Machine Learning in liver disease diagnosis: Current progress and future opportunities

Neha Tanwar¹ and Khandakar Faridar Rahman²

¹ Research Scholar, Banasthali Vidyapith, Rajasthan-304022, India
² Assistant Professor, Banasthali Vidyapith, Rajasthan-304022, India

E-mail: nehatanwar0707@gmail.com, kfrahman@banasthali.in

Abstract: There has been a rapid growth in the use of automatic decision-making systems and tools in the medical domain. By using the concepts of big data, deep learning, and machine learning, these systems extract useful information from large medical datasets and help physicians in making accurate and timely decisions regarding predictions and diagnosis of diseases. In this regard, this study provides an extensive review of the progress of applying Artificial Intelligence in forecasting and detecting liver diseases and then summarizes related limitations of the studies followed by future research.

Keywords: Liver Diseases, Machine Learning, Data Mining, Deep Learning

1. Introduction:
The liver is the largest organ in the human body. It is responsible for all metabolic functions within the body from the conversion of nutrients within the diet into usable body substances to storing these substances and then supplying them to the cells when required. It is also responsible for the conversion of toxic substances into harmless substances. Other vital functions of the liver include bile production, protein production, storing and releasing glucose, processing haemoglobin, blood cleaning, immune factor production, clearing bilirubin, etc. Thus, it is the primary and most crucial body organ, and the maintenance of its health is essential for improved overall health. But the fact is that people generally overlook it in the case of health. Due to unhealthy lifestyle routines, most of the population across the globe are suffering from acute to severe liver problems.

Liver diseases include:

- Viral infections, e.g., hepatitis A, hepatitis B, and hepatitis C,
- Immune system problems, e.g., autoimmune hepatitis, primary biliary cholangitis, primary sclerosing cholangitis,
- Diseases caused due to drugs, poisons, or high alcohol consumption, e.g., fatty liver disease, non-alcoholic fatty liver disease (NAFLD), non-alcoholic steatohepatitis (NASH) and, cirrhosis,
Inherited diseases, e.g., hemochromatosis, hyperoxaluria, alpha-1 antitrypsin deficiency, and Wilson disease and
Cancer and tumour, e.g., Liver Cancer, bile duct cancer, and liver cell adenoma.

The prevention of liver failure is possible by diagnosing and treating liver diseases at an early stage. There are four stages of liver disease, among which the first stage is marked by inflammation, which may or may not show any symptoms in patients. Prolonged inflammation replaces the healthy liver tissue with the scar tissue, which causes the disease to enter in the second stage, i.e., fibrosis, which is also mostly asymptomatic. Severe scarring on the liver causes cirrhosis, which is the third stage. In this stage, the patient starts to experience symptoms like abdominal pain, fatigue, weakness, jaundice, etc. When there is a drastic deterioration in liver function, then it is End-stage liver disease (ESLD). In this, the patient shows severe complications but can be treated without transplanting the liver. In the fourth stage, unhealthy cells start to develop and expand, which causes Liver cancer. These conditions pointed to the need for methods that provide early prediction of liver diseases so that the effect can be mitigated, and damage can be controlled by providing appropriate treatment at an early stage.

The commonly used diagnosis and tests for liver diseases or Hepatic diseases include a liver blood test, complete blood count (CBC), Abdominal and Pelvic CT, Abdominal Ultrasound, Elastography, ERCP (Endoscopic Retrograde Cholangiopancreatography), Lactate Dehydrogenase (LDH) Isoenzymes Test, Lactate Dehydrogenase (LDH) Test, Liver Biopsy, Liver Function Tests (LFT), Magnetic Resonance Cholangiopancreatography (MRCP), MRI of the body (Chest, Abdomen, Pelvis), and Needle Biopsy.

The conventional methods of treating liver diseases suffer from various limitations. Some of them are as follows:

- A large amount of medical data is produced, whereas the number of expert observers to analyze this data is significantly less. Moreover, a physician may or may not have expertise in the analysis of various types of data and images.
- Discovering hidden patterns and relationships in massive medical data is often considered unimportant.
- The ultrasound images may contain noise, can be of bad quality, and are often different from different machines that affect the diagnosis. Also, manual detection can be slow and ambiguous.
- Liver biopsy is risky and often interpreted differently by different observers.
- Liver Transplantation in hepatocellular carcinoma (HCC) is rarely performed because of organ shortage, and hence other effective treatments such as liver resection (LR) are given more importance. But this can lead to a reoccurrence of the disease in patients with high risk.
- A single biomarker is not sufficient for disease prediction. Thus, it is essential to use a combination of biomarkers to increase diagnostic accuracy. Moreover, the Biomarkers which are commonly used for diagnosis of diseases may give false results.
- Detection of NAFLD at an early stage can protect the liver from further damage and thus can prevent the patient from severe diseases such as HCC. There are various non-invasive methods of detecting NAFLD, but these often lack accuracy because of inaccurate blood-marker tests and may include costly imaging processes.
- HCC can be significantly predicted by microvascular invasion (mVI), but the pre-operative assessment of mVI is difficult.
- Prediction of postoperative survival is generally not given much importance, which leads to the reoccurrence of the disease.
- Disease diagnosis by physicians is less accurate and time-consuming.

Automatic systems can be used to overcome these limitations of conventional disease diagnosis methods. The main aim of clinical decisions is to provide a correct and timely diagnosis. Thus, an automatic disease diagnosis system can help a physician or even patients to make precise and accurate predictions of various liver diseases. Clinical decision support system (CDSS) [1] aids physicians and
patients in making a timely diagnosis and taking the necessary treatment to mitigate the effects of diseases. Various data mining and machine learning (ML) algorithms have been used in these CDSSs to obtain the desired outcomes.

Advancement in computer science is increasing rapidly, and especially the concepts of artificial intelligence (AI), namely deep learning (DL), ML, and big data, are being used in various fields for making predictions by gaining a large amount of information from data.

Due to its unmatched powers, AI is also used in Healthcare for analyzing records of patients to deduce information, early disease diagnosis, and thus assisting physicians and patients in deciding treatments. A well-known field of computer science and engineering called Knowledge discovery in large databases (KDD) has been widely used for extracting essential and useful information from large medical datasets, which usually are challenging to be deduced by a medical expert in a short period time. Thus, the application of ML classification algorithms for handling extensive medical data and making clinical predictions is beneficial for identifying factors that help in the prognosis of patients with various liver diseases. KDD involves a series of steps that are to be carried out properly, and all levels are equally important. The mechanism of KDD includes- selection, pre-processing, transformation, data-mining, interpretation, and evaluation.

2. Creating a ML model:
This section describes all the steps involved in creating a machine learning model for the diagnosis and prediction of liver diseases. These steps are defined as follows:

2.1. Data collection and pre-processing- The records of patients are used for deducing information, which helps in the prognosis of mild to severe liver diseases. The commonly used data include Ultrasound images, CT Scans, MRI scans, Blood-test reports, general information about the patient including age, gender, BMI, alcohol consumption, eating habits, child-pugh stage, histologic grading, etc.

The data used can be either primary or secondary, single-centred or multi-centred, prospective, or retrospective. Primary data is the data collected from hospitals, medical institutions, or testing centres by the researcher himself/herself. Secondary data is the one which is managed by the researcher from online repositories or from some other researcher for performing the analysis. Single-centred data is the one that is collected from one single location, e.g., a hospital or a medical institution from a specific country. In contrast, multi-centre data is the one that is collected from multiple sites, which can even be from different countries. Here, most of the studies reviewed were performed on secondary data acquired from a single-centre. Few studies were also conducted on both large-scale and small-scale primary data obtained from either single-centre or multiple centres. By the term prospective data, we mean data collected when there is a change in the characteristics with time. In contrast, retrospective data is the one that has already been raised in the past.

In data mining and machine learning, the predictive performance can be optimized by pre-processing of data, which converts the raw data into clean data. It includes cleaning, normalization, transformation, filling the missing values, checking inconsistent values, removing duplicate values, removing noisy data, feature aggregation, feature sampling, dimensionality reduction, feature encoding. It can be performed by using various tools such as R framework, Weka, Rapidminer, Python, SPSS, etc.

2.2. Region of interest (ROI) segmentation- When the dataset to be used contains images, then the pre-processing is done by image segmentation. It is a process of separating the principal and unimportant areas within the image and thus making it more meaningful and hence, easy to analyze by applying statistical data mining methods. The ROI is the collection of the segmented regions of the image, which are to be focused on extracting useful information. It is either marked by a radiologist or expert physician or is determined by adjusting the pixel intensity values.

2.3. Feature Selection- Feature selection can be defined as the process of selecting the significant features that are strongly related to the output from the dataset for faster training of the model, reduced
complexity, reduced dimensionality, straightforward interpretation, and improved accuracy. Significant variables/biomarkers are extracted from the records of clinical information and laboratory tests of patients by using statistical data mining and ML algorithms. Above mentioned tools for pre-processing also contains packages that allow feature selection.

2.4. Cross-Validation and Splitting of data set– In cross-validation, a portion of the dataset is reserved for testing the model, and the remaining data is used for training the model. It avoids the problem of overfitting and underfitting. Commonly used cross-validation techniques are ‘Hold-out cross-validation (early stopping)’ and ‘K-fold cross-validation’ [2]. Data splitting is a division of a dataset into two or three parts, i.e., Training data, Validation data, and Test data to eliminate bias. It is done according to a split ratio that depends on the dataset being used and the type of model being created. It is performed according to a sampling technique, which can be either Simple random sampling, trail-and-error sampling, systematic sampling, convenience sampling, CADEX, DUPLEX, or stratified sampling [2]. The most commonly used sampling technique by ML algorithms is simple random sampling. The ML algorithms are trained by using the training data to develop the predictive model, which is then validated by using the validation data. The model learns from the training data, whereas it does not learn from the validation data and only uses it for improving the hyperparameters. After deciding the hyperparameters, the test data is used to check the performance of the created model.

2.5. Classification algorithms - Classification can be defined as the process of creating a model of class attributes that are obtained from a dataset so that a class label can be assigned to an unseen record with high accuracy. ML algorithms use the training data with significant features for performing classification. There are various types of ML algorithms that are selected according to the kind of predictive and diagnostic model to be created. The ML algorithms can be broadly classified into three categories, namely- Supervised learning, Unsupervised Learning, and Reinforcement Learning. A brief description of them is as follows:

2.5.1. Supervised learning- These algorithms map an input to the desired output and work according to a result that is to be achieved. With these algorithms, the training process is continued until the model is created. Examples of these algorithms include Decision Tree (DT), Random Forest (RF), K-nearest neighbour (KNN), Logistic Regression, Support Vector Machine (SVM), Naïve Bayes (NB), etc.

2.5.2. Unsupervised learning- In these algorithms, no target result has to be achieved and is mainly used for solving clustering problems. Examples of these algorithms include the Apriori algorithm, k-means, etc.

2.5.3. Reinforcement learning- These algorithms are used when specific decisions are to be made by the machine, which trains itself using trial and error method and tries to improve its performance by learning from experience. Examples of these algorithms include Markov Decision Process.

Here the aim is creating a predictive and diagnostic model that is expected to deliver results within a specific boundary. Therefore, supervised learning algorithms are generally used for creating an automatic disease detection model.

2.6. Predictive model creation- This step produces the ML model, which serves the desired results and hypothesis. The performance of algorithms is compared to select the one that gives the highest accuracy by using a cross-validation method. By using the tuned dataset containing significant attributes is used for training and validation of the ML classification algorithm, which yields the final predictive or diagnostic model giving automatic decisions about the new and unseen dataset.

2.7. Evaluation- The evaluation of the produced model is carried out based on confusion matrix, accuracy, precision, recall, specificity, F1 score, ROC (receiver operating characteristics) curve, true-positive rate (TPR), false-positive rate (FPR).
3. Related Studies:

This section includes studies related to the use of AI in the field of liver disease diagnosis in the last six years. The Indian Liver Patient Dataset (ILPD) from the UCI machine repository [3] has been used in many studies ([4,5,14–16,6–13]). The dataset contains 583 records for 11 attributes in which class 1 includes 416 liver patients, and class 2 includes 167 non-liver patients. Most of these studies used 80% of the dataset for training the model, and the remaining 20% for testing the model and the classification was carried out by using 10-fold cross-validation. BUPA dataset from the UCI machine repository [3] was used by ([17–19]). It consists of 345 different instances of 7 attributes, and one extra attribute was added that indicate the presence and severity of the disease. Many researchers also used primary data from hospitals and medical institutions. The review of these studies is as follows:

3.1. Related Studies in 2015:

The predictive capability of C5.0 and C4.5, SVM, DT, KNN, LR, NB, RF, and NN classifiers was compared using ILPD with Rapid Miner and IBM SPSS Modeler tools. From Rapid Miner Software, SVM showed the highest accuracy (72.54%), and from the SPSS Software, C5.0 showed the highest accuracy (87.91%) [4]. In another study [6], NB and SVM algorithms were compared, and it was observed that SVM outperformed the NB algorithm by providing more accurate classification on the ILPD dataset. A hybrid model called NeuroSVM was produced by combining SVM and ANN, which achieved an accuracy of 98.83% in the classification of ILPD [7]. Using the CT scans of 80 patients, the significant features were extracted from the ROI using SFSS and GA (genetic algorithm). Probabilistic Neural Network (PNN), Linear Vector Quantization (LVQ) Neural Network, and Back Propagation Neural Network (BPNN) were used for the classification of fatty and cirrhosis liver. PNN outperformed LVQ and BPN with accuracy (recognition rate) of 95% [20].

In a comparison of SWE measurements and biopsy scores in estimating fibrosis in patients with chronic liver disease, the empirical results showed that SWE estimates of liver stiffness are correlated with the severity of fibrosis and can be used as a tool for differentiating fibrosis patients [21].

3.2. Related Studies in 2016:

A novel Neuro-Fuzzy inference system provided an accuracy of 79/83% after classification with significant factors being total bilirubin (TB), direct bilirubin (DB), alkaline phosphate (Alkphos), serum glutamic-oxaloacetic transaminase (Sgot) from ILPD [8]. A ML model that could predict the stage of liver fibrosis using the DT classifier was proposed by using the records of 100 HCV patients. After the pre-processing of data, the essential features were selected using the ANOVA and BOX plot tests. The achieved accuracy was 93.7% [22].

Instead of blood markers, ten other variables were used from the medical record of 2060 patients, namely sex, age, alcoholic cirrhosis, non-alcoholic cirrhosis, alcoholic hepatitis, viral hepatitis, different types of chronic hepatitis, alcoholic fatty liver disease, other types of fatty liver disease, and hyperlipidaemia. The dataset was divided into two groups of 70% and 30% for training and testing, respectively of ANN, LR, SVM, and DT classifiers. Among them, ANN showed the best results with AUROC of 0.873 [23]. 512 features were extracted from 100 liver ultrasound images with the help of GIST descriptors. The feature set was reduced to 18 by using the Marginal Fisher Analysis (MFA) integrated with the Wilcoxon signed-rank test. Then, the images along with the significant features were fed as input to the classifiers, namely DT, Adaboost, KNN, PNN, SVM, NB Fuzzy Sugeno, linear, and quadratic discriminant analysis with 10-fold cross-validation. The evaluation proved that PNN provided a better diagnosis of fatty liver disease with a classification accuracy of 98% [24].

J48, NB, MLP ANN, ZeroR, 1BK, and VFI algorithms were compared for classification of LFT dataset for predicting liver diseases. The analysis showed that MLP outperformed other algorithms [25]. KDD and ML approach could detect fibrosis in HCV patients [26]. ANOVA method was used for selecting essential features from the records of 100 HCV patients, and medium KNN, linear SVM, and cubic SVM algorithms were used for classification. The evaluation showed that cubic SVM provided the highest accuracy of 98%. A multi-layer classifier ensemble model named HMV (hierarchical majority
voting) was proposed [27]. The seven classifiers used were NB, LR, QDA, K-NN, SVM, DT using information gain, and DT using the Gini index with 10-fold cross-validation. The proposed framework was evaluated on various datasets of different diseases and showed high diagnosis efficiency for all. A novel diagnosis system was proposed for FLD using 124 ultrasound images of 62 patients. 128 significant features were extracted and used for classification by the BPNN algorithm after random partitioning. The achieved accuracy was 97.58% [28].

3.3. Related Studies in 2017:
Boosted C5.0 outperformed CHAID with 93.75% accuracy and produced 92 rules for disease detection using ILPD [5]. These rules were optimized and 24 rules were generated by applying GA, which increased the accuracy of the Boosted C5.0 algorithm [9]. Meta-learning algorithms, namely Adaboost, Logitboost, Bagging, and Grading, were compared for classification of ILPD [10]. It was concluded that grading was the best algorithm among these meta-learning algorithms as it provided the highest Correct Classification rate, minimum incorrect classification rate and showed the least Execution time. The gut microbiome can provide biomarkers that aid the diagnosis of advanced fibrosis in NAFLD. The feature selection stage provided 40 significant features, which are age, BMI, and 37 bacterial species. By using these features, the RF classifier achieved an AUROC of 0.936 in detecting advanced fibrosis [29].

A novel approach using LWA (learning with abstention) method classifies liver ultrasound images as normal or abnormal, and do not classify the data if its unsure of the accurate prediction. A method of structural risk minimization was adopted for developing LWA, and it was optimized using a stochastic gradient-based solver. The features were extracted from the ROI within 99 ultrasound images and were used for training the NN, SVM, and LWA classifiers with 5-fold cross-validation. The proposed LWA method outperformed other classifiers with AUROC of 78% [30].

A study [31] presented a decision-support system for detecting the hepatitis B stage, using 513 RTE images of patients. Age, gender, and 11 RTE features extracted from the images were used to train the NB, RF, KNN, and SVM classifiers. Among the classifiers, the RF algorithm achieved the best accuracy of 91.25%. Yip et al. built a model for detecting NAFLD by dividing the data of 922 patients into training (n=500) and testing (n=422) sets. The elastic net regularisation method extracted 6 significant features from 23 parameters. Four ML classifiers (logistic regression, ridge regression, decision tree, and AdaBoost) were trained using these features. They were evaluated in the context of selecting the model that gives the best prediction accuracy. The results showed that ridge regression achieved the best AUROC of 0.87 and 0.88 in the training and validation groups, respectively [32].

3.4. Related Studies in 2018:
In a study, the C5.0, CHAID (Chi-square Automatic Interaction Detector), and CART (classification and regression tree) algorithms were combined with MLPNN, and their performance was compared in the classification of ILPD. Among them, MLPNN-C5.0 showed the highest accuracy of 94.12% [33]. Among J48, logistic model tree (LMT), RF, Random tree, REPTree, Decision Stump, and Hoeffding Trees; the Decision Stump showed the highest accuracy of 70.67% than other techniques in the classification of ILPD [34]. J48, MLP, Bayesnet, RF, and SVM classifiers were used for classifying the ILPD dataset pre-processed by min-max normalization. The feature selection was performed by PSO algorithm. Among them, J48 showed the highest accuracy of 95.04% [11]. In a study, the predicting capability of particle swarm optimization (PSO), GA, MReg (multilinear regression), and alternating decision tree (ADT) algorithms were compared by using medical data of 39,567 patients. The Kruskal-Wallis and Chi-Square tests concluded that age, platelet count, AST (aspartate aminotransferase), and albumin have a strong relation with advanced fibrosis, and hence were used for model creation. By using random uniform sampling, the dataset was divided into training (22,690 patients), and testing (16,877) sets. The four algorithms were applied for classification using 10-fold cross-validation. The results depicted that the ADT model showed the highest efficiency and AUROC of 84.4% and 0.76, respectively [35].
A study [19] compared SMO, J46, NB, and Bayes net by using the BUPA dataset and concluded that SMO achieved the highest accuracy, but NB took the least time in creating the model. Thus, if the time factor is ignored then, SMO performed best. Another study [18] used RF and ANN for classification with 10-fold cross-validation and divided the BUPA dataset into 90% and 10% for training and testing set, respectively. Their performance was compared with Gaussian processes, linear LR, MLP NN, SVM, AIRS, and CBR-PSO. It was concluded that ANN showed the highest accuracy of 85.28%. DCGAN (Deep Convolutional Generative Adversarial Network) can generate synthetic data samples from the available dataset. Training features were extracted from the ROI in 64 CT scans. The obtained data were pre-processed for classification using the CNN algorithm with 3-fold cross-validation, which resulted in an accuracy of 57%. This accuracy was increased by enlarging the dataset using data augmentation. A comparison of classical data augmentation, DCGAN, and ACGAN was carried out, and it was observed that DCGAN performed better and increased that the accuracy of CNN to 85.7%. Thus, the proposed approach provided improved accuracy and solved the problem of limited data samples [36].

The factors provided by ATP III clinical criteria updated in 2005 for Metabolic Syndrome (MetS) along with age and gender were used for creating a model for prediction of NAFLD. After pre-processing, the data of 40637 patients were divided into 66% and 34% for training and testing, respectively. The classification was performed by the J48 algorithm using hold-out cross-validation, and the AUROC of 0.731 was achieved [37]. A ML model was presented for detecting advanced liver fibrosis and cirrhosis in patients with HBV or HCV. 70% of the dataset was used for testing and 30% for testing using DT, RF, and GB classifiers using 10-fold cross-validation. Out of these algorithms, GB showed the highest AUROC=85% [38].

3.5. Related Studies in 2019:
A study [39] developed a ML method for the diagnosis of Hepatitis disease by using the Self-Organizing Map (SOM) technique for clustering tasks, CART for feature selection, and ensembles of Adaptive Neuro-Fuzzy Inference System (ANFIS) for predicting the hepatitis disease with 10-fold cross-validation. Its performance was compared with ANFIS, K-NN, NN, and SVM by applying these techniques on the same dataset without incorporating SOM and NIPALS for the classification task. The proposed approach obtained the ROC of 93.06% and outperformed other technologies.

A novel approach that trains the model using GA, normalization, feature selection, and 5-fold cross-validation was presented. About ten ML algorithms were used for classification, and among them, SVM(type C-SVC) using the proposed method yielded the highest accuracy of 88.49% [1]. From a study which compared Fuzzy logic, Fuzzy NN, and DT in classifying the BUPA liver data set, concluded that the Neuro-Fuzzy system, i.e., Fuzzy NN, achieved the highest accuracy of 91% [17]. A Deep Learning Radiomics of Elastography (DLRE) diagnose the liver fibrosis by using CNN classifier. 1330 two-dimensional shear wave elastography (2D-SWE) images were augmented using random transformations to increase the dataset. A DLRE ROI with the whole 2D-SWE was selected for the classification of 660 images using CNN. The proposed model was compared with 2D-SWE, APRI, and FIB-4, and the results proved that DLRE provided better classification with AUROC of 0.97 for cirrhosis and 0.98 for advanced fibrosis [40].

According to [12], among RF, REP Tree, J48, NB, and DT, the highest accuracy of 100% and a complete precision of 1 was achieved by RF in the classification of ILPD. Although, NB took much less time to build the model, but was less accurate than the RF classifier. The Heat map was used by [13] for checking correlated columns within the ILPD dataset. After classification, an accuracy of 75%, 74%, 69%, 64%, 62%, and 53% was achieved by LR, RF, DT, SVM, KNN, and NB, respectively. The analysis result showed that the LR classifier achieved the highest accuracy.

In a study conducted by [14], the ANN model was able to predict whether a person is a liver patient or not with 99.00% accuracy. The most effective factors in the liver patient were: Alkphos, Albumin and Globulin Ratio, Albumin among 11 features in ILPD. The predictive model produced by [15] has shown a prediction accuracy rate of 72.4% by using ILPD and confirmed that laboratory tests such as Age,
DB, SGPT, TP, ALB were significant predictors of the categories of liver disease. In a study, predictive models were built by using KDD, random sampling, and classification algorithms over ILPD. Among the LR, RF, AutoNeural & k-NN classifiers, K-NN showed the best outcome with an accuracy of 99.79% [16]. Online search log, along with the search time, can be organized by using the data creation method developed by [41]. This dataset can be used to predict the occurrence of liver cancer by applying a classification algorithm. After a comparison of various SVM kernels, it was observed that the RBF kernel showed the highest accuracy of 94.3%. Then, the SVM-RBF model was compared with other ML algorithms with 10-fold cross-validation, and it was concluded that the RF algorithm gained the highest accuracy of 97.5%. A study [42] produced a ML tool using Naïve Bayes and SVM classifiers because the SVM showed the highest accuracy among all the algorithms, and the NB classifier produced the results in the lowest period time. Thus, combining the two algorithms produced desired results in terms of accuracy, precision time taken to deliver the prediction.

Clinical and non-elastographic MRI radiomic features can be used for predicting liver stiffness [43]. The SVM model after cross-validation and using only either clinical data or radiomic data extracted from MRI of patients gives an accuracy of prediction as 77% and 70%, respectively. Whereas, when both the clinical and radiomic data are used together, then the obtained accuracy is 84%. The derived model was validated both internally and externally on 225 and 84 patients, respectively. The performance was evaluated based on the AUROC curve. Thus, the SVM model with a combination of clinical and T2-weighted radiomic data classifies the MR elastography derived liver stiffness [43]. A study compared various data mining algorithms for precise prediction of liver diseases. The models namely, RF, MLP neural network, Bayesian network, SVM, and PSOSVM, followed the ELTA approach and selected relevant features for the disease prediction. As a result, the PSOSVM model showed the best outcomes than other algorithms after being evaluated on a 10-fold cross-validation method. It also provided seven significant parameters for better accuracy [44].

Another study provided a model that uses a VP-Expert shell for the diagnosis of liver diseases. After the reports of blood tests, symptoms regarding the complication, and time are fed into the system with the help of block and Mockler charts, the expert system diagnoses the disease and delivers results [45]. In [46] discussed the automatic tumor detection techniques in the area of computational pathology, which can be defined as interpreting and analyzing medical images by using powers of computers and advanced technologies such as modern image analysis algorithms. The use of AI in pathology image segmentation helps in identifying ROI in much less time by extracting the parameters from segmented regions and relating these parameters with pathologic, genomic, and clinical settings.

Depth attenuation correction to the ultrasound images of 114 patients was used to reduce the noise within the images. The resulted images were analyzed, and five features were then extracted from the ROI, and based on these features, three models were trained using LR, SVM, and LR-SVM classifiers with 10-fold cross-validation. The evaluation showed that the hybrid model showed the highest accuracy [47].

The data of 539 HCC patients and 1043 non-HCC patients were used for establishing a ML model that predicts HCC. Linear classification of data was made by using the LR model, and non-linear classification was carried out using SVM with RBF kernel, gradient boosting, RF, NN, and deep learning. The variables were selected using the Alkaike information method. After splitting the dataset, the performance of all the classifiers was compared using the Wilcoxon rank-sum and chi-square test for quantitative or categorical values, respectively. The highest accuracy of 87.34% was shown by the gradient boosting (non-linear) model, and it also lowers the chances of misclassification by about half in comparison with a single marker. It was also noticed that a linear model is suitable for linearly-separable data, and a non-linear model is suitable for non-linearly separable data [48].

3.6. Related Studies in 2020:

A ML tool based on a GB algorithm was developed that predicts primary sclerosing cholangitis (PSC) in patients [49]. Previously used PSC diagnosis methods consider SAP as the most significant factor, but according to the developed approach PREsto (primary sclerosing cholangitis risk estimate tool)
views bilirubin as the most significant parameter for the prognosis. PREsto provides much better results than the currently used approach with C-Statistics of 0.90% and enables the patients to self-access using an online calculator [49]. A non-invasive ML algorithm for the diagnosis of NAFLD and NASH named LIVERFAst is a summation of anthropometric and serum biomarkers [50]. It is made up of neural networks and provides an easy to carry out, cheap, and convenient modification of the SAF (Steatosis, Activity, and Fibrosis) system that evaluates NASH and NAFLD precisely. The algorithm was developed by using the records of patients from different Southeast Asian countries. 2826 unique biomarkers were used for the training of the algorithm. Molecular changes, i.e., biomarkers, were produced by various stages of liver diseases, and these changes were detected in the serum, which indicated the scene of the liver disease.

A densely connected deep NN for the screening of liver diseases was presented in another study [51]. It is an end-to-end deep NN model with four hidden layers. The data used for screening comprised 15 features, i.e., age, gender, and 13 LFT parameters of 76,914 patients. The model used 5-fold cross-validation, and its prediction performance was compared with that of LR, RF, and conventional DNN by using the same dataset. The comparison showed that the proposed dense DNN outperformed other models. It was also seen that moderate over-fitting aids the screening process. The disease-free survival (DFS) of patients with HCC after liver resection (LR) can be predicted by using ML algorithms. The RF model is made by using clinical and other parameters and pre-processing the training data. The performance was evaluated on the test data, which was divided into low-risk and high-risk patients by the RF prediction. The recursive feature elimination analysis phase of RF marked the relevant predictors, which facilitated the outcome of the model. Thus, this model can be used to decide the type of treatment required for the HCC patient, i.e., LR for low-risk patients and LT for high-risk patients [52].

According to [53], a non-invasive way to assess HCC was provided by evaluating the records of 4423 chronic hepatitis C (CHC) patients to find the significant parameters for the diagnosis of HCC. From the dataset records of 293 HCV patients, 53 patients had HCC. This lot of patients were used for validation. The author developed various ML models, and out of these models, the alternating decision tree showed the best results with an area under the ROC curve of 95.6% and an accuracy of 99%. Also, the study shows that age, AFP, ALP, albumin, TB parameters were associated with the presence of HCC. A study [54] presented a non-invasive method that classifies the patients as either healthy or having pulmonary hypertension, cirrhosis, HCC, or CRLM based on breath-based metabolites, i.e., volatile organic compounds (VOCs). The RF ML model was trained using 22 VOCs from the breath of 296 patients. It was observed that two VOCs, namely acetaldehyde and acetone, were higher in numbers in cirrhosis and HCC patients in comparison to healthy patients. Thus, this model helps in the screening of diseases and tumours within the liver with greater accuracy and sensitivity.

The multi-parametric MRI can be used to detect the status of mVI for HCC before operation or treatment by using ML classifiers [55]. Radiomic features were extracted by manually segmenting the HCC area. The ML classifiers were built on MRI sequences of features derived from either tumour only region or peritumoral region only or the combination of the two, and the achieved accuracy by the classifiers was 86.69%, 84.62%, and 84.19% respectively. Thus, parameters from the tumour-only region showed the highest accuracy. Therefore, it was observed that different sequences provide complementary information. In contrast, a combination of sequences offers high accuracy and points to the usefulness of a multi-sequence approach in predicting microvascular invasion (mVI) [55]. Transfer learning was used for the classification of pathological-images to diagnose various diseases and tumours within the liver. Transfer-learning based image recognition is more promising for the difficult to collect medical images, and the proposed model showed universality in classifying the images with HCC. Its performance is comparable to that of human experts. Its performance was further tested for two other tumours, namely CRC and BIDC. These tumours were also efficiently classified by the developed model [56].

In [57], a DLNN for ICC (Intrahepatic Cholangiocarcinoma) was developed by applying concepts of deep learning on four pathologic, six serologic, and two etiologic features of 1421 patients. These 12
features were fed as input to the NN, which was then validated with records of 234 patients. The results showed that the developed model was more accurate than the AJCC stage and COX model. A study [58] used imbalanced primary data to create a ML model that could predict the postoperative survival of patients after liver transplantation. Rules from MELD (Model for End-stage Liver Disease) score were used for data pre-processing, and feature selection and classification were performed using the RF algorithm. The proposed model was compared with GB, LR, and DT using 10-fold cross-validation. The experimental results indicated that RF produced greater accuracy (77.1%) than other algorithms. It was concluded that feature selection using RF (international normalized ratio (INR), lymphocytes, prothrombin time (PT), platelets, white blood cell (WBC), Magnesium (Mg), Sodium (Na), age, and BMI) was more beneficial than MELD. Another study [59] used BPNN, RBFN, and KNN for designing a model for diagnosing and predicting the severity of hepatitis by dividing the dataset into equal portions for training and testing. The results concluded that KNN outperformed the other two algorithms with the accuracy of 99.33%. Also, BPNN and RBFN required more training data and hyperparameter settings to achieve comparable accuracy. [60] investigated the impact of gut microbial profiles in the diagnosis of liver disease (LD) and alcoholic liver disease (ALD). The experimental results showed that AUROCs of 0.834 for LD and 0.956 for ALD was achieved by augmenting the conventional risk factors of the GB classifier with microbiome predictors. In [61] developed various models using multi-omics and clinical records of 1514 patients from which the required features were selected on the least absolute shrinkage and selection operator (LASSO). The models were created using the liver fat content derived from MRI scans of patients along with either clinical parameters, omics parameters, and a combination of both by applying RF analysis with 5-fold cross-validation. A total of 18 models that predicts early-stage NAFLD were developed in which the models created with input data of both clinical and omics parameters produced the highest accuracy with AUROC curve=0.84.

A study [62] investigated the benefit of using a ML-based radiomic model in the prediction of metachronous metastases in colorectal cancer (CRC) patients. The dataset containing CT scans of 91 patients was divided into two groups of patients based on having or not having metachronous liver metastases. Then the data was analyzed and segmented, which extracted 1767 radiomic features for each patient, and inter-correlated element were removed based on the Kruskal-Wallis test. Three models were created using clinical data, radiomic data, and a combination of both clinical and radiomic data with RF classifier along with wrapper feature selection and Bayesian optimization. The AUC for the three models is 86% (with radiomic data), 71% (with clinical data), 86% (with a combination of both data). Thus, radiomic analysis of data provides significant features, i.e., biomarkers for the identification of high-risk patients in developing colorectal liver metastases.

Tumour cells release DNAs in the blood, which are called circulating tumour DNAs (ctDNAs). These carry tumour-specific genomic aberrations. [63] used the SCNA (somatic copy number aberrations) present in the ctDNA to create a weighted RF driver (wRF) ML model that detects HCC in its early stage. Plasma samples 384 hepatitis B virus (HBV) patients were used to profile SCNAS by using low-depth whole-genome sequencing. Out of 384 patients, a discovery set with 209 patients was used to train the model along with information retrieved from extensive external data. The obtained overall AUC was 0.893, with 0.874 for early-stage cancer and 0.933 for advanced-stage cancer. The diagnosis capability of the model was also evaluated on two validation sets with 76 and 99 patients who had early-stage disease, and for them, the achieved AUC was 0.920 and 0.812, respectively.

LDA, GA, and SVM algorithms can be combined for creating a hybrid diagnostic model that provides an accuracy of 90.30%. Here, LDA provided reduced dimensions in data, SVM provided classification with 5-fold cross-validation, and GA optimized the SVM model [64]. A novel and generic computational framework, named DeepProg, which processes multiple types of omics data sets with a combination of deep-learning (autoencoder) and machine-learning algorithms, specifically for survival prediction. DeepProg showed better predictive accuracies when compared to SNF based multi-omics integration method and the baseline Cox-PH method [65].
The computational model proposed by [66] can discriminate HCC and non-cancerous cells from CwoHCC (cirrhosis without HCC) even with inaccurately sampled fewer biopsy specimens. The accuracy of HCC diagnosis was improved by obtaining 11-gene-pair from the 19-gene-pair using mRMR (maximum redundancy minimum relevance). In [67], an ANN model was produced for predicting 1-year progression-free survival (PFS) in HCC patients which showed (AUROC) of 0.866 in HCC patients, which was higher than predicted by TNM, BCLC, Okuda, CLIP, CUP, JIS, and ALBI scores (p < 0.0001). A study [68] concluded that KNN (51.06%, R) and random forest (54.56%, Python) in multi and binary class labels respectively among NB, KNN, NN, RF and SVM with 10-fold cross-validation. Feature selection had no impact on accuracy. Another study [69] proposed a trainable multi-contrast windowing method for the optimal choice of contrast windows for CT segmentation based on deep learning. It then analyzed its impact on the segmentation network that measures tumours within the liver. The proposed window produced better outcomes as compared to the window with fixed parameters defined by radiologists. Thus, it can be concluded that pre-defined windowing parameters are not the best parameters for performing segmentation and a single-window range is not sufficient. Thus, it is optimum to use a window with multiple contrasts, which is trainable by the machine itself. An IncRNA signature could predict the survival of HCC patients [70].

The clinical data of 180 patients with HCC were used to train the model, which could divide the list of HCC patients into two groups with varied prognosis. The prediction capability of the IncRNA signature was compared with that of TNM (tumour/node/metastasis) stage based on pROC, timeROC, survival, and randomforestSRC. The analysis showed that the four-IncRNA signature gave more efficient performance than the TMA stage in prediction. According to COX analysis, a four-IncRNA signature could independently predict the survival of HCC patients.

A summary of the studies discussed above is presented in Table 1. as follows:

| Year | References | Objective | Main Accomplishments |
|------|------------|-----------|----------------------|
| 2015 | [4] [6] [20] [7] [21] | Using ML and Elastography techniques for Prediction and Detection of liver diseases. Developing GUI that aids in self-assessment. | Used datasets with records of clinical tests of patients and compared various ML algorithms using the significant features from the data for the prediction of various liver diseases. The concepts of elastography were also explored. Different ML models for diagnoses were developed. |
| 2016 | [8] [22] [23] [24] [25] [26] [27] [28] | Identify significant features from the data. Prediction and diagnosis by ML models using clinical and ultrasound tests. | Standard significant features were provided by using data mining techniques on the clinical records for training various ML classifiers. Ultrasound images were also used for extracting useful information. Efficient models were produced for the timely diagnosis of diseases. |
| 2017 | [9] [5] [10] [29] [30] [31] [32] | Improve the diagnostic accuracy of models using boosting and optimization techniques. Provide novel, efficient detection models using gut-microbiome samples. | The detection accuracy of classifiers increased by using boosting techniques. Introduced a novel approach that uses learning with abstention. Various ML algorithms were trained using small and large-scale clinical and imaging data. Their outcomes provided efficient models. The features from the gut-microbiome were also explored. |
| 2018 | [33] [11] [19] [18] [36] [37] [38] [35] | Propose a data augmentation method. Introduce models for predicting disease progression. Expand the concept of elastography using DL techniques. | A data augmentation method named DCGAN was presented for solving the problem of a limited dataset. DLRE showed comparable performance to that of liver biopsy and provided features that are strictly related to liver fibrosis. Various improved ML models with significant features for predicting disease progression were presented. |
Detection of disease progression. Compare data mining algorithms for prediction using imbalanced data. Explore the concept of computational pathology and the impact of clinical tests on model creation. Develop a liver disease diagnostic model using ML concepts.

An expert system for quick diagnosis of diseases was presented. Imaging data were extensively used for gaining better insights into the application of computing powers in disease diagnosis. Various types of features selection methods were used by different studies that provided efficient features related to diseases. An image segmentation technique using AI was proposed for quickly locating the ROI. Various indicators other than blood tests were investigated for achieving improved accuracy.

Predict the disease-free survival after treatment. Create efficient models to predict HCC and its reoccurrence. Predict mVI using ML concepts. Classify patients using the breath-based metabolites. Apply DL concepts to pathological images for predicting HCC.

Accurate models using dimensionality reduction and optimization were presented. Genomic, pathologic, multi-omics, and serum-based parameters were investigated for creating ML models. Various models were also validated on extensive primary datasets. Conventional features were augmented with other significant features extracted from microbiome samples. Models were presented for detecting the progression-free survival of patients.

| Year | References | Summary |
|------|------------|---------|
| 2019 | [39][1][17][40] [12][13][14][15][16] [41][42][43][44] [45][47][48] | Detection of disease progression. Compare data mining algorithms for prediction using imbalanced data. Explore the concept of computational pathology and the impact of clinical tests on model creation. Develop a liver disease diagnostic model using ML concepts. |
| 2020 | [49][50][51][52] [53][54][55][56] [61][67][68][69] [70][62][63][64] [66][65][57][58] [59][60] | Predict the disease-free survival after treatment. Create efficient models to predict HCC and its reoccurrence. Predict mVI using ML concepts. Classify patients using the breath-based metabolites. Apply DL concepts to pathological images for predicting HCC. |

4. Research Gaps and Future Scopes:

It is observed that some approaches do not show efficient outcomes when used on a large dataset and classifies a sample of datasets with greater accuracy than the remaining dataset. Different methods have a varied range of efficiency and sensitivity. It is also seen that specific approaches do not show the same level of accuracy when applied to real-time data. Collecting large-scale medical data for research and selecting the most significant features are the challenging aspects of using ML concepts for liver disease detection. These problems can be solved by carrying out large scale studies on multiple centres with emphasizing the collection, segmentation, and pre-processing of data.

The proposed approach in [1,9,12,17,18,32,42,49,58] should be tested on a large independent test set to prove its robustness. Combining different ML algorithms and implementing them on large-scale secondary and real-time data from various locations will be beneficial in proving their robustness ([4,6,67,68,7,37,38,40,53,55,57,61]). In a study [5], only two classifiers were compared, and the features were ranked according to the gain ratio. More accurate feature selection methods are required to prove the reliability of the model. Moreover, an extensive comparison of classifiers is important to prove the robustness of the selected algorithm.

The capabilities of the ML models can be increased from disease prediction to disease detection and providing treatment options [14,45,56,60]. In [15], the objective is the prediction of the reoccurrence of the disease. However, there is no measure of detecting whether the condition had reoccurred or not. In this regard, continuous monitoring of the patients’ records is required though an online portal by increasing the aim from just prediction to actual progression detection.

Most of the studies have used the random-sampling method for dataset division for training and testing. However, there is a lack of studies comparing the results of different sampling techniques over the dataset. Moreover, it can also be possible that different sampling methods suit different situations and datasets. Relying on only a specific way limits the possibilities of further enhancements [16].

A lot of research has been done in forecasting liver diseases, whereas the studies focussing on the progress of the disease are very few. High emphasis should be laid on detecting the stage of the illness so that the physicians can directly instruct the treatment to patients without wasting time in costly tests [8]. Various studies including [10] used ILPD. A comparison of all these studies using the same tool will be beneficial in delivering the best performing model. A comparison of feature selection techniques would be advantageous in extracting significant and disease-related parameters ([11,28]). The use of data mining techniques can provide hidden aspects in genetic profiles, blood test results, bone density, heart rate, blood pressure, diabetic status, urine samples and other related data. The impact of these
parameters should be checked on various liver diseases, and appropriate classifiers should be used to achieve an efficient automatic diagnostic system [19,51,62,63]. A small dataset for training and testing of the model was used by [20]. It can be increased by collecting primary or secondary data, or by augmenting the available data. The implementation of the proposed model on other data types, e.g., CT scans, MRI, and ultrasound images, is also required. The work of [21,34,64] can be incorporated with other ML models for inventing improved diagnostic techniques.

In [22,59], the proposed ML classifiers should have been compared with different classifiers in the prediction of the liver fibrosis stage. Moreover, the significance of other biomarkers can be calculated and added to increase the training set and to achieve improved accuracy. In [23], the parameters from the blood test were not considered for model creation. Thus, combining the blood test features along with the features already used in the study can provide new directions. Moreover, a comparison of the produced model with the new model, including blood-related features, can offer a promising approach for disease diagnosis.

A large dataset should be used for training and testing the models because the models created with limited data lack trustworthiness [24,26,54]. The studies ([25,27]) can be extended for multi-class classification and predicting liver disease stages. The significant features specified in [29,31,32,65] can be combined with other clinical, serum-based, and imaging-based features to provide a more generic model. A relatively small dataset has been used for model creation in [30,52]. There is a need to test the proposed approach on a large unseen dataset to prove its efficiency. Moreover, it can be combined with other ML models to provide multi-class classification. The scope of approach can be increased to provide the details regarding the stage of diseases. An extensive study was carried out by using extensive training and test set for model creation [35]. The methodology used can be applied to other types of medical records of patients, e.g., imaging data. Also, the proposed approach should be compared with other proposed models using the same baseline dataset. It would yield the most efficient model in the diagnosis of liver diseases.

The technique proposed by [36] uses separate GANs for each lesion class, which makes the training process very complicated. The problem can be solved by using a GAN framework that supports multiple lesions. Also, the use of unlabelled data for training the model will provide good quality lesion samples. Various DT algorithms have been compared for classification using the same dataset. These should also be compared with other ML algorithms.

The investigation of the effects of clustering, processing, and ensemble learning along with ML methods on the accuracy can be beneficial in producing reliable predictive systems that can automatically classify the unseen data with high efficiency [39]. A publicly accessible web-server implementation of the models would assist physicians and patients across the globe to get a timely diagnosis. It will also help in evaluating the models on a large-scale real-time data, and thereby making modifications if required according to the feedback of physicians, will yield a robust diagnostic model [41,48,54]. The concepts of deep learning can be combined with the ML algorithms in the development of predictive models [43]. This will improve the outcomes in one or the other way by introducing new areas of research. The is a lack of studies on using meta-heuristic methods for achieving optimized classification such as Genetic methods and the Ant colony optimization method [44].

Various aspects should be investigated before processing so that more critical issues are prioritized for achieving better quality and dependable models. Some of these areas include tissue transmission coefficient, imaging frequency, equipment manufacturer, testing equipment, data sampling method, segmentation method, cross-validation technique, feature-selection method, data mining tool, model creation tool, etc. [47]. The selection of accurate and removal of less significant biomarkers is the most challenging phase of model creation. High emphasis must be laid on this step by using appropriate selection methods. Incorporating more striking features will ultimately improve the accuracy of the system [50,66]. There is a need for an approach that provides enhanced segmentation of pathological images, which will allow faster segmentation and reduced sample size [55]. The clinical records can be integrated with other information, e.g., genomics, molecular, genetics, environmental, past operational
data, lifestyle, etc. [63]. Investigating important texture features related to lesions is vital for producing reliable classification models [69].

5. Conclusion:
ML approach is widely used in analyzing clinical information, genetic records, and medical images of patients. Various studies have proved the unmatchable potential of data mining and ML tools in the medical domain. These tools can discover hidden significant predictive parameters from medical datasets that provide early prediction and diagnosis of diseases. In this study, we tried to find the research gap of the various researches that use ML concepts in liver disease diagnosis. Moreover, the future scope has also been mentioned regarding the same. From the above review, it is evident that ML techniques are highly promising in diagnosing liver diseases. But further data proving its validity and efficiency is required for its constant use by physicians.

Acknowledgement:
The authors wish to thank Dr Kunal Pal Singh (MBBS), for providing necessary information related to liver diseases and their treatments.

References
[1] Ksiażek W, Abdar M, Acharya U R and Pławiak P 2019 A novel machine learning approach for early detection of hepatocellular carcinoma patients Cogn. Syst. Res. 54 116–27
[2] Reitermanov Z 2010 Data Splitting 31–6
[3] Dua, Dheeru and Graff C 2017 {UCI} Machine Learning Repository Univ. California, Irvine, Sch. Inf. Comput. Sci.
[4] Abdar M 2015 A Survey and Compare the Performance of IBM SPSS Modeler and Rapid Miner Software for Predicting Liver Disease by Using Various Data Mining AlgorithmsCumhur. Sci. J. 36 3230–41
[5] Abdar M, Zomorodi-Moghadam M, Das R and Ting I H 2017 Performance analysis of classification algorithms on early detection of liver disease Expert Syst. Appl. 67 239–51
[6] Dr. S. Vijayarani1 M S D 2015 Liver Disease Prediction using SVM and Naïve Bayes Algorithms Int. J. Sci. Eng. Technol. Res. 4 816–20
[7] Nagaraj K and Sridhar A NeuroSVM: A Graphical User Interface for Identification of Liver Patients Kalyan Nagaraj 1* and Amulyashree Sridhar 2 1* 9
[8] Farokhzad M R and Ebruhimi L 2016 A Novel Adaptive Neuro Fuzzy Inference System for the Diagnosis of Liver Disease Int. J. Acad. Res. Comput. Eng. 1 61–6
[9] Hassoon M, Kouhi M S, Zomorodi-Moghadam M and Abdar M 2017 Rule Optimization of Boosted C5.0 Classification Using Genetic Algorithm for Liver disease Prediction 2017 Int. Conf. Comput. Appl. ICCA 2017 299–305
[10] Pasha M and Fatima M 2017 Comparative Analysis of Meta Learning Algorithms for Liver Disease Detection J. Softw. 12 923–33
[11] Banu Priya M, Laura Juliet P and Tamilselvi P R 2018 Performance Analysis of Liver Disease Prediction Using Machine Learning Algorithms Int. Res. J. Eng. Technol. 5 206–11
[12] Ramaiah M, Baranwal P, Shastri S B, Vanitha M and Vanmathi C 2019 Analytical Comparison of Machine Learning Techniques for Liver Dataset 2019 Innov. Power Adv. Comput. Technol. i-PACT 2019 9–13
[13] Rahman A K M S, Javed Mehed Shamrat F M, Tasnim Z, Roy J and Hossain S A 2019 A comparative study on liver disease prediction using supervised machine learning algorithms
[14] Musleh M M, Alajrami E, Khalil A J, Abu-nasser B S, Barhoom A M and Abu-naser S S 2019 Predicting Liver Patients using Artificial Neural Network Int. J. Sci. Technol. Res. 8 419–22

[15] Abdalradaa A S, Yahya O H, Alaidi A H M, Hussein N A, Alrikabi H T and Al-Quraishi T 2019 A predictive model for liver disease progression based on logistic regression algorithm Period. Eng. Nat. Sci. 7 1255–64

[16] Arbain A N and Balakrishnan B Y P International Journal of Data Science and Advanced Analytics

[17] Singh A K 2019 A Comparative Study on Disease Classification using Machine Learning Algorithms SSRN Electron. J. 114 1–10

[18] Haque M R, Islam M M, Iqbal H, Reza M S and Hasan M K 2018 Performance Evaluation of Random Forests and Artificial Neural Networks for the Classification of Liver Disorder Int. Conf. Comput. Commun. Chem. Mater. Electron. Eng. IC4ME2 2018 1–5

[19] Arshad I, Dutta C, Choudhury T and Thakral A 2018 Liver Disease Detection Due to Excessive Alcoholism Using Data Mining Techniques Proc. 2018 Int. Conf. Adv. Comput. Commun. Eng. ICACCE 2018 163–8

[20] Mala K, Sadasivam V and Alagappan S 2015 Neural network based texture analysis of CT images for fatty and cirrhosis liver classification Appl. Soft Comput. J. 32 80–6

[21] Samir A E, Dhyani M, Vij A, Bhan A K, Halpern E F, Méndez-Navarro J, Corey K E and Chung R T 2015 Shear-Wave elastography for the estimation of liver fibrosis in chronic liver disease: Determining accuracy and ideal site for measurement Radiology 274 888–96

[22] Ayeldeen H, Shaker O, Ayeldeen G and Anwar K M 2016 Prediction of liver fibrosis stages by machine learning model: A decision tree approach Proc. 2015 IEEE World Conf. Complex Syst. WCCS 2015

[23] Rau H H, Hsu C Y, Lin Y A, Atique S, Fuad A, Wei L M and Hsu M H 2016 Development of a web-based liver cancer prediction model for type II diabetes patients by using an artificial neural network Comput. Methods Programs Biomed. 125 58–65

[24] Acharya U R, Fujita H, Bhat S, Raghavendra U, Gudigar A, Molinari F, Vijayananthan A and Hoong Ng K 2016 Decision support system for fatty liver disease using GIST descriptors extracted from ultrasound images Inf. Fusion 29 32–9

[25] Baitharu T R and Pani S K 2016 Analysis of Data Mining Techniques for Healthcare Decision Support System Using Liver Disorder Dataset Procedia Comput. Sci. 85 862–70

[26] SAI 2016 Proceedings of 2016 SAI Intelligent Systems Conference (IntelliSys) Intellisyss 1

[27] Bashir S, Qamar U, Khan F H and Naseem L 2016 HMV: A medical decision support framework using multi-layer classifiers for disease prediction J. Comput. Sci. 13 10–25

[28] Saba L, Dey N, Ashour A S, Samanta S, Nath S S, Chakraborty S, Sanches J, Kumar D, Marinho R T and Suri J S 2016 Automated stratification of liver disease in ultrasound: An online accurate feature classification paradigm Comput. Methods Programs Biomed. 130 118–34

[29] Loomba R, Seguritan V, Li W, Long T, Klitgord N, Bhatt A, Dulai P S, Caussy C, Bettencourt R, Highlander S K, Jones M B, Sirlin C B, Schnabl B, Brinkac L, Schork N, Chen C H, Brenner D A, Biggs W, Yooseph S, Venter J C and Nelson K E 2017 Gut Microbiome-Based Metagenomic Signature for Non-invasive Detection of Advanced Fibrosis in Human Nonalcoholic Fatty Liver Disease Cell Metab. 25 1054-1062.e5
[30] Hamid K, Asif A, Abbasi W, Sabih D and Minhas F U A A 2017 Machine Learning with Abstention for Automated Liver Disease Diagnosis Proc. - 2017 Int. Conf. Front. Inf. Technol. FIT 2017 2017-Janua 356–61

[31] Chen Y, Luo Y, Huang W, Hu D, Zheng R qin, Cong S zhen, Meng F kun, Yang H, Lin H jun, Sun Y, Wang X yan, Wu T, Ren J, Pei S F, Zheng Y, He Y, Hu Y, Yang N and Yan H 2017 Machine-learning-based classification of real-time tissue elastography for hepatic fibrosis in patients with chronic hepatitis B Comput. Biol. Med. 89 18–23

[32] Yip T C F, Ma A J, Wong V W S, Tse Y K, Chan H L Y, Yuen P C and Wong G L H 2017 Laboratory parameter-based machine learning model for excluding non-alcoholic fatty liver disease (NAFLD) in the general population Aliment. Pharmacol. Ther. 46 447–56

[33] Abdar M, Yen N Y and Hung J C S 2018 Improving the Diagnosis of Liver Disease Using Multipler Preceptor Neural Network and Boosted Decision Trees J. Med. Biol. Eng. 38 953–65

[34] Nahar N and Ara F 2018 Liver Disease Prediction by Using Different Decision Tree Techniques Int. J. Data Min. Knowl. Manag. Process 8 01–9

[35] Hashem S, Esmat G, Elakel W, Habashy S, Raouf S A, ElHefnawi M, Eladawy M and ElHefnawi M 2018 Comparison of Machine Learning Approaches for Prediction of Advanced Liver Fibrosis in Chronic Hepatitis C Patients IEEE/ACM Trans. Comput. Biol. Bioinforma. 15 861–8

[36] Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J and Greenspan H 2018 GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification Neurocomputing 321 321–31

[37] Perveen S, Shahbaz M, Keshavjee K and Guergachi A 2018 A Systematic Machine Learning Based Approach for the Diagnosis of Non-Alcoholic Fatty Liver Disease Risk and Progression Sci. Rep. 8 1–12

[38] Wei R, Wang J, Wang X, Xie G, Wang Y, Zhang H, Peng C Y, Rajani C, Kwee S, Liu P and Jia W 2018 Clinical prediction of HBV and HCV related hepatic fibrosis using machine learning EbioMedicine 35 124–32

[39] Nilashi M, Ahmadi H, Shahmoradi L, Ibrahim O and Akbari E 2019 A predictive method for hepatitis disease diagnosis using ensembles of neuro-fuzzy technique J. Infect. Public Health 12 13–20

[40] Wang K, Lu X, Zhou H, Gao Y, Zheng J, Tong M, Wu C, Liu C, Huang L, Jiang T, Meng F, Lu Y, Ai H, Xie X Y, Yin L P, Liang P, Tian J and Zheng R 2019 Deep learning radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: A prospective multicentre study Gut 68 729–41

[41] Joarder T A, Ahmed B and Sattar A S 2019 Detecting Liver Cancer from Online Search Logs 2019 4th Int. Conf. Electr. Inf. Commun. Technol. EICT 2019 1–5

[42] Durai V Liver disease prediction using machine learning 5 1584–8

[43] He L, Li H, Dudley J A, Maloney T C, Brady S L, Somasundaram E, Trout A T and Dillman J R 2019 Machine learning prediction of liver stiffness using clinical and T2-Weighted MRI radiomic data Am. J. Roentgenol. 213 592–601

[44] Joloudari J H, Saadatfar H, Dehzangi A and Shamshirband S 2019 Computer-aided decision-making for predicting liver disease using PSO-based optimized SVM with feature selection Informatics Med. Unlocked 17 100255

[45] Mirmozaffari M 2019 Developing an Expert System for Diagnosing Liver Diseases Eur. J.
Engineering I 2019 www.kosse.or.kr 15 72–8

Chen C I, Chen T B, Lu N H, Du W C, Liang C Y, Liu K I, Hsu S Y, Lin L W and Huang Y H 2019 Classification for liver ultrasound tomography by posterior attenuation correction with a phantom study Proc. Inst. Mech. Eng. Part H J. Eng. Med. 233 1100–12

Sato M, Morimoto K, Kajihara S, Tateishi R, Shina S, Koike K and Yatomi Y 2019 Machine-learning Approach for the Development of a Novel Predictive Model for the Diagnosis of Hepatocellular Carcinoma Sci. Rep. 9 1–7

Shung D L and Assis D N 2020 Machine Learning in a Complex Disease: PREsTo Improves the Prognostication of Primary Sclerosing Cholangitis Hepatology 71 8–10

Aravind A, Bahirvani A G, Quiambao R and Gonzalo T 2020 Machine Learning Technology for Evaluation of Liver Fibrosis, Inflammation Activity and Steatosis (LIVERFASt\&\sup&T;M\&\lt;/sup\&\gt;) J. Intell. Learn. Syst. Appl. 12 31–49

Yao Z, Li J, Guan Z, Ye Y and Chen Y 2020 Liver disease screening based on densely connected deep neural networks Neural Networks 123 299–304

Schoenberg M B, Bucher J N, Koch D, Börner N, Hesse S, De Toni E N, Seidensticker M, Angele M K, Klein C, Bazhin A V., Werner J and Guba M O 2020 A novel machine learning algorithm to predict disease free survival after resection of hepatocellular carcinoma Ann. Transl. Med. 8 434–434

Hashem S, ElHefnawi M, Habashy S, El-Adawy M, Esmat G, Elake W, Abdelazziz A O, Nabeel M M, Abdelmukoud A H, Elbaz T M and Shousha H I 2020 Machine Learning Prediction Models for Diagnosing Hepatocellular Carcinoma with HCV-related Chronic Liver Disease Comput. Methods Programs Biomed. 196

Miller-Atkins G, Acevedo-Moreno L, Grove D, Dweik R A, Tonelli A R, Brown J M, Allende D S, Aucejo F and Rotroff D M 2020 Breath Metabolomics Provides an Accurate and Noninvasive Approach for Screening Cirrhosis, Primary, and Secondary Liver Tumors Hepatol. Commun. 4 1041–55

Nebbia G, Zhang Q, Arefan D, Zhao X and Wu S 2020 Pre-operative Microvascular Invasion Prediction Using Multi-parametric Liver MRI Radiomics J. Digit. Imaging

Chen W-M, Fu M, Zhang C-J, Xing Q-Q, Zhou F, Lin M-J, Dong X, Zheng Q-Z, Hong M-Z and Pan J-S 2020 Deep Learning-Based Universal Expert-Level Recognizing Pathological Images of Hepatocellular Carcinoma and beyond

Jeong S, Ge Y, Chen J, Gao Q, Luo G, Zheng B, Sha M, Shen F, Cheng Q, Sui C, Liu J, Wang H, Xia Q and Chen L 2020 Latent Risk Intrahepatic Cholangiocarcinoma Susceptible to Adjuvant Treatment After Resection: A Clinical Deep Learning Approach Front. Oncol. 10

Liu C L, Soong R S, Lee W C, Jiang G W and Lin Y C 2020 Predicting Short-term Survival after Liver Transplantation using Machine Learning Sci. Rep. 10 1–10

Muhi S H, Abdullah H N and Abd B H 2020 Modeling for predicting the severity of hepatitis based on artificial neural networks Int. J. Intell. Eng. Syst. 13 154–66

Liu Y, Meric G, Havulina A S, Teo S M, Ruuskana M, Sanders J, Zhu Q, Tripathi A, Verspoor K, Cheng S, Jain M, Jousilahti P, Vazquez-Baeza Y, Loomba R, Lahti L, Nizanen T, Salomaa V, Knight R and Inouye M 2020 Early prediction of liver disease using conventional risk factors and gut microbiome-augmented gradient boosting medRxiv 2020.06.24.20138933

Atabaki-Pasdar N, Ohlsson M, Viñuela A, Frau F, Pomares-Millan H, Haid M, Jones A G, Thomas E L, Koivula R W, Kurbasic A, Mutie P M, Fitipaldi H, Fernandez J, Dawed A Y,
Giordano G N, Forgie I M, McDonald T J, Rutters F, Cederberg H, Chabanova E, Dale M, Masi F De, Thomas C E, Allin K H, Hansen T H, Heggie A, Hong M G, Elders P J M, Kennedy G, Kokkola T, Pedersen H K, Mahajan A, McEvoy D, Pattou F, Raverdy V, Häussler R S, Sharma S, Thomsen H S, Vangipurapu J, Vestergaard H, ’t Hart L M, Adamski J, Musholt P B, Brage S, Brunak S, Dermitzakis E, Frost G, Hansen T, Laakso M, Pedersen O, Ridderstråle M, Rutten T, Hattersley A T, Walker M, Beulens J W J, Mari A, Schwenk J M, Gupta R, McCarthy M I, Pearson E R, Bell J D, Pavo I and Franks P W 2020 Predicting and elucidating the etiology of fatty liver disease: A machine learning modeling and validation study in the IMI DIRECT cohorts PLoS Med. 17 e1003149

[62] Taghavi M, Trebeschi S, Simões R, Meek D B, Beckers R C J, Lambregts D M J, Verhoef C, Houwers J B, van der Heide U A, Beets-Tan R G H and Maas M 2020 Machine learning-based analysis of CT radiomics model for prediction of colorectal metachronous liver metastases Abdom. Radiol.

[63] Tao K, Bian Z, Zhang Q, Guo X, Yin C, Wang Y, Zhou K, Wan S, Shi M, Bao D, Yang C and Xing J 2020 Machine learning-based genome-wide interrogation of somatic copy number aberrations in circulating tumor DNA for early detection of hepatocellular carcinoma EBioMedicine 56

[64] Ali L, Wajahat I, Amiri Golilarz N, Keshtkar F and Bukhari S A C 2020 LDA–GA–SVM: improved hepatocellular carcinoma prediction through dimensionality reduction and genetically optimized support vector machine Neural Comput. Appl. 2

[65] Poirion O, Chaudhary K, Huang S and Garmire L X 2020 DeepProg : an ensemble of deep-learning and machine- learning models for prognosis prediction using multi-omics

[66] Zhang Z M, Tan J X, Wang F, Dao F Y, Zhang Z Y and Lin H 2020 Early Diagnosis of Hepatocellular Carcinoma Using Machine Learning Method Front. Bioeng. Biotechnol. 8 1–9

[67] Liu X, Hou Y, Wang X, Yu L, Wang X, Jiang L and Yang Z 2020 Machine learning-based development and validation of a scoring system for progression-free survival in liver cancer Hepatol. Int. 14 567–76

[68] Nandipati S C R, Xinying C and Wah K K 2020 Hepatitis C Virus ( HCV ) Prediction by Machine Learning Techniques 4 89–100

[69] Kwon J and Choi K 2020 Trainable multi-contrast windowing for liver CT segmentation Proc. - 2020 IEEE Int. Conf. Big Data Smart Comput. BigComp 2020 169–72

[70] Jiang H, Zhao L, Chen Y and Sun L 2020 A four-long noncoding RNA signature predicts survival of hepatocellular carcinoma patients J. Clin. Lab. Anal. 1–9