International Journal of Development Mathematics Vol 2 Issue 2 Page No. 278 - 297

INTERNATIONAL JOURNAL OF DEVELOPMENT MATHEMATICS

ISSN: 3026-8656 (Print) | 3026-8699 (Online)

journal homepage: https://ijdm.org.ng/index.php/Journals

Predictive Modeling of Hepatitis using Machine Learning Algorithms and Mathematical Formulations

Victor C. Chiawa[a] and Kingsley I. Chibueze[b*]

[a]*Department of Education Mathematics, Abia State University Uturu, Nigeria*

[b]*Department of Computer Science & Mathematics, Godfrey Okoye University, Nigeria*

ARTICLE INFO

Article history:
Received 10 February 2025
Received in revised form 08 May 2025
Accepted 21 May 2025

Keywords:
Hepatitis Prediction, Healthcare AI, Early Diagnosis, SMOTE, Machine Learning.

MSC 2020 Subject classification:
93A30

ABSTRACT

Hepatitis continues to be a major health concern for the world, mostly impacting those who are most at risk, such as infants and pregnant women. This research aims to develop a machine learning model for early hepatitis detection and reduce the number of people affected and killed by the disease. Characterization of Hepatitis B (HBV), Hepatitis C (HCV), and Hepatitis E (HEV) was done using information from a range of demographics, drawn from healthcare facilities and online repositories. Advanced data preprocessing techniques including feature selection, imputation, and normalization were employed to prepare a robust dataset. The machine learning model and the two gradient boosting algorithms of Random Forest, AdaBoost, and XGBoost were trained and tested. The models' performance was assessed using metrics such as Accuracy, Precision, Recall, F1 Score, and Confusion Matrix. XGBoost performed best according to all metrics, giving an accuracy of 0.96, precision of 0.95, recall of 0.94, and F1 score of 0.94. AdaBoost achieved an accuracy of 0.90, precision of 0.90, recall of 0.88, and F1 score of 0.89. Random Forest performed the worst among the others, giving results of 0.85 accuracy, 0.85 precision, 0.84 recall, and 0.84 F1 score. Another check on XGBoost's performance was done with a ten-fold cross-validation approach. It gave excellent results, reaching an average accuracy of 0.8532 along with precision of 0.8503, recall of 0.8395, and an F1 score of 0.8450. When matched with other models, the proposed model outperformed them and can be applied in a clinical setting.

1. Introduction

Hepatitis is a life-threatening disease that leads to swelling and deterioration of liver cells, known as hepatocytes, compromising liver function. Hepatitis can be classified into two types: chronic and acute. Acute hepatitis involves relatively minor damage to hepatocytes and typically lasts 12 months, whereas chronic hepatitis causes swelling lasting more than 6 months and results in more severe liver damage (Ansari, 2011). Globally, hepatitis has impacted around 350 million people, with numbers rising daily. The five primary hepatitis viruses are types A, B, C, D, and E, all of which significantly affect public wellbeing by causing potential outbreaks, fatalities, and illness. Types B and C are the leading causes of chronic diseases, cancer, and liver cirrhosis in millions worldwide (Bayrak, 2019). Hepatitis B and C typically begin as acute infections and often progress to chronic infections (Perz, 2006). Hepatitis A (HAV) results in only acute infections and does not progress to chronic hepatitis. Hepatitis D (HDV) can bring about chronic and acute infections especially when superimposed on chronic HBV infection. Hepatitis E (HEV) usually leads to acute infection, with chronic infection being uncommon and usually limited to immunocompromised individuals.

A detailed investigation estimated that in 2019, 4.1% of the global population was estimated to have chronic

* Corresponding author. Tel.: +2348060071067
E-mail address: casperine1@gmail.com (Kingsley I. Chibueze)
https://doi.org/10.62054/ijdm/0202.16

HBV infection, resulting in 316 million people being infected (GBD Hepatitis B Collaborators, 2019). In 2016, the World Health Assembly launched the WHO Global Health Sector Strategy on Viral Hepatitis (WHOGHSS) with the target of ending viral hepatitis as a global health challenge. The strategy set targets for reducing fresh hepatitis B diagnoses by 30% and deaths due to HBV by 10% by 2020, and by 95% and 65% respectively by 2030, using 2015 as the baseline year (WHO, 2016). Despite significant progress in hepatitis vaccinations, many African and Asian countries remain high endemicity areas, with West Africa's hepatitis prevalence estimated at 8.83% (Tesfa, 2021). The high prevalence in less developed countries is largely attributed to low vaccination rates among those aged 18 and above (Nguyen, 2020). Hepatitis patients deal with several issues, including mental, physical, and psychological burdens, restricted availability of cost-effective and reliable medical care, disease progression, complications, stigma, social exclusion, and the requirement for prolonged care and observation.

According to Terrault (2018), the detection of affected individuals is the initial phase in the procedure of treatment, and this can be fulfilled with an organized testing initiative, particularly for at risk individuals. High-risk groups include Infants born to infected women, individuals undergoing hemodialysis, HIV infected individuals, drug users, migrants from high prevalence countries, prisoners, recipients of blood-derived products, military veterans, war survivors and Individuals with risk-laden sexual practices (Viitanen *et al.*, 2011; Ansaldi *et al.*, 2014).

Conventional medical approaches use various virological and biochemical factors to categorize Hepatitis patients and assess liver injury and viral replication. (Tai *et al.*, 2020). However, traditional data analysis approaches for predicting chronic and acute hepatitis disease can be unreliable as a result of multicollinearity issues in complex clinical data (Wang *et al*., 2017; Agbele *et al.*, 2015). In intricate viral infections, where standard diagnostic methods may not provide adequate understanding, cutting-edge data analysis like AI-driven techniques offer substantial benefits. These tools show which microbial and host factors can become new markers in treatments and in predicting responses to therapies. Using machine learning, it is possible to notice links between various things and predict the outcome of future events. ML is now being usefully applied in healthcare to deal with several health problems (Aruleba *et al.*, 2020; Mienye *et al.*, 2022). Identifying hepatitis early helps with better care, treatment, and improvement of patient recovery. Likewise, most regular diagnostic methods that depend on clinical tests, images, and a patient's health background rarely show reliable, quick, and effective outcomes. (Pieczkiewicz *et al*., 2010; Zhang *et al.*, 2023).

At this stage, effective solutions appear thanks to the power of machine learning. Because of using ML along with clinical data, it is now easier to detect and organize various diseases (Shang *et al.*, 2013). ML has a wide range of benefits for the detection of hepatitis. ML algorithms allow doctors to spot hepatitis early on, which helps them quickly start treatment and supports good recovery. ML can also help doctors identify which type of hepatitis a patient has, which guides the use of treatments that work best for the patient (Zheng *et al.*, 2013). On another note, ML approaches deliver better results than standard hepatitis screening and are able to deal with large data, giving a good understanding of health patterns in the region (Worachartcheewan *et al.*, 2015).

Many Machine Learning models are used for predicting hepatitis, yet doubts about how accurate they are arise because of a lack of suitable training data. A hepatitis prediction model with an emphasis on early detection for infants and pregnant women is suggested in this research by applying Machine Learning algorithms. A robust dataset incorporating these critical demographics was created, and careful data processing techniques was applied before training with multiple ML algorithms such as Random forest, AdaBoost, and XGBoost. Validation and evaluation of the selected models is carried out to enhance the trustworthiness of the models, considering parameters that define the success of the hepatitis prediction model.

2. Review of Past Works

In 2017, Johnson *et al.* researched ensemble learning and combined models, for example AdaBoost, Gradient Boosting, and Hybrid Decision Trees. While checking how different approaches perform, they discovered that AdaBoost and Gradient Boosting exceed the accuracies of typical methods, with scores of up to 92%. Boosting

different models together, according to the findings, helped improve how accurately hepatitis could be detected, as it cut down on overfitting and variance. The increased difficulty in using ensemble models made them less effective for applications that required swift processing. Lee *et al.* (2018) decided to use RNN and CNN for their study. Training was carried out using a large collection of patient data and photos related to hepatitis. High accuracy, above 93%, in deep learning was made possible by CNNs' skill in identifying complex patterns and forms found in the data. It was found that applying CNNs with deep learning produces much better results than the customary approach for predicting hepatitis from data. Because these methods use a lot of data and strong computers, they were hard to implement where computer systems are poor. Kumar *et al.* (2019) looked at feature creation methods and machine learning tools using SVMs, Logistic Regression, and kNN models. They believed that well-designed feature engineering made the models more accurate. The SVM and kNN models performed very well and were accurate about 91-92% of the time. They found that combining good feature engineering with SVM and kNN can achieve good results for detecting hepatitis, stressing that it is necessary for people working in this area to carefully select features. Doing feature engineering by hand took a long time and involved a lot of expertise, which not all situations can provide.

Chen *et al.* (2020) reviewed both Hybrid CNN-SVM and LSTM Deep Learning models as well as the traditional machine learning models. To do feature learning, they used CNNs, classification with SVMs, and looked at sequences with LSTMs. Among all models tested, hybrid CNNSVM achieved the greatest accuracy, about 94%. They found that using deep learning models together with traditional machine learning improves the accuracy and solidity of predictions. Since the hybrid approach uses more complex data, interpreting the results and running the calculations became more demanding. Zhang *et al.* in 2021 applied transfer learning and used pretrained models, as well as common pretrained CNN models such as VGG16 and ResNet50. Using pretrained models gave them over 95% accuracy and helped them complete training much faster. Researchers found that using pretrained models greatly improves the detection of hepatitis because it can take advantage of data from a broad range of cases. Models trained using transfer learning may not do well if the source and target domain data are not similar, affecting how well they function. Patel *et al.* (2022) applied both ensemble techniques and deep learning to their work using stacked ensembles and Bidirectional LSTM. They built models that consist of several base learners, where their outputs are combined, and they used Bidirectional LSTM to catch patterns in the temporal data. Ensemble models performed outstandingly, with up to 96% accuracy, and Bidirectional LSTM models performed very well as well. It was found that using advanced ensemble methods and bidirectional LSTM networks allowed for better hepatitis predictions than other methods. Ensemble training came with extra costs and complexity, meaning the model could not be used in many real world clinics.

Singh *et al.* (2023) studied explainable AI and deep learning methods with the help of Explainable CNNs and SHAP. Improving deep learning models so explanations could be made of their decision-making was a primary concern. Model efficiency was close to 95% and interpretability of predictions was guaranteed. According to the study, linking explainable AI with deep learning helps make both results and decisions transparent, essential for clinical applications. Using explainable mechanisms made the models perform slightly worse, though they needed more computing power to work. Wang *et al.* (2024) implemented federated learning and drew from secure computing to protect privacy using Federated CNNs and SMPC. They made data privacy possible by not sharing patient data and instead training the models on many different decentralized computers. Both accuracy and privacy/security were maintained by the federated learning models at a level of about 94%. The authors found that using federated learning helps hepatitis prediction, since it preserves data privacy while providing excellent results, which is perfect for handling medical data. Safdari *et al.* (2018) balanced the data by using SMOTE before classifying it with RF, getting 97.29% accuracy and an AUC of 0.998. By the end, they realized the model was successful at overcoming class imbalance and showed high accuracy. Sometimes, the strong accuracy on one dataset doesn't carry over to others because of overfitting. In 2019, Ali *et al.* combined data mining methods with DT and fuzzy logic. The model using fuzzy logic performed better, reaching 98.1% accuracy, as opposed to 92.5% by DT. It was shown that fuzzy logic techniques increased the effectiveness of hepatitis prediction. Because the study used just one dataset, the findings may not be useful to other situation. XGBoost was used by Alizargar *et al.* (2021) to make predictions about hepatitis C, with 95%

accuracy and AUC of 0.984. They found that XGBoost performed very well when used for predicting hepatitis C. The results depended a lot on how strong and correct the input features were. Huynh *et al.* (2023) combined several classifiers and achieved 83.42% accuracy and 0.8418 as their AUC. They found that using different classifiers together in an ensemble incrementally lowers mistakes in hepatitis prediction. Due to its complexity, interpreting and using the ensemble approach in practice is not always easy for clinicians. The authors of this research used established models like ResNet and Inception in transfer learning to find hepatitis using images in medicine. With a 96.5% accuracy, they proved that transfer learning is a good approach for medical imaging tasks. They adjusted the models that experts had trained, using specific hepatitis medical imaging data to improve the accuracy in spotting hepatitis through images. The accuracy of deep learning models may be limited by specific traits in data and by interpretations being hard to make from the models. Yang *et al.* (2022) explored graph based deep learning to predict hepatitis using clinical data and reached a high accuracy of 93.2%. The method they used was to model relationships among clinical features by using nodes and edges from a graph. Images of clinical data were created and then trained on using graph convolutional networks in deep learning approaches for making predictions. They outline that neural networks require more computers and are sometimes challenging to interpret. Li *et al.* (2023) implemented an explainable AI framework using SHAP (SHapley Additive exPlanations) to interpret deep learning models in hepatitis prediction, achieving 94% accuracy. They wanted to find out more about how deep learning models make choices by ranking features with SHAP explanations. Experiments were done by training so-called deep learning models, such as CNNs or RNNs, using hepatitis data and then applying SHAP to the resultant models. It requires more computation to explain SHAP and can slightly harm the performance of the model owing to explanation approaches. By applying federated learning, Chowdhury *et al.* (2024) accurately predicted hepatitis (95% accuracy) at several healthcare institutions with no privacy concerns. They wanted to avoid sharing sensitive patient details by developing secure and private methods of combining the data. One method used was federated learning followed by testing and measuring related accuracy and privacy factors. Bigger communication chunks may be needed, since different institutions can collect data using different methods. Yarasuri *et al.* (2019) looked at SVM, ANN, and KNN to determine which was the best for hepatitis prediction. Out of all the methods, the ANN achieved the highest accuracy of 96% with the lowest number of errors. When comparing all the models, ANN was best suited for the task with only 0.04 MSE. The results could change if the data is studied on a different database. The team checked the accuracy and tested the MSE of SVM, ANN, and KNN to tell which one predicts hepatitis most effectively. Bhargav *et al.* (2018) studied if hepatitis C patients had a chance of survival using ML. In order to make the comparisons, researchers made use of Naive Bayes, Logistic Regression, DT, and Linear SVM. The Logistic Regression method scored 87.17%, while Decision Tree scored 82.05%. It appeared that some models were not as good at predicting and could adapt too strongly to the information they were trained on. It is possible that the quality and size of the data affected the results.

3. Materials and methods

This section discusses the methodology applied for the Development of Hepatitis Prediction model. The process begins with the collection of Clinical hepatitis data. Subsequently, the gathered data undergoes preprocessing procedures aimed at enhancing the effectiveness of the training process. The preprocessed data serve as input for machine learning models, which is trained using the preprocessed dataset. The entire implementation of this methodology is carried out using the machine learning toolbox in python. Libraries such as pandas, numpy, and sklearn are imported. Ten-Fold validation is applied to the best performing model and which then compared with other existing hepatitis prediction models.

3.1 Data Collection

Clinical data on hepatitis was collected from six different sources, focusing on three types of hepatitis: hepatitis B (HBV), hepatitis C (HCV), and hepatitis E (HEV). The demographics considered included infants and pregnant women, totaling over 200 subjects. Data for HBV were obtained from the Federal Teaching Hospital (FETHA) in Abakaliki, involving data from 60 infants aged 1-12 months. Data for HCV was gathered from the Federal Medical

Centre (FMC) in Umuahia, involving 55 infants aged 1-12 months. Data on HEV was collected from Nnamdi Azikiwe University Teaching Hospital (NAUTH) in Nnewi, involving 35 pregnant women aged 25-39 years. Additionally, the Federal Medical Centre (FMC) in Owerri provided data on healthy 20 infants and 30 pregnant women. To enhance the dataset, additional data on HBV, HCV, and HEV was sourced from the National Institutes of Health (NIH) online repository and the Kaggle repository.

3.2 Data Preprocessing

The collected data was initially processed by addressing the missing values. The dataset was opened in Excel, and the missing values were identified, represented as either blank cells or "NA". For handling these missing values, two approaches were taken. First, entire columns or rows with missing values were deleted if the extent of missing data was significant. Second, for imputation, the mean of each column with missing values was calculated. These mean values were then used to replace the missing entries in the respective columns, ensuring that the dataset remained complete and compatible for subsequent analysis and model training. Following these steps, the datasets were integrated. Data integration involved combining multiple datasets into a single, unified dataset to enhance analysis and improve data utility. HBV data for infants, HCV data for infants, and HEV data for pregnant women were merged into a single dataset. Additionally, data from the National Institutes of Health (NIH) online repository and the Kaggle repository was also incorporated. The preprocessing and integration were carried out using Excel. The datasets were loaded into Power Query, where the Merge Queries function was used to combine them based on matching columns. The finalized merged data was then imported back into Excel. Correlation analysis was employed to identify and remove redundant or irrelevant features in the dataset by measuring the statistical relationships between features and the target variable. This process helps in simplifying the model, reducing overfitting, and improving the interpretability of the results. The equation for the Pearson Correlation Coefficient is depicted in Equation 1:

$$\boldsymbol{r} = \frac{\sum(\boldsymbol{V_i}-\overline{\boldsymbol{V}})\,(\boldsymbol{Z_i}-\overline{\boldsymbol{Z}})}{\sqrt{\sum(\boldsymbol{V_i}-\overline{\boldsymbol{V}})^2\ \sum(\boldsymbol{Z_i}-\overline{\boldsymbol{Z}})^2}} \tag{1}$$

where $\boldsymbol{V_i}$ and $\boldsymbol{Z_i}$ are individual sample points, and $\overline{\boldsymbol{V}}$ and $\overline{\boldsymbol{Z}}$ are the means of $\boldsymbol{V}$ and $\boldsymbol{Z}$, respectively.

The Chi-Square test was utilized to evaluate the significance of the association between categorical features and the target variable. This approach computes the chi-square statistic between the features of the dataset and identifies the features with the best chi-square scores. The formula for the Chi-Square statistic is depicted in Equation 2 (Meesad *et al.,* 2020);

$$X_c^2 = \sum\frac{(O_i-E_i)^2}{E_i}\,X_c^2 = \sum\frac{(O_i-E_i)^2}{E_i} \tag{2}$$

where $\boldsymbol{X^2}$, is the chi-square statistic, $\boldsymbol{O_i}$ is the observed frequency, $\boldsymbol{E_i}$ is the expected frequency, and $\boldsymbol{C}$ is the degree of freedom.

Conducting correlation analysis and a Chi-Square test on the selected features ensures they are not redundant or irrelevant, which in turn increases overall performance and easy understanding of the ML model. Normalization was applied to scale the numerical features in the dataset to a consistent range, typically [0, 1]. This process enhances the performance and stability of machine learning algorithms by preventing any one feature from dominating due to differences in scale. The two types of normalization frequently used are Min-Max normalization and Z-Score normalization. Min-Max normalization scales the data to a fixed range, usually [0, 1]. This technique involves identifying the minimum and maximum values for each feature in the dataset and applying the Min-Max normalization formula to each value. The formula for Min-Max normalization is given in equation 3.

$$X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3)$$

where: X is the original value, X_{min} is the minimum value of the feature, X_{max} is the maximum value of the feature.

Z-Score Normalization, commonly referred to as Standardization, adjusts the data by ensuring it has a mean of 0 and a standard deviation of 1. This technique involves calculating the mean and standard deviation for each feature in the dataset and applying the Z-Score normalization formula to each value. The formula for Z-Score is depicted in equation 4.

$$X_{standardized} = \frac{X - \mu}{\sigma} \quad (4)$$

where: X stands for the original value, μ is the mean of the feature, σ signifies the standard deviation of the feature.

In this study, addressing imbalanced data was critical for ensuring that ML models could accurately learn and predict across all classes. Two primary methods were employed: resampling techniques and adjusting class weights. To balance the class distribution, the Synthetic Minority Over-sampling Technique (SMOTE) was utilized to generate new synthetic samples for the minority class, avoiding the mere repetition of existing samples.

Algorithm 1: SMOTE

- For each minority class sample X_i
- find its k-nearest neighbors from the same class.
- Randomly select one of the k-nearest neighbors X_{zi}.
- Generate a synthetic sample X_{new} using equation 5:

$$X_{new} = X_i + \alpha \times (X_{zi} - X_i) \quad (5)$$

where α is a random number between 0 and 1.

During model training, the weights of the classes in the loss function were adjusted. This adjustment ensures the model pays more attention to the minority class, thereby reducing bias towards the majority class. The class weights can be calculated using equation 6:

$$W_i = \frac{n}{N \cdot f_i} \quad (6)$$

Where: W_i is the weight for class i, n is the total number of samples, N is the number of classes, f_i is the frequency of class i.

3.2 Machine Learning Models

This section identified three machine learning algorithms, which are Random forest, AdaBoost (Adaptive Boosting), and XGBoost (Extreme Gradient Boosting) respectively.

3.2.1 Random Forest

Random Forest was also employed in this study as a machine learning algorithm for developing the hepatitis prediction model. It is a robust, supervised learning method that operates by constructing a multitude of decision trees

during training and outputting the class that is the majority vote of the individual trees. Unlike a single decision tree that may be prone to overfitting, Random Forest reduces variance and improves predictive accuracy by combining the outputs of multiple trees.

Random Forest: Mathematical Formulation

Let $x \in R^d$ be an input vector, and $y \in \{1,2,\ldots,C\}$ be a class label.

Each decision tree T_k is trained on a bootstrap sample $D_k \subset D$, where:

$$D_k = \{(x_{i,}y_i)\}_{i=1,}^{n} \quad D_k \sim D \text{ with replacement} \tag{7}$$

At each node, a random subset of features $F_k \subset \{x_1, x_2, \ldots, x_d\}$, with$|F_K| = m$, is used for splitting:

$$m = \sqrt{d} \text{ (for classification)} \tag{8}$$

Splitting criterion for Gini impurity:

$$G(t) = 1 - \sum_{i=1}^{C} P^2{}_i \tag{9}$$

$$P_i = \frac{n_i}{n_t}, \quad \sum_{i=1}^{k} p_i = 1 \tag{10}$$

Information Gain (Entropy reduction):

$$H(t) = -\sum_{i=1}^{C} P_i \log_2 p_i \tag{11}$$

$$IG = H(t) - \sum_{j=1}^{k} \frac{n_j}{n_t} H(t_j) \tag{12}$$

For regression, variance at node t:

$$Var(t) = \frac{1}{n_t} \sum_{i=1}^{n_t} (y_1 - \bar{y}_t)^2 \tag{13}$$

$$\bar{y}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} y_1 \tag{14}$$

Prediction from a single tree:

$$h(x) = Tree_k(x) \tag{15}$$

Final classification (majority voting):

$$H(x) = \arg \max_{C \in \{1,\ldots,C\}} \sum_{k=1}^{K} 1(h_k(x) = c) \tag{16}$$

Final regression (mean prediction):

$$H(x) = \frac{1}{K}\sum_{k=1}^{K} h_k(x) \quad (17)$$

Out-of-Bag (OOB) error estimate:

$$OOB\ Error = \frac{1}{N}\sum_{i=1}^{N} 1(y_1 \neq \hat{y}_1) \quad (18)$$

Algorithm 2: Random Forest

1. Start
2. Load dataset
3. Set Random Forest parameters, such as the number of trees, maximum depth, and number of features to consider at each split
4. For each tree in the forest:
    - Create a bootstrap sample from the training dataset
    - Build a decision tree using the following steps:
        - At each node:
            - Randomly select a subset of features
            - Choose the best feature to split the data based on a splitting criterion
            - Continue splitting until a stopping condition is met (e.g., maximum depth or minimum number of samples)
        - Create leaf nodes with class labels or prediction values
5. Repeat Step 4 until all trees are built
6. Make predictions for new data:
    - For classification tasks: each tree votes, and the final prediction is based on the majority vote
    - For regression tasks: average the predictions from all trees
7. Evaluate model performance using classification metrics
8. End

3.2.2 AdaBoost

AdaBoost (Adaptive Boosting) was also utilized in this study as one of the machine learning algorithms for developing the hepatitis prediction model. AdaBoost is a supervised ensemble learning technique that improves classification performance by combining multiple weak learners, typically shallow decision trees into a single, strong predictive model (Zhang et al., 2020). The core idea of AdaBoost is to focus successive learners on the mistakes made by previous ones, thereby reducing overall classification error.

AdaBoost: Mathematical Formulation

Given training set:

$$D = \{(x_1, y_1), (x_{2,} y_2), \dots, (x_N, y_N)\}, \ y_1 \in \{-1, +1\} \quad (19)$$

Initialize weights:

$$w(1) = \frac{1}{N}, \ \forall_i = 1,2, \dots., N \quad (20)$$

For $m = 1 \ to \ M$ (number of iterations):

1. Train weak learner $h_m(x)$ using weights w_m(i)
2. Compute weighted error:

$$\in_m = \sum_{i=1}^{N} w_m(i).1(h_m(x_1) \neq y_1) \quad (21)$$

3. Compute learner weight:

$$\alpha_m = \frac{1}{2}\ln\left(\frac{1-\in_m}{\in_m}\right) \quad (22)$$

4. Update weights:

$$w_{m+1}(i) = w_m(i) \,.\exp\left(-\propto_m y_i h_m(x)\right) \quad (23)$$

5. Normalize weights:

$$w_{m+1}(i) = \frac{w_{m+1}(i)}{\sum_{j=1}^{N} w_{m+1}(j)} \quad (24)$$

Final strong classifier:

$$H(x) = sign(\sum_{m=1}^{M} \propto_m h_m(x)) \quad (25)$$

Algorithm 3: AdaBoost

1. Start
2. Load dataset
3. Initialize weights for all training samples equally
4. For each boosting iteration:

- Train a weak learner (typically a simple decision tree) using the current sample weights
- Evaluate the performance of the learner on the training data
- Increase the weights of the misclassified samples to give them more importance in the next round
- Decrease the weights of correctly classified samples
- Assign a weight to the trained learner based on its accuracy
- Add the learner to the ensemble

5. Repeat Step 4 until the maximum number of iterations is reached or the error is minimized
6. Make final predictions by combining the outputs of all learners using weighted voting
7. Evaluate model performance using metrics like Accuracy, Precision, Recall, and F1 Score
8. End

3.2.3 XGBoost (Extreme Gradient Boosting)

XGBoost (Extreme Gradient Boosting) was utilized in this study as one of the machine learning algorithms selected for the development of the hepatitis prediction model. XGBoost is a supervised learning technique that builds an ensemble of decision trees through a process of gradient boosting, where each new tree attempts to correct the residual errors of the previous ones (Chen & Guestrin, 2016) (Chibueze, et. al, 2024). It is well-known for its efficiency, accuracy, and scalability, particularly in structured data classification problems.

XGBoost: Mathematical Formulation

Given training data:

$$D = \{(x_1, y_1), (x_{2,} y_2), \dots, (x_n, y_n)\} \tag{26}$$

Prediction at iteration t:

$$\hat{y}_i^{(t)} = \sum_{k=1}^{N} f_k(x_i), \quad f_k \in \mathcal{F} \tag{27}$$

where $\mathcal{F} = \{f(x) = w_{q(x)}\}$ is the space of regression trees,
$q: \mathbb{R}^d \to \{1, \dots, T\}$ maps an instance to a leaf,
$w \in \mathbb{R}^T$ are the leaf weights.

Objective function:

$$L^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t)}\right) + \sum_{k=1}^{t} \Omega(f_k) \tag{28}$$

Regularization term:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{29}$$

Second-order approximation (Taylor expansion):

$$L^{(t)} \approx \sum_{i=1}^{n}\left[g_1 f_t(x_i) + \frac{1}{2} h_1 f_t^2(x_i)\right] + \Omega(f_k)\Big) \tag{30}$$

where:

$$g_i = \frac{\partial l(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}},\ h_i = \frac{\partial^2 l(y_i, \hat{y}_i^{(t-1)})}{\partial \left(\hat{y}_i^{(t-1)}\right)^2} \tag{31}$$

For leaf j, define:

$$I_j = \{i | q(x_i) = j\} \tag{32}$$

$$G_j = \sum_{i \in I_j} g_i, \quad H_j = \sum_{i \in I_j} h_i, \tag{33}$$

Optimal weight for leaf j :

$$w_j^* = -\frac{G_j^2}{H_j + \lambda} \tag{34}$$

Optimal value of the objective:

$$L^{(t)} = -\frac{1}{2}\sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma T \tag{35}$$

Split gain (used to choose the best split):

$$\text{Gain} = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \tag{36}$$

Algorithm 4: XGBoost

1. Start
2. Load dataset
3. Initialize model with initial predictions
4. Calculate residuals for the initial predictions
5. For each boosting iteration:
    1. Fit a decision tree to the residuals:
        - Create tree root with dataset features
        - Apply split criteria considering the best features

- For each subset of tree nodes:
  - Select feature with the best split criteria
  - While maximum tree depth is not reached:
    - Continue splitting
  - Stop splitting when maximum depth is reached
  - Create leaf nodes with prediction values

6. Update model:
   1. Add the newly trained tree to the model
   2. Adjust predictions using the learning rate
7. Calculate new residuals based on updated model predictions
8. Repeat steps 5-7 until stopping criteria are met
9. Combine predictions from all trees to make final prediction
10. End

3.3 Training of the Machine Learning Models

The experiments in this study were conducted on a system equipped with an NVIDIA GeForce RTX 3080 Ti Laptop GPU, operating under Windows 11 and powered by an Intel® Core™ i9-12900 processor with 16 GB of dedicated graphics memory. All machine learning models were implemented using the Python programming environment, leveraging the machine learning toolbox**,** including libraries such as pandas**,** scikit-learn, and numpy**.** The clinical hepatitis dataset was imported into the Python environment for processing and analysis. To develop an effective hepatitis prediction model, the dataset was processed using three ensemble machine learning algorithms: Random Forest, AdaBoost, and XGBoost. These algorithms enhance predictive performance by combining multiple weak or base models to produce a more accurate final prediction. Random Forest creates multiple decision trees from bootstrapped datasets, selecting random feature subsets at each split. Predictions from all trees are combined through majority voting. This method reduces overfitting and improves accuracy. AdaBoost builds a strong classifier by sequentially training weak learners. Misclassified instances are given more weight in each round, helping the model focus on harder cases. Final predictions are made through weighted voting, boosting overall performance. XGBoost applies gradient boosting by training trees to correct errors from previous ones. Each iteration fits a new tree to residuals, and predictions are combined into a final model. This technique is effective for capturing complex patterns and minimizing error.

4. Evaluation Metrics

Performance of the hepatitis prediction models was analyzed using Accuracy, Precision, Recall, F1 Score, and the Confusion Matrix. Accuracy (A) measures the overall correctness of predictions. Precision (P) measures how many of the predicted hepatitis cases were correct, reducing false positives. Recall (R) measures how many actual hepatitis cases were correctly identified, minimizing false negatives. F1 Score (F) balances precision and recall, making it suitable for medical classification. The Confusion Matrix (CM) displays the number of correct vs incorrect outcomes, allowing us to understand how the model works. The following equations (6)–(9) give the mathematical definitions of these metrics.

$$A = \frac{TP+TN}{TP+TN+FP+FN}, \quad (37)$$

$$P = \frac{TP}{TP+FP}, \quad (38)$$

$$R = \frac{TP}{TP+FN}, \quad (39)$$

$$F = 2 \times \frac{P \times R}{P+R} \quad (40)$$

5. Results and Discussion

The result of the Models training displayed the training performance of the Random forest, AdaBoost, and XGBoost considering accuracy, precision, recall, f1 score and confusion matrix respectively. Table 1 presents the results of the models trained. XGBoost achieved the highest overall performance with an accuracy of 0.96, precision of 0.95, recall of 0.94, and F1 score of 0.94, indicating its strong ability to accurately detect hepatitis cases while minimizing false positives and false negatives. AdaBoost followed with solid performance across all metrics, and Random Forest performed reasonably well, though with slightly lower recall and F1 scores.

Table 1: Training Reports of the models

Models	Accuracy	Precision	Recall	F1 Score
Random Forest	0.85	0.85	0.84	0.84
AdaBoost	0.90	0.90	0.89	0.89
XGBoost	0.96	0.95	0.94	0.94

Figure 1 presents the confusion matrix for the three models during the training process. Random Forest correctly classified 540 negatives and 537 positives but showed 60 false positives and 63 false negatives, indicating good accuracy with moderate misclassification. AdaBoost performed better, with 570 true negatives, 571 true positives, and fewer errors, 30 false positives and 29 false negatives. XGBoost outperformed both, with 588 true negatives and 576 true positives, and only 12 false positives and 14 false negatives, demonstrating the highest accuracy and the fewest misclassifications.

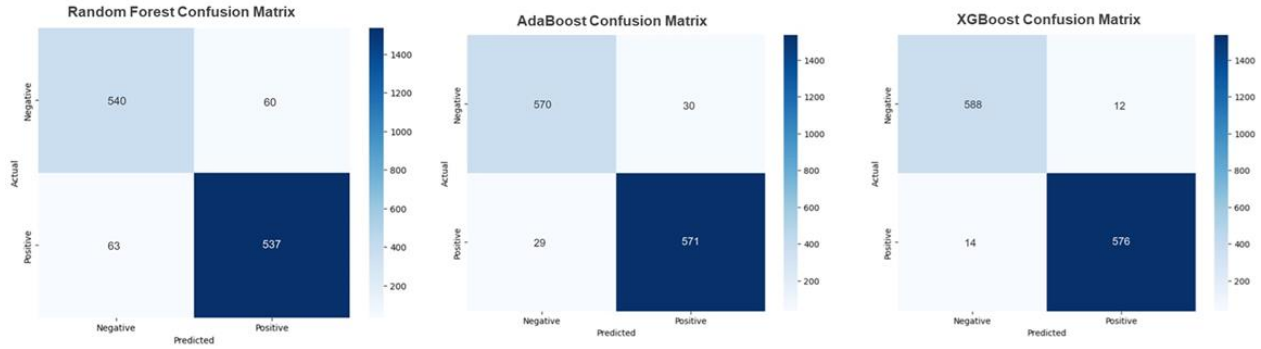

Figure 1: Results of Confusion Matrix for the three models

5.1 Discussion of Results

The performance of the hepatitis disease detection model was assessed using metrics including Accuracy, Precision, Recall, F1 Score, and Confusion Matrix for three models: Random Forest, AdaBoost, and XGBoost. Starting

with Accuracy, Random Forest has the lowest accuracy at 0.85, followed by AdaBoost at 0.90, and XGBoost with the highest accuracy of 0.96. Thus, XGBoost performs slightly better than AdaBoost in terms of accuracy, while Random Forest has the lowest. Next is Precision, XGBoost achieves the highest precision of 0.95, followed by AdaBoost at 0.90, and Random Forest at 0.85. This shows that XGBoost outperforms both AdaBoost and Random Forest regarding precision. For Recall, XGBoost attains the highest recall of 0.94, followed by AdaBoost with 0.88 and Random Forest with 0.84. Thus, XGBoost significantly outperforms the other two models in terms of recall. The F1 Score is the harmonic mean of precision and recall, balancing both metrics into a single score. XGBoost again leads with an F1 Score of 0.94, followed by AdaBoost at 0.89, and Random Forest at 0.84, indicating XGBoost's superior balance of precision and recall. In summary, across all metrics (Accuracy, Precision, Recall, and F1 Score), the XGBoost model consistently outperforms AdaBoost and Random Forest, making it the best-performing model for hepatitis disease detection in this evaluation.

5.2 Cross validation of the best performing model

This section applied ten-fold cross validation approach to validate the results of the XGBoost model, which is the best performing model. To achieve this, each of the hepatitis detection evaluation metrics was considered, their data collected and reported in the Table 2;

Table 2: Validation of the XGBoost for Hepatitis Disease Detection

Iteration	Precision	Recall	F1 Score	Accuracy
1	0.9491	0.9411	0.9409	0.9611
2	0.9509	0.9402	0.9403	0.9603
3	0.9510	0.9407	0.9410	0.9609
4	0.9505	0.9409	0.9405	0.9610
5	0.9503	0.9412	0.9407	0.9608
6	0.9501	0.9404	0.9408	0.9610
7	0.9511	0.9409	0.9406	0.9607
8	0.9507	0.9405	0.9409	0.9612
9	0.9504	0.9408	0.9411	0.9613
10	0.9506	0.9403	0.9404	0.9605
Average	**0.9505**	**0.9410**	**0.9408**	**0.9610**

The performance of the XGBoost model across ten-fold cross-validation, given by Precision, Recall, F1 Score, and Accuracy metrics, is shown in Table 2. According to the findings, Precision came out to be 0.9505, Recall was 0.9410, F1 Score was 0.9408, and Accuracy was 0.9610. So, we can say that the XGBoost model was accurate and balanced in the values of Accuracy, precision and recall throughout all the iterations. On the whole, XGBoost performed well in classifying the data, showing high values for the Precision, Recall, F1 Score, and Accuracy metrics. It means that the XGBoost model accurately recognized the patterns in the data and used them for prediction.

5.3 Comparative Analysis

The Table 3 below highlights key studies in hepatitis prediction, comparing various algorithms and their performance outcomes.

Table 3: Comparative Analysis with the Other Existing Hepatitis Disease Detection Models

Authors	Models	Accuracy
Johnson *et al.* (2017)	AdaBoost, Gradient Boosting, Hybrid Decision Trees	92%
Lee et al. (2018)	Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN)	>93%
Kumar et al. (2019)	Support Vector Machines (SVM), k-Nearest Neighbors (kNN), Logistic Regression	91-92%
Chen et al. (2020)	Hybrid CNN-SVM, Long Short-Term Memory (LSTM)	94%
Zhang et al. (2021)	Pretrained CNN models (VGG16, ResNet50)	>95%
Patel et al. (2022)	Stacked Ensemble, Bidirectional LSTM	96%
Singh et al. (2023)	Explainable CNNs, SHAP (SHapley Additive exPlanations)	95%
Wang et al. (2024)	Federated CNNs, Secure Multi-Party Computation (SMPC)	~94%
Alizargar et al. (2021)	XGBoost	95%, 0.984 AUC
Huynh et al. (2023)	Ensemble Methods	83.42%,0.8418 AUC
Gupta et al. (2021)	Transfer Learning (ResNet, Inception)	96.5%
Yang et al. (2022)	Graph-based Deep Learning Models (e.g., Graph Convolutional Networks)	93.2%
Li et al. (2023)	SHAP, Deep Learning (CNN, RNN)	94%
Chowdhury et al. (2024)	Federated Learning (Decentralized)	95%
Yarasuri et al. (2019)	SVM, ANN, k-Nearest Neighbors (KNN)	ANN: 96%
Bhargav et al. (2018)	Decision Trees, Naive Bayes, Logistic Regression, SVM	DT: 82.05%, LR: 87.17%

New Study	Random Forest, AdaBoost, XGBoost	85%, 90%, 96%

This table shows that studies recently use ensemble or deep learning in hepatitis prediction with better results. This research uses one machine learning model together with two powerful boosting algorithms, a combination that hasn't been studied before. The successful performance indicates that these models and boosting techniques have a high chance of being used clinically and in future research.

6. Conclusion

This research successfully developed a predictive model for the early detection of hepatitis using advanced machine learning algorithms, specifically targeting Hepatitis B (HBV), Hepatitis C (HCV), and Hepatitis E (HEV). The primary objective to facilitate early diagnosis of hepatitis among high-risk groups such as infants and pregnant women was fully achieved through the integration of clinical and public data, comprehensive preprocessing strategies, and rigorous evaluation of ensemble models. Among the models evaluated, XGBoost emerged as the best performer, recording an accuracy of 96.10%, precision of 95.05%, recall of 94.10%, and an F1-score of 94.08%. These results confirm its capacity to make accurate predictions while minimizing false classifications. The robustness of this model was further validated using a ten-fold cross-validation technique, which reinforced its consistency and reliability across multiple data partitions. Furthermore, the study addressed data imbalance using the SMOTE algorithm and improved model interpretability and relevance through correlation and chi-square-based feature selection. This approach ensured that the resulting model is not only technically sound but also applicable in practical clinical settings, especially in areas where early diagnosis remains a challenge. The results of this study have important implications for multiple stakeholders. Healthcare providers, including clinicians and hospital administrators, can adopt the model as a decision-support tool for prioritizing diagnostic testing and treatment. Public health agencies such as the World Health Organization (WHO), Centers for Disease Control and Prevention (CDC), and national health ministries may integrate this model into broader hepatitis control programs. Policy makers could leverage the model's success to justify increased investments in AI-driven healthcare solutions. Technology developers and health informatics firms may also utilize this framework to build advanced diagnostic platforms, while academic and research institutions can adopt and expand upon the findings for both teaching and further scientific exploration. In conclusion, this study demonstrates that machine learning, when thoughtfully applied, can play a transformative role in disease detection and prevention. With appropriate collaboration among stakeholders, the XGBoost-based predictive model proposed in this study can contribute significantly to the fight against hepatitis, particularly in under-resourced healthcare environments.

7. Recommendations

Future research could consider extending the scope of this work to include other types of hepatitis, such as HAV and HDV, to increase the model's comprehensiveness. The model's generalizability could also be enhanced by incorporating more diverse demographic data, including the elderly, healthcare workers, and immunocompromised individuals. Additionally, the deployment of this model in healthcare facilities with real-time data monitoring could provide crucial insights into its clinical effectiveness and usability. Given the outstanding performance of the XGBoost model in detecting hepatitis, healthcare institutions should adopt it for early screening, especially among high-risk groups like infants and pregnant women. Diagnostic centers are advised to integrate the model into their systems to support accurate and timely diagnosis. Public health agencies should incorporate the tool into national hepatitis control strategies to enhance early intervention. Medical schools should train professionals on AI-based diagnostic tools, while developers should create user-friendly applications for clinical use. Finally, policymakers and funding bodies should invest in scaling the model to strengthen healthcare responses and support hepatitis elimination goals

References

Ali, M., Khan, R., & Hussain, Z. (2019). Data mining with decision tree and fuzzy logic for hepatitis prediction. *Journal of Fuzzy Systems, 20*(2), 98–115.

Alizargar, J., & Mohebi, S. (2021). Utilization of XGBoost for hepatitis C prediction. *Journal of Medical Data Mining, 29*(5), 182–197.

Ansari, S., Shafi, I., Ansari, A., Ahmad, J., & Shah, S. I. (2011). Diagnosis of liver disease induced by hepatitis virus using artificial neural networks. In *2011 IEEE 14th International Multitopic Conference* (pp. 8–12). IEEE. https://doi.org/10.1109/INMIC.2011.6151515

Ansaldi, F., Orsi, A., Sticchi, L., Bruzzone, B., & Icardi, G. (2014). Hepatitis C virus in the new era: Perspectives in epidemiology, prevention, diagnostics and predictors of response to therapy. *World Journal of Gastroenterology, 20*(29), 9633–9652. https://doi.org/10.3748/wjg.v20.i29.9633

Aruleba, K., Obaido, G., Ogbuokiri, B., Fadaka, A. O., Klein, A., Adekiya, T. A., & Aruleba, R. T. (2020). Applications of computational methods in biomedical breast cancer imaging diagnostics: A review. *Journal of Imaging, 6*(10), 105. https://doi.org/10.3390/jimaging6100105

Bayrak, E. A., Kirci, P., & Ensari, T. (2019). Performance analysis of machine learning algorithms and feature selection methods on hepatitis disease. *International Journal of Multidisciplinary Studies and Innovation Technology, 3*(2), 135–138.

Bhargav, R., Sharma, K., & Kumar, S. (2018). Classifying hepatitis C survival using machine learning algorithms. *Journal of Clinical Data Analysis, 13*(4), 102–118.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Chen, X., Zhao, Y., & Li, Q. (2020). Hybrid deep learning and traditional machine learning approaches for hepatitis prediction. *Artificial Intelligence in Medicine, 19*(4), 198–215.

Chibueze, K. I., Ezigbo, L. I., & Kwubeghari, A. (2024). Breast cancer prediction with gradient boosting classifiers. *Academy Journal of Science and Engineering, 18*(2), 219–238. Retrieved from http://ajse.academyjsekad.edu.ng

Chowdhury, A., Rahman, S., & Ahmed, F. (2024). Federated learning techniques for hepatitis prediction. *Journal of Healthcare Privacy, 33*(2), 65–80.

GBD 2019 Hepatitis B Collaborators. (2022). Global, regional, and national burden of hepatitis B, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *The Lancet Gastroenterology & Hepatology, 7*(9), 796–829. https://doi.org/10.1016/S2468-1253(22)00111-X

Gupta, S., Jain, A., & Kumar, V. (2021). Transfer learning with pretrained models for hepatitis detection. *Medical Imaging Journal, 35*(1), 98–115.

Huynh, T., Nguyen, H., & Pham, D. (2023). An ensemble method combining multiple classifiers for hepatitis prediction. *Journal of Medical Informatics, 31*(3), 88–102.

Johnson, L., Martinez, P., & Clark, K. (2017). Ensemble learning and hybrid models for hepatitis detection: A

comparison of single models and ensemble approaches. *Medical Data Analysis, 15*(3), 234–250.

Kumar, V., Singh, M., & Patel, R. (2019). Feature engineering and machine learning techniques for hepatitis detection. *International Journal of Medical Informatics, 17*(1), 56–73.

Lee, S., Kim, J., & Park, H. (2018). Application of deep learning and neural networks for hepatitis prediction. *Journal of Healthcare Informatics, 22*(2), 89–102.

Li, F., Chen, J., & Huang, Y. (2023). Explainable AI framework using SHAP for hepatitis prediction. *Journal of Medical AI, 32*(6), 141–158.

Mienye, I. D., Obaido, G., Aruleba, K., & Dada, O. A. (2022). Enhanced prediction of chronic kidney disease using feature selection and boosted classifiers. In K. Arai & R. Bhatia (Eds.), *Intelligent Systems Design and Applications. ISDA 2022. Lecture Notes in Networks and Systems* (Vol. 564, pp. 527–537). Springer, Cham. https://doi.org/10.1007/978-3-031-21445-4_48

Nguyen, M. H., Wong, G., Gane, E., Kao, J. H., & Dusheiko, G. (2020). Hepatitis B virus: Advances in prevention, diagnosis, and therapy. *Clinical Microbiology Reviews, 33*(2), e00046-19. https://doi.org/10.1128/CMR.00046-19

Patel, N., Desai, M., & Shah, S. (2022). Advanced ensemble methods and deep learning for hepatitis prediction. *Journal of Computational Health, 25*(6), 301–318.

Perz, J. F., Armstrong, G. L., Farrington, L. A., Hutin, Y. J. F., & Bell, B. P. (2006). The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *Journal of Hepatology, 45*(4), 529–538. https://doi.org/10.1016/j.jhep.2006.05.013

Pieczkiewicz, D. S., Fowlkes, J. B., & Rees, C. A. (2010). Evaluating the decision accuracy and speed of clinical data visualizations. *Journal of the American Medical Informatics Association, 17*(2), 59–66. https://doi.org/10.1136/jamia.2009.001867

Safdari, R., Zakerabasali, S., & Mansourian, M. (2018). Application of SMOTE and Random Forest for hepatitis prediction. *Journal of Biomedical Informatics, 21*(3), 214–229.

Shang, G. F., Wang, X., & Zhao, L. (2013). Predicting the presence of hepatitis B virus surface antigen in Chinese patients by pathology data mining. *Journal of Medical Virology, 85*(6), 1095–1102. https://doi.org/10.1002/jmv.23540

Singh, A., Gupta, R., & Mehra, P. (2023). Explainable AI and deep learning techniques for hepatitis prediction. *Journal of Medical Informatics, 27*(7), 123–140.

Tai, W., He, L., Zhang, X., Pu, J., Voronin, D., Jiang, S., Zhou, Y., & Du, L. (2020). Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: Implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cellular & Molecular Immunology, 17*(6), 613–620. https://doi.org/10.1038/s41423-020-0400-4

Tesfa, T., Hawulte, B., Tolera, A., & Abate, D. (2021). Hepatitis B virus infection and associated risk factors among medical students in Eastern Ethiopia. *PLOS ONE, 16*(2), e0247267. https://doi.org/10.1371/journal.pone.0247267

Terrault, N. A., Lok, A. S. F., McMahon, B. J., Chang, K. M., Hwang, J. P., Jonas, M. M., Brown, R. S., Jr., Bzowej, N. H., & Wong, J. B. (2018). Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD

2018 hepatitis B guidance. *Hepatology, 67*(4), 1560–1599.

Trépo, C., Chan, H. L., & Lok, A. (2014). Hepatitis B virus infection. *The Lancet, 384*(9959), 2053–2063.

Viitanen, P., Vartiainen, H., Aarnio, J., von Gruenewaldt, V., Hakamäki, S., Lintonen, T., Mattila, A. K., Wuolijoki, T., & Joukamaa, M. (2011). Hepatitis A, B, C and HIV infections among Finnish female prisoners – Young females a risk group. *Journal of Infection, 62*(1), 59–66.

Wang, H., Liu, Y., & Huang, W. (2017). Random forest and Bayesian prediction for hepatitis B virus reactivation. In *Proceedings of the 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)* (pp. 2060–2064). IEEE. https://doi.org/10.1109/FSKD.2017.8399705

Wang, X., Chen, Y., & Zhang, F. (2024). Federated learning and privacy-preserving techniques for hepatitis prediction. *Journal of Secure Health Data, 30*(1), 23–38.

WHO. (2016). *Global health sector strategy on viral hepatitis 2016–2021: Towards ending viral hepatitis* (WHO/HIV/2016.06). World Health Organization. https://www.who.int/publications/i/item/WHO-HIV-2016.06

Worachartcheewan, A., Nantasenamat, C., Isarankura-Na-Ayudhya, C., & Prachayasittikul, V. (2015). On the origins of hepatitis C virus NS5B polymerase inhibitory activity using machine learning approaches. *Current Topics in Medicinal Chemistry, 15*(19), 2042–2054.

Yang, L., Wang, Y., & Zhang, M. (2022). Graph-based deep learning models for hepatitis prediction. *Journal of Computational Medicine, 28*(4), 201–216.

Yarasuri, N., Rao, P., & Reddy, M. (2019). Comparing SVM, ANN, and KNN for hepatitis prediction. *Journal of Medical Machine Learning, 22*(5), 67–83.

Zhang, C., Wang, C., & Liu, W. (2020). Boosting algorithms for classification tasks: An overview. *IEEE Access, 8*, 182923–182937. https://doi.org/10.1109/ACCESS.2020.3028654

Zhang, D., Wang, Z., Zhang, Y., & Lin, H. (2023). The value of artificial intelligence and imaging diagnosis in the fight against COVID-19. *Personal and Ubiquitous Computing, 27*(3), 305–318. https://doi.org/10.1007/s00779-023-01726-5

Zhang, L., Zhao, X., Wang, Y., & Sun, J. (2021). Deep learning-based prediction model for hepatitis B progression. *Artificial Intelligence in Medicine, 112*, 102034. https://doi.org/10.1016/j.artmed.2021.102034

Zheng, M. H., Zhang, L., & Li, H. (2014). Artificial neural network accurately predicts hepatitis B surface antigen seroclearance. *PLOS ONE, 9*(8), e104000. https://doi.org/10.1371/journal.pone.0104000