepatitis C, predictions, KNN, Decision Tree, SVC, MLP**

I. Introduction

HCV stands for Hepatitis C Virus. It is a virus that infects the liver and can cause inflammation, liver damage, and other serious health problems [1]. HCV is spread primarily through contact with the blood of an infected person, such as through sharing needles or other drug injection equipment, or less commonly, through unprotected sexual contact or from mother to child during childbirth [2]. Many people with HCV may not show any symptoms initially and the virus can remain in the body for many years without causing any noticeable damage. However, over time, chronic infection with HCV can lead to liver cirrhosis, liver cancer, and other serious liver diseases [3]. In the late 1960s, it was assumed that only hepatitis A and B types existed. In the early 1970s, a study began on transfusion recipients assuming that hepatitis B is responsible for this transfusion hepatitis. At the time, as viruses of A and B hepatitis were discovered, it came out that neither of them was responsible for transfusion hepatitis. They named new hepatitis non-A non-B hepatitis(NANB). After 15 years, HCV found to be the reason for NANB hepatitis. Chronic hepatitis was developed more frequently and represented as viral infection [4]. Hepatitis C is the festering of the liver, which is usually caused by a viral infection. It is one of the major causes of an increase in the mortality rate worldwide, as its the 12th most common cause of death. A study says approximately 0.17 billion people or 3% of the people infected by HCV. Hepatitis C can be cured if it detects within 8 to 12 weeks [5]. So it declared a viral epidemic. It is five times as widespread as HIV-1 [6].

The liver is the human body's vital organ for filtering blood. A reddish-brown, wedge-shaped organ lies behind the lower part of the ribs and is 7-10.5cm in size. Ribs help to protect it. The liver's main job is filtration of blood that is coming from the digestive tract so that filtered blood is transferred to other parts of the body; by metabolizing drugs, chemicals, fats, etc., and removing wastes [7]. The liver stores necessary chemicals and nutrients and releases them into the bloodstream whenever the body needs them. The liver also produces bile, i.e., a fluid containing cholesterol and bile acids (that help in the digestion of fats) and some essential chemicals (that help in healing and blood clotting). When the liver is inflamed became difficult to perform its job [8]. HCV, is the inflammation or swelling of the liver. Hepatitis C is not the only cause of HCV. It spreads through blood transfusion with an infected person and increases liver cancer risk by 2.5%.

Doctors divided it into two types based on "how long someone has this disease" and "intensity," acute hepatitis C in the early stage within six months. After that, it becomes chronic hepatitis C. 85% of humans with acute hepatitis eventually experience chronic hepatitis C (CHC) disease. HCV is directly connected with liver fibrosis. It is the amassing of extracellular matrix protein, including collagen, and causes the most liver diseases. It eventually developed into cirrhosis, liver failure, hypertension [9], and it has become the common cause of liver transplants in the world. Information on liver fibrosis is essential to predict hepatitis C.

As Hepatitis C is increasing with time, people are using noninvasive methods in staging chronic liver diseases to avoid biopsy drawbacks [10]. Till now, nothing has been developed to detect very early-stage liver fibrosis. The study said that 6-7% of adults live without knowing they have liver fibrosis. They come to know when it grew and became chronic [11].

The problem of Hepatitis C prediction has been extensively discussed in the literature, but it is still not on the list of unsolved problems. This research aims to help people with HCV infection to pay proper attention to their health. The ultimate goal is to find the factors that cause HCV infection, the Relationship between these factors, and the formation of an application that would predict HCV infection at the early stage.

In this article, we focused on predicting Hepatitis C by measuring different factors responsible for HCV infection. The primary goal of this research is to predict Hepatitis C at an early stage. By attaining this goal, we want to assist people in predicting Hepatitis C based on combined parameters efficiently. To achieve this goal, we focused on the following specific objectives:

- Collect the data of people having Hepatitis C as well as healthy people.
- Study of various factors that cause Hepatitis C.
- Relationship of factors that cause Hepatitis C.
- Apply the different machine learning algorithms on the collected dataset and train the model to predict Hepatitis C at early stage.
- Construct an application.

The organization of the rest of the paper is as follows. Section 2 is about related work. The data collection method is explained in Section 3. The preparation of the awareness dataset is discussed in Section 4. Predictive methods are described in Section 5. The results are presented in Section 6. Finally, conclusions are drawn, and future work is discussed in Section 7.

II. Related Work

In this section, I will discuss some critical factors that can cause HCV infection, its effects on the liver, and some related disorders and machine learning models applied in the domain of HCV. I will also discuss treatment and preventive measures to control this disease.

A. Prediction Models in HCV

Hashem et al., [10] proposed the model to predict advanced liver fibrosis by combining the clinical information and serum biomarkers. Dataset of HCV patients was taken and made into two sets. An alternative decision tree method is used to develop two models. One way used six, while the second used four parameters, similar to FIB-4 features. "The model achieved 86.2% negative predictive value and 0.78 ROC with 84.8% accuracy, which is better than FIB-4".

Al Kindhi et al. tested six SVM methods and came out with the best concerning kernel performance in sequence detection in isolated DNA. The dataset consisted of half isolated homo sapiens and half isolated HCV. Edit Levenshtein Distance method used to get the mutation, result saved in SVM, and the target has a positive and negative value of the isolated against HCV. Results of all six tests showed that all methods have variable performance accuracy, but the SVM method was the best to predict the HCV mutation in isolated DNA [12].

Liver biopsy made it very difficult to predict HBV and HCV by predicting hepatic fibrosis (HF) and cirrhosis. It can be diagnosed/indicated without invasive blood test methods and cutting-edge machine algorithms. Hashem et.al., [13] constructed and then tried to compared different ML methods by FIB-4 score on an HBV patient dataset (n = 490). Models were validated on the dataset (n=86). To check the capability, they further check models on HCV patient datasets (n=254, n=230). Gradient boosting (GB) performance is more accurate than other methods along with FIB-4, which scores ($p < 0.001$). The AUROC of FIB-4 and GB models was calculated to classify early and advanced HF, 0.841 and 0.918. Similarly, the GB model shows 0.871 AUROCV, and FIB-4 shows 0.830 for the classification of cirrhosis and non-cirrhosis [13].

Recently, non-invasive techniques have been used to stage chronic liver diseases to avoid biopsy drawbacks. They evaluate different ML methods of predicting advanced hepatic fibrosis to develop classification models using serum-biomarker and clinical information. In [14] HCV patient's dataset of (n = 39,567), divided the dataset into two sets, mild to moderate fibrosis and advanced fibrosis, according to MATAVAR score. Many methods like decision trees, genetic algorithms, and multi-linear regression models were proposed to predict advanced fibrosis. Their performance is evaluated by performing Receiver operating characteristic curve analysis on them. Age, platelet count, AST, and albumin are significant for advanced hepatic fibrosis. This study concluded ML algorithms could predict advanced HF inpatients with AUROC having accuracy between 66.3% and 84.4%.

Most prediction models are complicated and need separate algorithms for specific liver fibrosis and cirrhosis prediction. Using lab data, Chun-Tao Wai aimed to construct a simple model for predicting significant liver fibrosis and liver cirrhosis. The dataset consisted of 25 months with CHC patients details who underwent liver biopsies. He divided the dataset and made validation and training sets. He used the predictive model, including platelets, AST, and Alkaline phosphatase for both (liver fibrosis and cirrhosis) with AUCROC of 0.82 and 0.92, respectively, on the training set. The AUC for APRI on the training set was 0.80 and 0.89. Similarly, The AUC for APRI on the validation set was 0.88 and 0.94, respectively. With more optimization, accuracy could be increased up to 51% for fibrosis and 81% for cirrhosis patients. This study concluded that laboratory results could accurately predict significant liver fibrosis and cirrhosis patients. It also decreases liver biopsy usage among CHC patients [15].

III. Data Collection Method

A dataset with (n = 500) samples is prepared for analysis. The dataset consists of two groups, i.e., experimental and control. The experimental group has (n = 250) samples of HCV-infected people collected from different Punjab hospitals. The control group also contained (n = 250) samples of healthy people. The study was descriptive by nature. A questionnaire was designed, which was based on a binary scale. Different sort of binary scale-based questions, i.e. gender, weight, drug

addiction history, shared syringe usage, medical history, and severeness, were asked in the questionnaire. People from both groups filled out the questionnaires. Data from both groups were compared and correlated.

A. Input Features

Table 1. shows input features used for the prediction of Hepatitis C. It describes the range along with input type for better understanding. Different binary scale-based features as a input, i.e. gender, drug addiction history, shared syringe usage, blood transfusion, piercing of ear, nose or tattoo, partner has Hepatitis C, needle prick, experience of easy bleeding, leg swelling, coloured urine, jaundice, poor appetite, fluid build up in belly, hepatic encephalopathy, spider-like blood vessels have taken. Age is categorize in 5 group, weight is taking number, continuous fever, fatigue, stomach pain, joint pain, early bruises, weight loss, rashes are the intensity based input features(0 referring low).

TABLE I
INPUT FEATURE FOR THE PREDICTION OF HEPATITIS C

Input Features	Properties	
	Type	*Range*
gender (q1)	Binary	0/1[a]
age (q2)	Numbers	0 to 4[b]
weight (q3)	Numbers	n
drugAddiction (q4)	Binary	0/1[c]
sharedSyringe (q5)	Binary	0/1
bloodTransfusion (q6)	Binary	0/1
piercedEarNoseOrTattoo (q7)	Binary	0/1
partnerHasHepatitisC (q8)	Binary	0/1
needlePrick (q9)	Binary	0/1
continousFever (q10)	Numbers	0 to 3[d]
fatigue (q11)	Numbers	0 to 2
continousStomachPain (q12)	Numbers	0 to 3
continousJointPain (q13)	Numbers	0 to 3
continousVomiting (q14)	Numbers	0 to 3
feelNauseaMostOften (q15)	Numbers	0 to 3
itchySkin (16)	Numbers	0 to 2[d]
easilyBruise (q17)	Numbers	0 to 3
easyBleeding (q18)	Binary	0/1
weightLoss (q19)	Numbers	0 to 2
swellingInLegs (q20)	Binary	0/1
rashesOnSkin (q21)	Numbers	0 to 2
darkColorUrine (q22)	Binary	0/1
jaundice (q23)	Binary	0/1
rednessOnHands (q24)	Binary	0/1
clayColouredPoop (q25)	Binary	0/1
poorAppetite (q26)	Binary	0/1
fluidBuildUpInBelly (q27)	Binary	0/1
hepaticEncephalopathy (q28)	Binary	0/1
spiderLikeBloodVessel (q29)	Binary	0/1
has hepatitis (q30)	Binary	0/1

[a] male/female
[b] 0= under 15, 1= 15-20 , 2= 21-40 , 3= 41-60, 4= above 60
[c] yes/no
[d] intensity
n = any number

Some graphs are listed to show the statistics of the data collected from different hospitals.

Figure 1. shows that different age groups have Hepatitis C. Most patients lie in 40 plus age group, which indicates that this

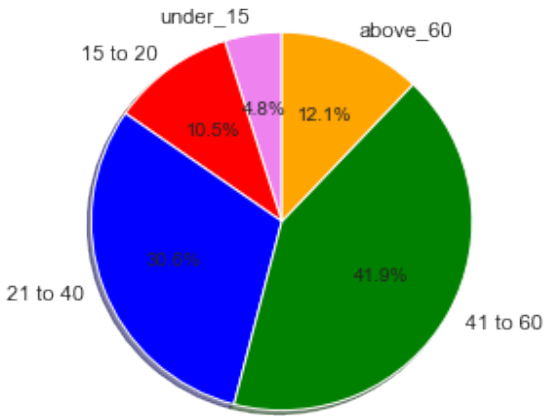

Fig. 1. Age Analysis.

age group has more chances of being affected. As they have more probability of getting in contact with infected people's blood by different means, like teeth treatment, one can be infected due to the carelessness of a dentist. So, you need to be very careful when having any blood transfusion cause activity. Hepatitis C shows significantly fewer symptoms in its early stages, but once it is detected in its early stage (within 8 to 12 weeks), it can be straightforward to cure. For this, it needs to be tested if someone indulged in hepatitis C-causing activity.

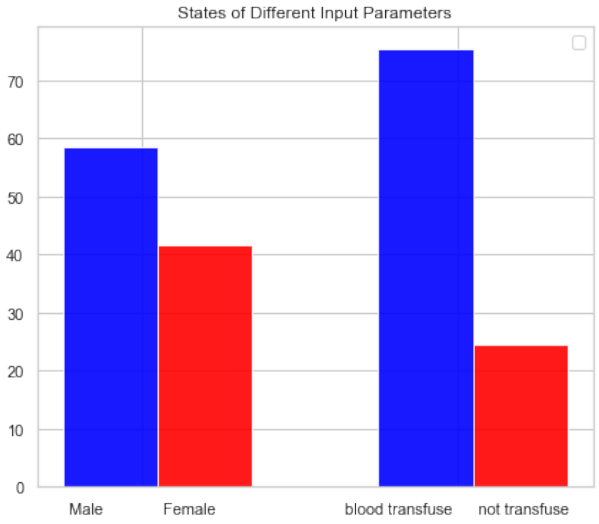

Fig. 2. Statistical analysis of different parameters.

Figure 2. shows that males have more numbers than females having hepatitis C. This graph shows that more people having hepatitis C did any kind of blood transfusion activity like blood donation, ear/nose piercing, tattoo making, syringe Sharing, etc.

IV. Preparation of the HCV Dataset

A. Formatting

The collected data were in raw form, so the process of data formatting was very crucial. Each feature of a data point had a different textual name. For example, gender had two inputs, male and female, which were converted into binary form, i.e., 1/0. Similarly, weight, age, continuous fever, and other similar features were transformed into binary format. Data entry had a yes response representing "1" and no describing "0". The purpose of converting all the features to the binary scale was to apply different ML algorithms.

B. Data Cleaning

In this process, our primary goal uses to deal with missing entities and unwanted characters. We replaced missing entities with the most repeating and occurring value of the particular feature subjected to its related feature. That is, we select all the data points that have the same output, and check their corresponding attribute's value to gain the most occurring value, then replace the missing value with the most occurring value. For instance, Itchy skin is divided into three features depending on its intensity, i.e., severe, moderate, no itching, and their missing feature's value replaced with 0, i.e., most occurring value, only one can have one value as at the same time someone could not have severed and moderate itchy skin or no itchy skin. So, every feature will have binary input. One represents yes have the mentioned problem, and 0 means no do not have.

C. Feature Extraction

For this purpose, Variance Threshold and SelaectKBest feature selection methods were applied. By this, the Variance Threshold method eliminated features having slight variance. Up to 12 of the 30 features took part in HCV predictions, and only features have been selected that contributed more to achieving the target value. We achieved an accuracy of 86.95% on unseen data. When we chose 19 from 30 features, the accuracy decreased to 86.32% on test data. So the conclusion is that it is significantly lower than the 92.38% MLP model. That is why we chose MLP to predict HCV.

V. Prediction Methods

we used different machine learning algorithms including K-nearest neighbour, Decision tree, random forest, Support vector machine, Multi-layer perceptron to train collected data. Deep learning and enforcement learning require large number of data points and takes time to train dataset.

A. k-nearest neighbors (KNN) algorithm

KNN stands for k-nearest neighbors used in the search, classification, and pattern reorganization. It is a type of supervised machine learning algorithm. The data is classified based on given functions and distances, which basically works on feature similarity to predict the class for new data points. Distances are calculated for the unique data point from all k classes, find the shortest distance from the class, and assign the data point to the closest class. The algorithm gets significantly slower as the data increases. It is mostly used for the classification purpose. [16]

B. Decision Tree Algorithm

The decision trees algorithm is also belonging to the supervised learning algorithm family. Decision trees are very flexible and are used for both types of classification and regression. It works downward and decides with the "If this then that" behavior. Decision trees are easy to explain, straightforward, fast, and suitable for large datasets.

Decision Tree Classifier is the algorithm in which trees are built through a process of binary recursive partitioning, and a tree-like structure is formed. The topmost node is the root node, branches represent decision rules, and the leaf node shows the final decision/output. The decision tree gives an optimal solution at each step without determining the last level's optimal solution. [17]

C. Support Vector Classifier (SVC)

SVC analyzes data for classification or regression; however, most of the time used for classification. SVC is used to find the best fit hyper-plane in N-dimensional space for new data points, then predict the class by adding some features to your classifier. SVC is useful for high-dimensional spaces using a subset of training points. In the case of extensive dataset and overlapped data, it does not work on the mark. [18]

D. Random Forest (RF)

Random forests (random decision forests) are the collection of Decision Trees. It is widely used for data clustering, feature selection, and statistical reasoning. This forest mainly uses numerical and categorical data. Random forest is very slow to use and has over-fitting issues. [19]

The word forest shows a collection of trees someplace. The same thing is happening in a random forest algorithm. When new data comes for 42 classifications, several trees are formed, then each tree classifies that point based on decision rules. As a result, the new comes when assigned to that class, with a maximum number of votes by the trees.

E. Multilayer Perceptron (MLP)

A Multilayer Perceptron (MLP) is the FeedForward Artificial Neural Network classification. As its name shows, MLP has multiple layers, input, and output layers, for receiving signals and deciding results; between them, a bundle of hidden layers for the computation is also called the computation engine of MLP. Perceptron is a linear classifier algorithm to divide input into two parts with a straight line. If an MLP has a linear activation function in each neuron, then linear algebra shows that any layers can be reduced to an input-output layer, two-layered.

Figure 3.1 shows the MLP model used for the Hepatitis C prediction. total of 30 features are used on the input layer. The count of the hidden layers is eight, and getting binary output, which means that person has Hepatitis C or not. Activation

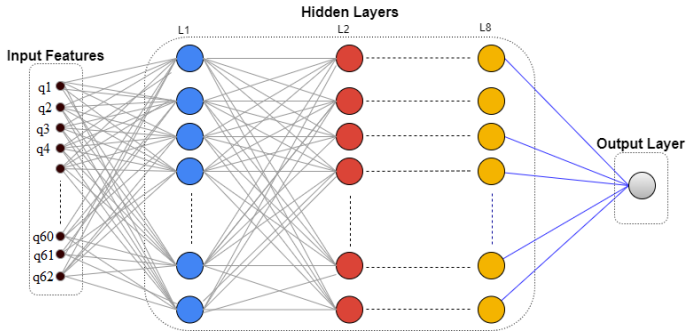

Fig. 3. Optimal MLPs model for Hepatitis C Prediction.

function Relu is used in this model. All the parameters on which this model is prepared is obtained by Hyperparameter tuning of these factors, and the best factors are selected in these hyperparameter tuning. The number of epochs is used from 0 to 40. The best accuracy is obtained when the epoch value is 35 to 39.

VI. Experiments and Results

Different techniques were applied to collect data for cleaning, scaling, and filtration. A model was trained based on the data through various machine learning algorithms. This section depicts the after effects of this research study. The results of the test data to predict Hepatitis C are given below.

TABLE II
ACCURACY OF DIFFERENT MACHINE LEARNING ALGORITHMS

Algorithms	Calculated Parameters				
	Accuracy	*Precision*	*Recall*	*F-Measure*	*Hyper-Parameter*
Decision Tree Classifier	83.80	0.90	0.80	0.85	max depth = 8, min sample leaf = 1, min sample split = 2
Random Forest	86.66	0.89	0.89	0.89	max depth = none, n estimators = 5
KNeighbors Classifier	86.30	0.88	0.68	0.80	n neighbors = 7
SVC	89.52	0.93	0.88	0.90	C = is 1, gamma = 0.001, kernel = linear
MLP	92.38	0.94	0.81	0.87	activation = relu, alpha = 0.0001, hidden layers = 8, perceptron = 18, LR[a] = adaptive, solver = adam

[a]Learning rate.

Table 4.1 shows that the best accuracy is achieved on the MLP model. Therefore, we tune MLP with different parameters to obtain maximum accuracy. Figure 4.1 shows the tuning of the model by using a different number of layers. Figure 4.2 shows the tuning of the model by using a different number of perceptrons. Figure 4.3 is about AUC (Area under the Graph) to show the trend between the true positive rate and false-positive rate. Figure 4.4 shows the accuracy of MLP on each number of iterations.

Figure 4. shows the accuracy of the MLP model with a different number of layers. The number of layers varying from

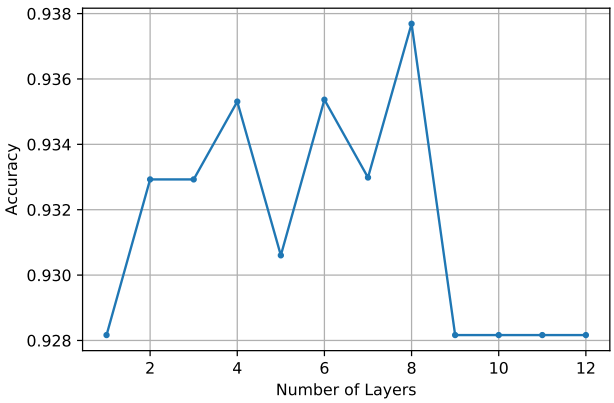

Fig. 4. Accuracy with different number of layers for hepatitis C prediction

1 to 12 is used to check the MLP model's accuracy. The graph shows maximum accuracy at several layers 8, which is 93.7. Minimum accuracy is observed at several layers 9 to 12, which is below 93.0 percent.

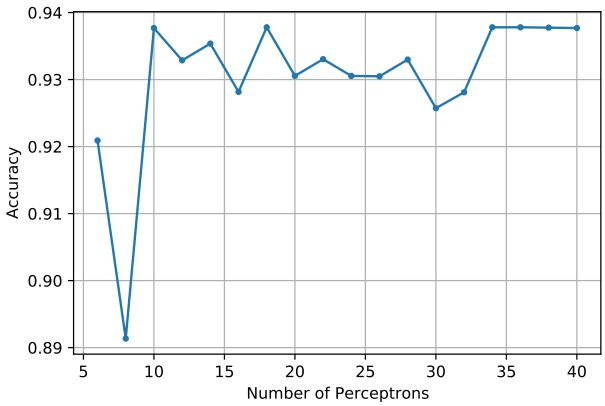

Fig. 5. Accuracy with different number of perceptrons for hepatitis C prediction

Figure 5. shows the MLP model's accuracy with a different number of perceptrons varying from 1 to 40 on each layer. Maximum accuracy is observed when several perceptrons are 18, which is 93.9%, and minimum accuracy is observed at 8, below 90%.

Figure 6. AUC (Area under Curve) shows how accurately the data is separated. It means checking how many people have hepatitis C and whether they are classified as hepatitis C patients and vice Versa. AUC value is 97 percent, which means that data is classified more correctly and accurately.

Figure 7. shows the tuning of the epochs ranging from 0 to 100. At the start, from 0 to 10 graph is more fluctuated, rapid change occurs in the graph, and accuracy increases rapidly. Afterward, the graph becomes smooth, and accuracy is achieved when epochs are 78.

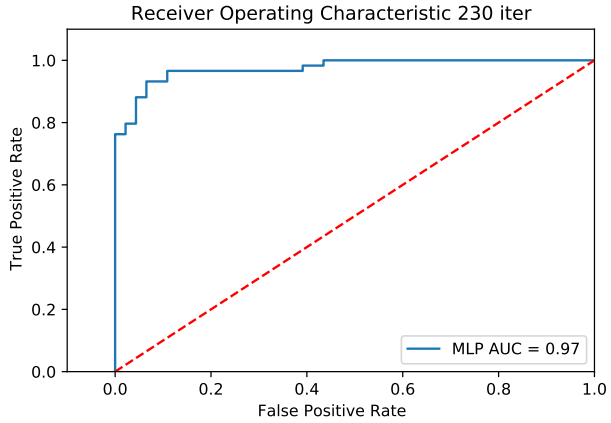

Fig. 6. AUC to check how accurately data is classified.

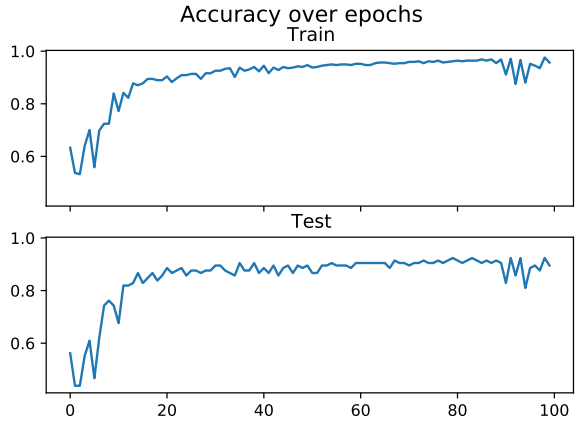

Fig. 7. Accuracy with different numbers of epochs in MLP model.

VII. Conclusion

After looking at data points from various angles, the research concluded that MLP is an optimal model for classifying and predicting Hepatitis C data. Due to a shortage of time, I collected 500 samples for the hepatitis C analysis and trained the model on it. For future research, I recommend that the experiment can be performed on more data to be more confident about the predictions. More males have Hepatitis C than females. Be careful during a blood transfusion, ear/nose piercing, tattoo making, blood donation, dentist visit, etc. Avoid syringe sharing and all instruments reuse without sterilization that can directly connect with blood. It can transmit the virus from one infected person to another healthy person. Our analysis shows more infected people share or did any blood transfusion activity in their past. If you did any blood transfusion activity, you should have done a Hepatitis C test on having any minor doubt.

Acknowledgment

Dr. Faisal Bukhari gives the main idea of the paper. He also contributed to the writing of the paper. Mehreen Shahzadi contributed to data collection, cleaning, scaling, analyzing results, and writing the paper. Numan Shafi helped in terms of applying different Machine Learning Algorithms to the dataset in the prediction of hepatitis C. The author(s) received no funding for the research,

References

[1] W. Elgretli, T. Chen, N. Kronfli, and G. Sebastiani, “Hepatitis c virus-lipid interplay: Pathogenesis and clinical impact,” Biomedicines, vol. 11, no. 2, p. 271, 2023
[2] E. M. Redwan, A. A. Aljadawi, and V. N. Uversky, “Hepatitis c virus infection and intrinsic disorder in the signaling pathways induced by toll-like receptors,” Biology, vol. 11, no. 7, p. 1091, 2022.
[3] R. Q. Ferrufino, C. Rodrigues, G. M. Figueiredo, et al., “Factors associated with spontaneous clearance of recently acquired hepatitis c virus among hiv-positive men in brazil,” Viruses, vol. 15, no. 2, p. 314, 2023.
[4] L. B. Seeff, ”The history of the ”natural history” of hepatitis c (1968–2009),” Liver International, vol. 29, pp. 89–99, 2009.
[5] https://www.cdc.gov/hepatitis/hcv/index.html
[6] G. M. Lauer and B. D. Walker, ”Hepatitis c virus infection,” New England journal of medicine, vol. 345, no. 1, pp. 41–52, 2001.
[7] M. Matthew Hoffman. (). Liver anatomy, [Online]. Available:https://www.webmd.com/digestive-disorders/picture-of-the-liver2.
[8] D. Nyberg. (). Understand liver, [Online]. Available: https://www.medicalnewstoday. com/articles/293082.
[9] R. Bataller and D. A. Brenner, ”Liver fibrosis,” The Journal of clinical inves- litigation, vol. 115, no. 2, pp. 209–218, 2005.
[10] S. Hashem, G. Esmat, W. Elakel, S. Habashy, S. Abdel Raouf, S. Darweesh, M. Soliman, M. Elhefnawi, M. El-Adawy, and M. ElHefnawi, ”Accurate prediction of advanced liver fibrosis using the decision tree learning algorithm in chronic hepatitis c Egyptian patients,” Gastroenterology research and practice, vol. 2016,2016.
[11] P. Gin‘es, I. Graupera, F. Lammert, P. Angeli, L. Caballeria, A. Krag, I. N. Guha, S. D. Murad, and L. Castera, ”Screening for liver fibrosis in the general population: A call for action,” The lancet Gastroenterology & hepatology, vol. 1, no. 3, pp. 256–260, 2016.
[12] B. Al Kindhi, T. A. Sardjono, and M. H. Purnomo, “Optimasi support vector machine untuk memprediksi adanya mutasi pada dna hepatitis c virus,” Jurnal Nasional Teknik Elektro dan Teknologi Informasi, vol. 7, no. 3, pp. 317–323, 2018.
[13] R. Wei, J. Wang, X. Wang, G. Xie, Y. Wang, H. Zhang, C.-Y. Peng, C. Rajani, S. Kwee, P. Liu, et al., ”Clinical prediction of HBV and HCV related hepatic fibrosis using machine learning,” EBioMedicine, vol. 35, pp. 124–132, 2018.
[14] S. Hashem, G. Esmat, W. Elakel, S. Habashy, S. A. Raouf, M. Elhefnawi, M. I. Eladawy, and M. ElHefnawi, ”Comparison of machine learning approaches for prediction of advanced liver fibrosis in chronic hepatitis c patients,” IEEE/ACM transactions on computational biology and bioinformatics, vol. 15, no. 3, pp. 861– 868, 2017.
[15] D. Nyberg. (). Understand liver, [Online]. Available: https://www.medicalnewstoday. com/articles/293082.
[16] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, “Comparative performance analysis of k-nearest neighbour (knn) algorithm and its different variants for disease prediction,” Scientific Reports, vol. 12, no. 1, pp. 1–11, 2022.
[17] F. Wu, X. Liu, Y. Wang, X. Li, and M. Zhou, “Research on evaluation model of hospital informatization level based on decision tree algorithm,” Security and Communication Networks, vol. 2022, pp. 1–9, 2022
[18] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, “A comprehensive comparative study of artificial neural network (ann) and support vector machines (svm) on stock forecasting,” Annals of Data Science, vol. 10, no. 1, pp. 183–208, 2023.
[19] N. Jain and P. K. Jana, “Lrf: A logically randomized forest algorithm for clas- sification and regression problems,” Expert Systems with Applications, vol. 213, p. 119 225, 2023