

## A CASE STUDY ABOUT APPLICATION OF MACHINE LEARNING IN HEALTH DOMAIN

M. Baghaei Oskuei    B. Ture Savadkoohi

*Seraj Higher Education Institute, Faculty of Computer and Electrical Engineering, Tabriz, Iran  
mahdibagaiei@gmail.com, bita.turesavadkoohi@gmail.com*

**Abstract-** Monitoring health state is a crucial and vital component of human life. Hepatitis C is a viral infection that is caused by the Hepatitis C Virus (HCV), which can result in severe liver damage in certain situations. Dark urine, yellow-tinged skin, fever, fatigue, weight loss, vomiting, headache, bone ache, and epigastric pain are a few negative effects associated with HCV. On the other hand, patients are faced with obstacles to effective treatment, disease progression prevention, complication reduction, and life-saving intervention due to the lack of accurate and timely diagnosis of hepatitis. Thus, healthcare agencies around the world are continuously improving their healthcare services with the help of Machine Learning (ML) approach. For this aim, ML technologies are utilized in this study as a technical solution in health care departments in order to process essential patient-related data related to HCV and play a significant role in early detection.

**Keywords:** Hepatitis C, Health Care, Machine Learning, Predictive Model.

### 1. INTRODUCTION

HCV is a liver disease that affects the human population. Thus, this is one of the most significant global health problems in the world [1]. The HCV is a slowly progressing disease that can cause liver cirrhosis and hepatocellular carcinoma. Actually, this kind of virus is a bloodborne virus that is transmitted through unsafe healthcare practices or through contact with the blood of a person who is infected [2]. Data science involves extracting and interpreting specific information [3]. As an example, we can mention Machine Learning (ML) as a subset of computer science that focuses on innovation recovery of computers. In the medical domain, ML has a wide variety of fields and applications, such as heart disease [4], diabetes [5], colon cancer [6], breast cancer [7], hepatitis [8-9], and much more.

Moulaei, et al. [8] conducted a meta-analysis study to predict viral hepatitis. Yaganoglu [9] applied different ML algorithms to predict hepatitis C. Kumar Sachdeva, et al. [10] proposed a systematic method for diagnosing HCV by applying a ML approach. Nilashi, et al. [7] applied a fuzzy rule method to build prediction models.

Sayfaah, et al. [11] introduced a mythology-based ML classification method to assist in the diagnosis of hepatitis. C. Ghazal, et al. [12] described an artificial intelligence system based on ML algorithms for detecting HCV.

Hashem, et al. [1] applied ML methods to predict advanced liver fibrosis. Agarwal, et al. [13] specified the risk factors of HCV in HIV-infected patients. Das, et al. [14] proposed a velocity-enhanced whale optimization approach for the classification of hepatitis disease. Another study has been conducted by Nandipati, et al. [15] to investigate and manage HCV infection. The remainder of this paper is arranged in the following way: in Section 2, the dataset used is presented, while in Section 3, the research technique and results are described. Finally, the conclusion is given in Section 4.

### 2. DATA SET

In this work, the dataset from the Gaggle website [16] is used. The columns include features such as Category, Age, Sex, Albumin (ALB), Alkaline phosphatase (ALP), Alanine amino-transferase (ALT), Aspartate amino-transferase (AST), Bilirubin (BIL), Choline esterase (CHE), Cholesterol (CHOL), Creatinine (CREA),  $\gamma$  - glutamyl-transferase (GGT) and Protein (PROT). Category features are divided into five classes: Blood Donor (0), suspected Blood Donor (0s), Hepatitis (1), Fibrosis (2), and Cirrhosis (3).

- Blood Donor class shows the people who are classified as regular blood donors. These groups have been undergoing the screening process and are considered as a healthy person without any signs of HCV or important health issues.
- Suspect Blood Donor class indicates those who are suspected of having potential health problems related to HCV or other conditions. They may show the risk factors or certain early symptoms that are required further evaluation.
- Hepatitis class represents the people who have been diagnosed with HCV. It is possible to have either acute hepatitis C (short-term) or chronic hepatitis C (long-term).
- Fibrosis class refers the people who have progressed to a stage of liver disease called fibrosis. Fibrosis occurs

when the liver attempts to heal itself by forming scar tissue due to long-term inflammation caused by HCV.

• Cirrhosis class mentions the people who have advanced liver disease known as cirrhosis. Cirrhosis is characterized by extensive scarring and disruption of normal liver function. It can occur as a result of long-standing HCV and other factors. Cirrhosis can lead to serious complications, including liver failure and an increased risk of liver cancer. The features of data set density are illustrated in Figures 1-11.

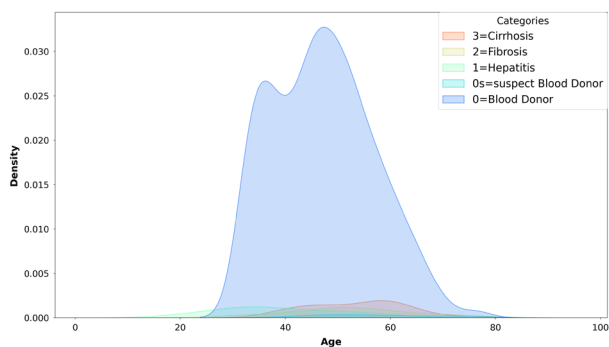


Figure 1. The density of age

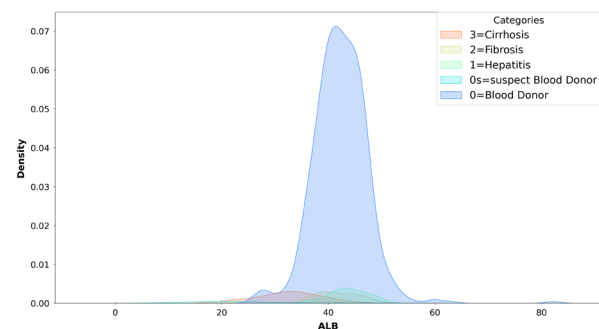


Figure 2. The density of Albumin

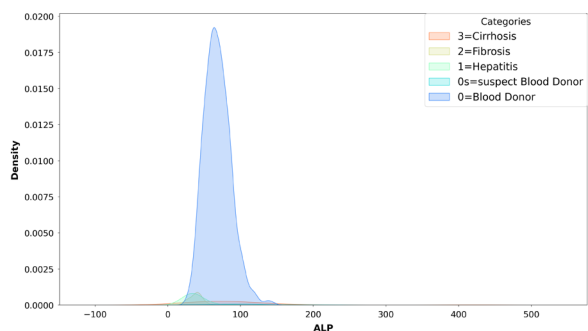


Figure 3. The density of Alkaline phosphatase

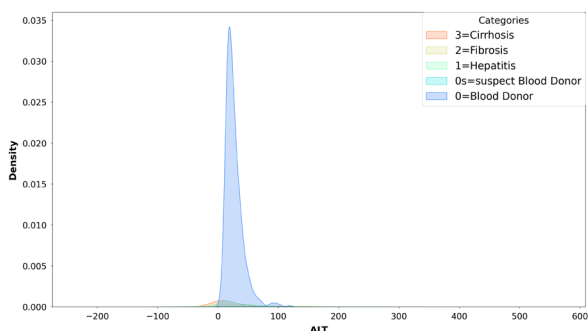


Figure 4. The density of Alanine amino-transferase

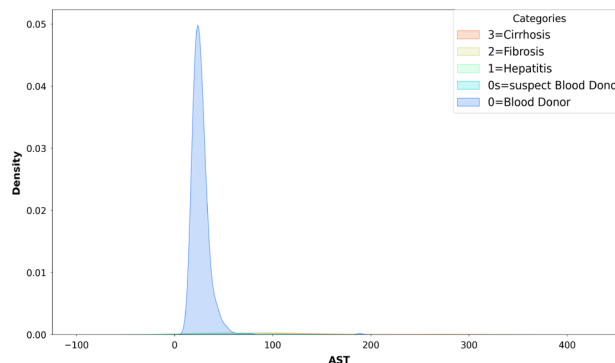


Figure 5. The density of Aspartate amino-transferase

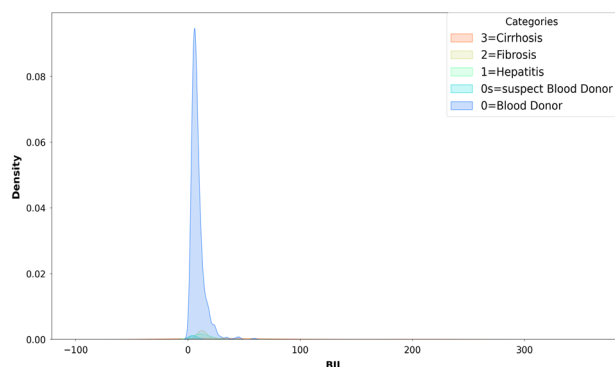


Figure 6. The density of Bilirubin

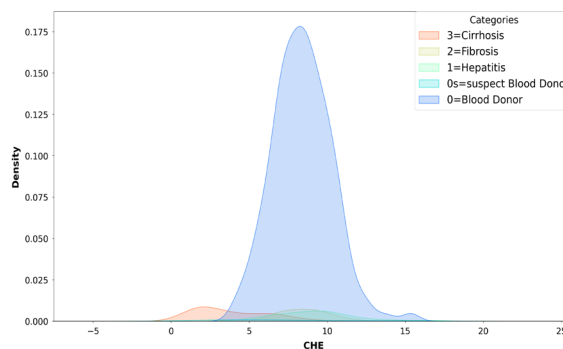


Figure 7. The density of Choline esterase

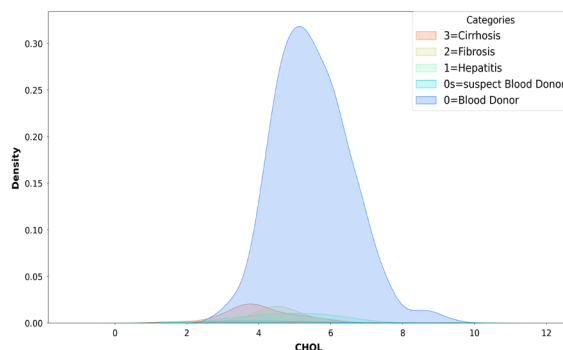


Figure 8. The density of Cholesterol

Figure 12 shows the number of null values in each column. Null values should either be replaced by new data or removed. Since some ML methods cannot work properly with these data, the number of these data is not significant and they do not have a great impact on the results, thus these data have been removed.

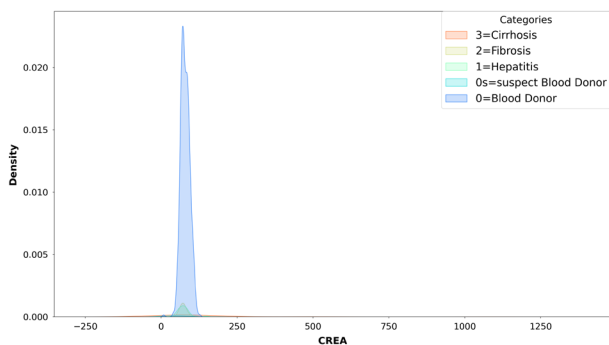


Figure 9. The density of Creatinine

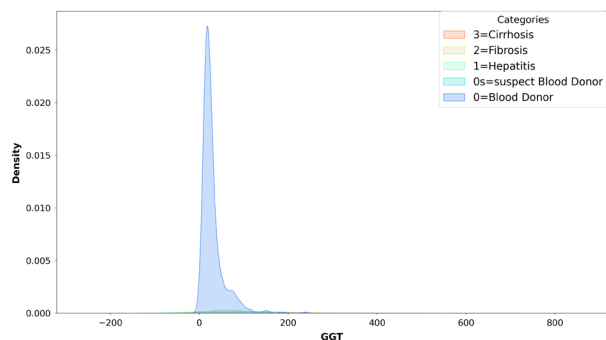


Figure 10. The density of  $\gamma$ -glutamyl-transferase

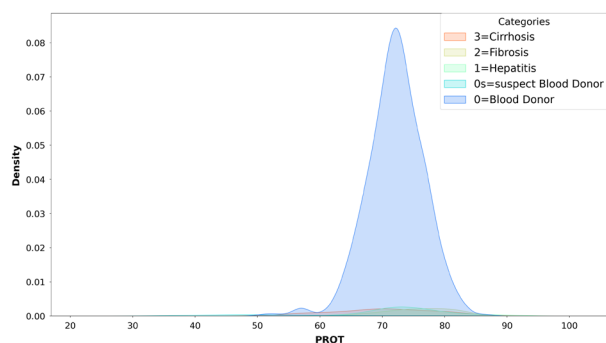


Figure 11. The density of Protein

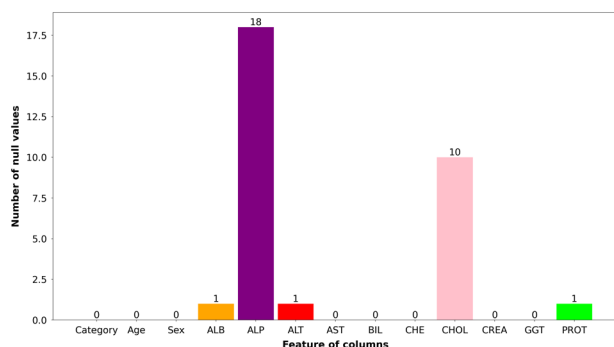


Figure 12. Total number of null values of each feature in each column

Since, the data set includes both categorical and numerical variables, thus the feature of category's column is converted to a numerical variable. Figure 13 illustrates the attributes that are applied for predicting HCV disease. For example, the connection value between category and AST is 0.65.

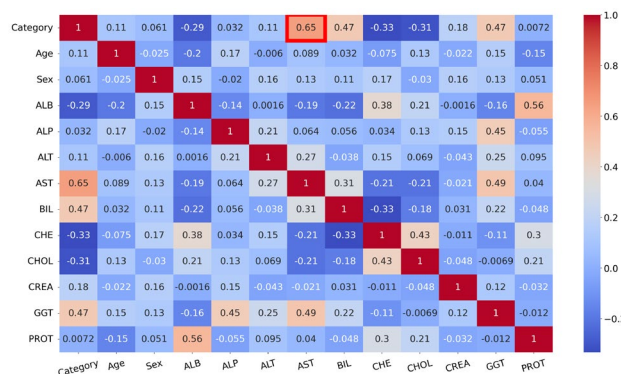


Figure 13. Correlation plot between feature-HCV diseases [9]

Figure 14 shows the outlier. Outliers are significantly different from the rest of the data. In learning process of ML classifiers outliers' effects of ML classifiers. Moreover, in data distribution they are very sensitive. Thus, the problems of the outliers should be solved or their impact on the overall results should be reduced as much as possible.

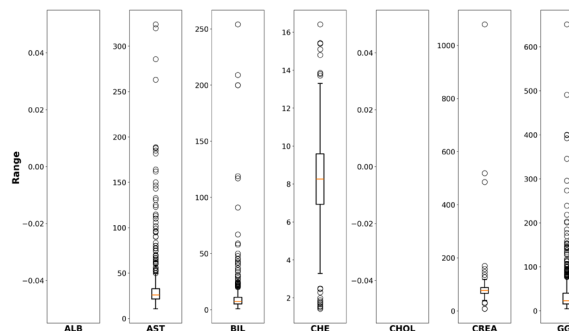


Figure 14. Boxplot of the outliers

### 3. RESEARCH TECHNIQUE

Medical assessments are an important obligation that require accurate recognition. So that, the diagnosis of diseases at the proper time causes the patient is cured at the earliest. On the other hand, utilization of ML as a newly headline in the health care because of its powerful data analysis capabilities causes to get accurate and timely disease diagnosis. In the following, the process of methodologies from the research of [4-5], [8-9] that are used in this paper will be explained.

#### 3.1. Preprocessing

Preprocessing is a crucial component of data science applications, because it can transform the input data into the proper format and make it ready for processing. So that, this step makes the processing steps more accurate and complete. The preprocessing stages are described in the following subsections.

##### 3.1.1. Normalizing Data by Computing Z-Score

Mean and standard deviation of each attribute are calculated in order to normalize data using a Z-score. Suppose,  $x = \{x_1, x_2, \dots, x_n\}$  be the number of attribute,

Z be Z-score value,  $x_i$  be data sample,  $\bar{x}$  be the mean of attribute,  $\sigma$  be standard deviation. The Z-score value and standard deviation are computed as follow, respectively [17]:

$$Z = \frac{x_{ij} - \bar{x}_j}{\sigma} \tag{1}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \tag{2}$$

**3.1.2. Logarithmic Filter**

Logarithmic filter is based on the changing all numeric value with their logarithms,  $\ln(n)$ . One of the two filters is applied as follows in order to test the value of log-filtering [18]:

- None for instance no modify, or
- *LogNums*, for instance, logarithmic filtering. All the number under 0.000001 are changed with  $\ln(0.000001)$  in order to avoid numerical errors with  $\ln(0)$

**3.1.3. SMOTE**

Synthetic Minority Over-Sampling Techniques (SMOTE) is a statistical sampling technique that is utilized to increase the number of samples in the data set uniformly [19]. Increasing the number of minority class is achieved through the use of oversampling. The component makes new instances according to the minority cases. When it receives the all datasets as an input, the percentage of only the minority instances increases. The ratio of the majority of instances stays unchanged when SMOTE is implemented. From the feature space, each proposed class and its near-close neighbors are sampled. Then, the algorithm makes new instances that incorporate properties from both the purpose case and its neighbors. The goal is to ensure that every class has additional characteristics and that the samples are more inclusive.

**3.2. Machine Learning Methods**

In various fields with different data bases, ML methods are now widely employed. On the other hand, this approach has gained the attention of researchers in healthcare domain. In the next subsection some ML methods in order to predict HCV are described.

**3.2.1. K-Nearest Neighbor (KNN)**

Similar things are assumed to be close to each other by the K-Nearest Neighbor (KNN) algorithm [20-21]. In this algorithm, the parameter k is specified as closest neighbors. Next, the distance between the new data point and the training data is calculated. Then, sorting the minimum distance of the nearest neighbor is done, and the closest neighbor is categorized based on the value of k. At the end, the classification outcomes are defined based on the major label.

**3.2.2. Native Bayesian (NB)**

Bayes' theory serves as the basis for the Naive Bayesian (NB) method. The Bayes' theorem specifies the probability of an event occurring presented the probability of another event that has already occurred.

Assume  $A_3$  and  $A_4$  be the events,  $P(A_3)$  and  $P(A_4)$  be the independent probabilities of  $A_1$  and  $A_2$ ,  $P(A_3 | A_4)$  be the probability of  $A_3$  that is given  $A_4$  be true and  $P(A_4 | A_3)$  be the probability of  $A_4$  that is given  $A_3$  be true. The Bayes theorem is defined as follows [5]:

$$P(A_3 | A_4) = \frac{P(A_4 | A_3)P(A_3)}{P(A_3)} \tag{3}$$

**3.2.3. Decision Tree (DT)**

A collection of predictors is used to predict the values of a target variable using a tree-like graph in the Decision Tree (DT) [22]. Divide-conquer problem solving methods are employed in this method. So, the tree starts with the root node that contains all the data and applies intelligent methods to separate it into multiple branches. The goal is to generate groups that are more similar at the child node.

**3.2.4. Random Forest (RF)**

Random Forest (RF) is used to create decision trees by referring to each tree [23]. By choosing a specific number of data points from the dataset and merging them, the decision is made. In the case of more guesses available, the average of all predictions is applied.

**3.2.5. Support Vector Machine (SVM)**

Support Vector Machine (SVM) is a method that maps variables to an n-dimensional feature space [24]. In order to avoid overfitting, SVM uses a hyperplane to divide the feature space by class when presented with labeled training data. So that, for a sample data set that is presented, SVM determines a new category for any of the labeled classes.

**3.2.6. Boosting Algorithms**

By generating a strong classification from a set of weak classifications, boosting algorithms aim to enhance the predictive performance of models [25]. To refine the errors from the primary model, these algorithms build a model from the training set and then refine it with a second model. Models are added until the training data is accurately predicted or classified. Adaptive Boosting (AdaBoost) algorithm is a method that is employed for binary classification. So that, the algorithm is designed to predict the main data point and give equal weight to all observations. If the first model is unable to predict it, the observation that was wrongly predicted is given a higher weight. Until the threshold accuracy is reached, this process extends and the models are added.

Gradient Boosting Machine (GBM) is considered to complete the prediction from several decision trees. Extreme Gradient Boosting (XGBoost) is an improved version of the GBM algorithm [26]. Furthermore,

XGBoost suggest a different kind of regularization methods that reduce overfitting. Regularization technique can be chosen by improving the hyperparameters of the XGBoost algorithm. Gbtree is chosen as booster during hyperparameter optimization.

**3.3. Cross Validation**

Cross Validation (CV) is applied in order to optimize the parameters of the method [4-5], [8]. The primary goal of CV is a generalization of the ML model for the independent dataset. So that, this technique consists of the split data into different parts and iteratively testing and training on them. K-flod CV is the most common type that is utilized by ML community in order to manipulate the bias-variance trade-off of the ML models. This type of technique is based on the dividing the available dataset into random groups of K. So that, K is the number of repetitions. In each repetition, one group of data will be utilized for testing and other remaining groups will be utilized for training the model. CV has been carried out by selecting the value of the number of folds as 10.

**3.4. Grid Search**

One of the most prominent methods to find hyperparameter configuration space is Grid Search (GS) [5], [27-28]. Furthermore, it can be remarked a Brute-force search approach in order to evaluate all hyperparameter combinations presented to the grid of configurations. GS approach is based on the Cartesian product of user defined finite set of values. Since, the GS approach cannot extract the well-performing sections by itself, thus, in order to determine the global optimums manually, first, consider a large search space and step size, then, limit the search space based on the priors' outcomes of well-performing hyperparameter configurations, next, iterate the previous step until an optimum is achieved.

**4. EVALUATION**

To determine how well the ML model is doing, it is necessary to evaluate which model [29, 30]. The model's evaluation is characterized by accuracy, precision, recall and F1-score. For this aim, the number of True-Positive (TP) refers to the result when the model correctly anticipates the positive class, False-Negative (FN) occurs when the class is positive, but the model anticipates it being negative, False-Positive (FP) occurs when the class is negative, but the model anticipates it to be positive and True Negative (TN) is achieved when the model accurately predicts the negative class. Accuracy is the sufficiency of the model in order to predict true labels. That is determined as follows [30]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

Precision is the model's ability not to predict a positive label that is truly negative. This is measured as [30]:

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

Recall is the ability of the model to predict positive labels. This is defined as follows [30]:

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

The weighted harmonic of precision and recall is known as F1-score. It is determined as follows [30]:

$$F1\text{-score} = \frac{2(Recall \times Precision)}{Recall + Precision} \tag{7}$$

The evaluation of proposed methods with preprocessing, k-fold CV and grid-search CV are illustrated in Figures 15-18. Figure 15 indicates that the KNN method with preprocessing and XGBoost with grid search-CV have the highest accuracy, while the lowest accuracy is associated with Adaboost with preprocessing. The accuracy value decreases in most methods when using k-fold CV and grid-search CV, but in Adabsoot, hyperparameter lead to an increase in accuracy. KNN is associated with the highest value of precision and recall with preprocessing, as demonstrated in Figures 16 and 17. On the other hand, the precision value has increased by applying grid-search CV in AdaBoost and XGBoost.

As shown in Figure 18 the highest and lowest value of F1-score are related to KNN and Adaboost with preprocessing. Figures 19-22 demonstrate comparisons between some methods that are used in this study with those of Wang, et al. [30]. The highest and lowest accuracy values are associated with KNN and AdaBoost method with preprocessing which are indicated in Figure 19. The highest precision value in all method related to Wang, et al. [30] can be seen in Figure 20. The highest and lowest values of recall and F1-score correspond to KNN and AdaBoost with preprocessing which are shown in Figures 21 and 22.

**5. CONCLUSION**

The incidence of Hepatitis C Virus is one of the leading that causes the major of chronic liver disease. On the other hand, the lack of accurate and proper diagnosis of Hepatitis C Virus results in a lack of efficient and timely treatment. Medical diagnosis is a crucial and complex task that necessitates accurate identification. This is where the power of machine learning data analysis capabilities comes into play, where the researchers have applied as non-invasive in order to process patient-related data to effective medical treatments of Hepatitis C Virus. This study applied different kind of machine learning techniques such as K-Nearest Neighbor, Native Bayesian, Decision Tree, Random Forest, Support Vector Machine, AdaBoost and XGBoost based on various evaluation for predicting Hepatitis C Virus. In this study, 98.48% accuracy was achieved with K-Nearest Neighbor with preprocessing and XGBoost with grid search-CV. Furthermore, employing this approach can be assisted in obtaining a timely diagnosis.

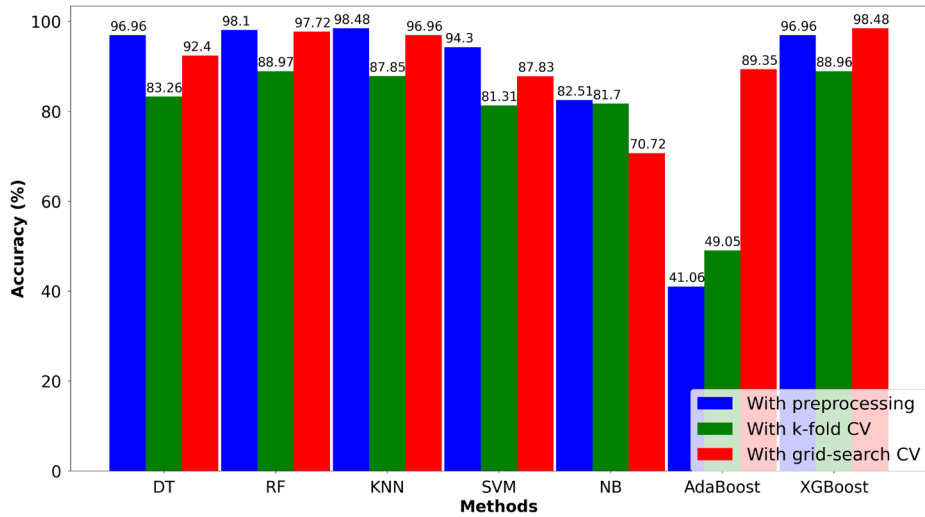


Figure 15. Comparing accuracy between different methods

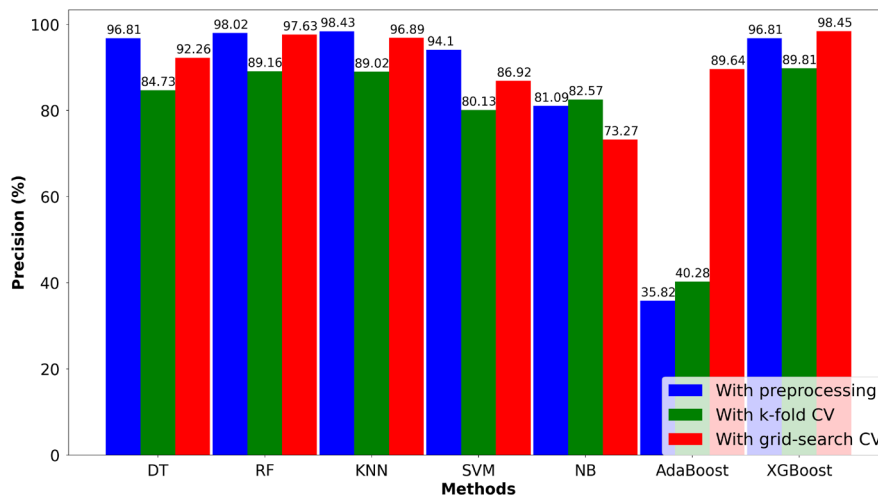


Figure 16. Comparing precision between different methods

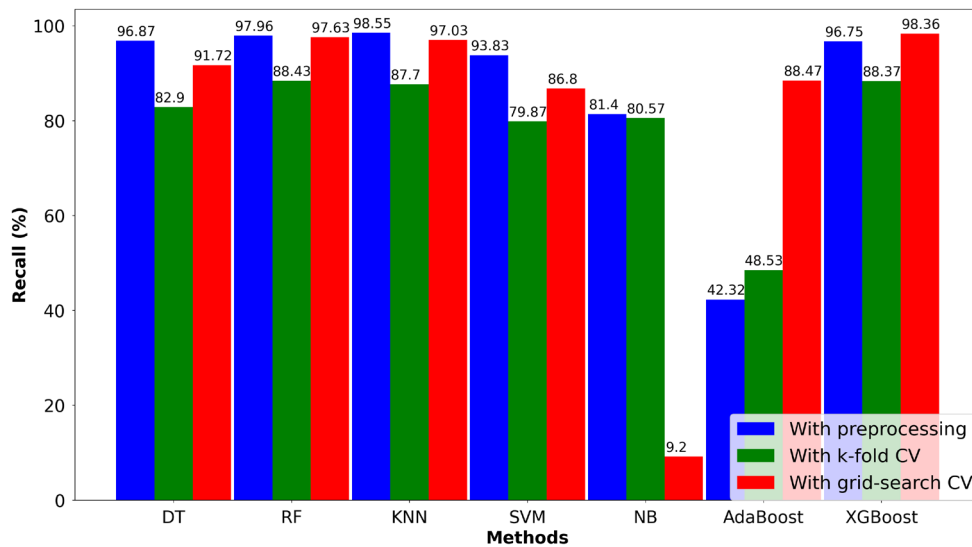


Figure 17. Comparing recall between different methods

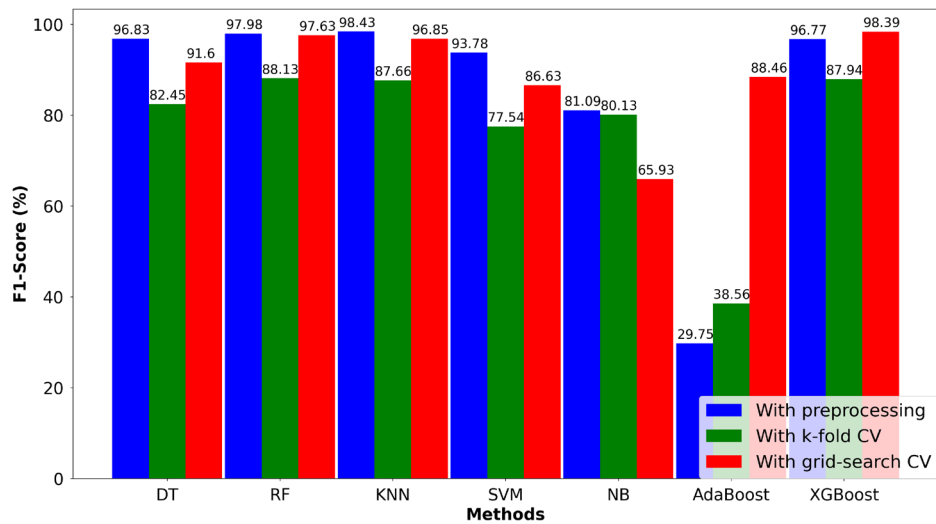


Figure 18. Comparing F1-score between different methods

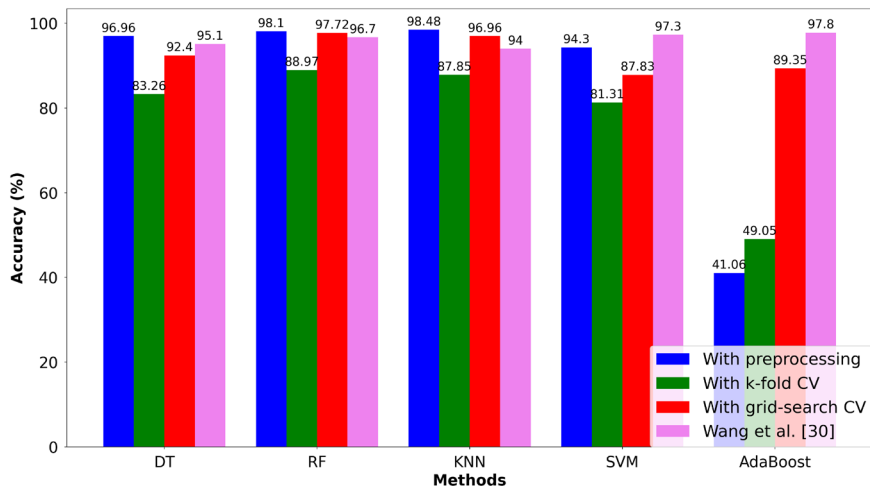


Figure 19. Comparing the accuracy of some methods used in this paper with those of Wang, et al. [30]

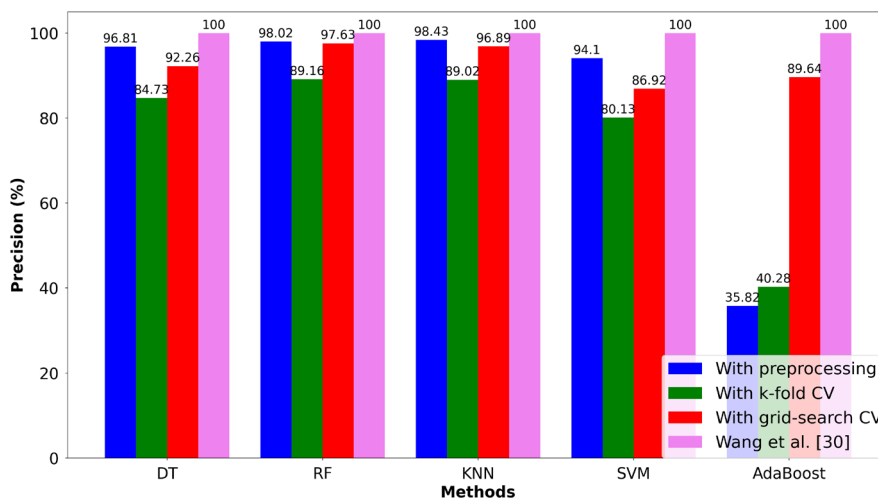


Figure 20. Comparing the precision of some methods used in this paper with those of Wang, et al. [30]

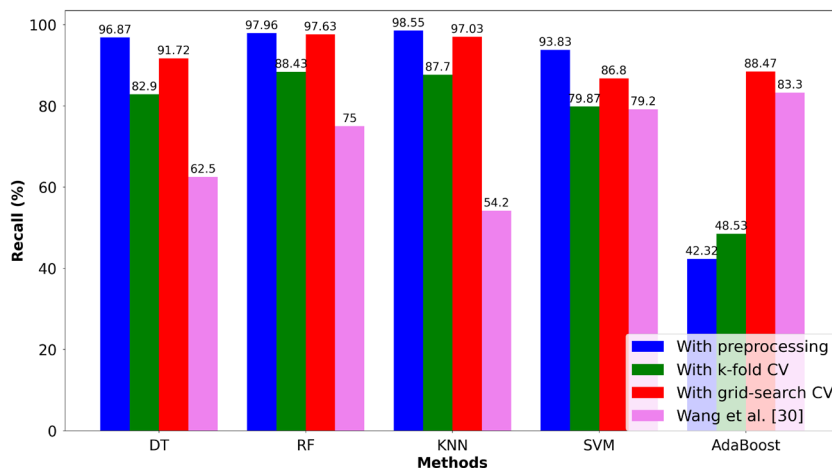


Figure 21. Comparing the recall of some methods used in this paper with those of Wang et al. [30]

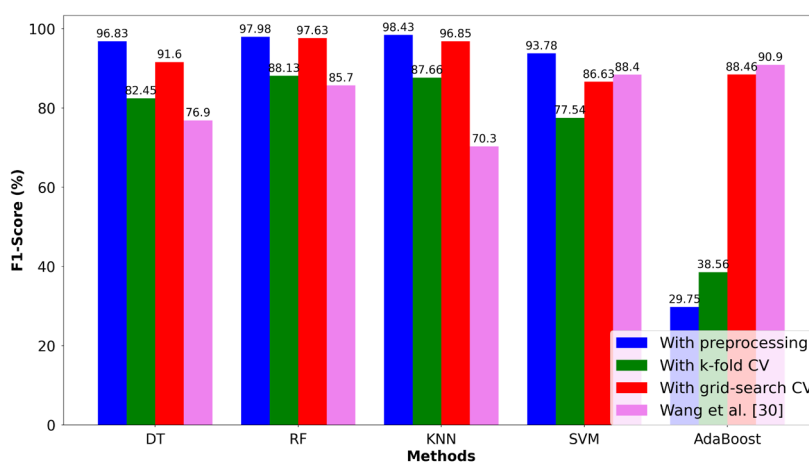


Figure 22. Comparing the F1-Score of some methods used in this paper with those of Wang, et al. [30]

**REFERENCES**

[1] S. Hashem, G. Esmat, W. Elakel, S. Habashy, S.A. Raouf, M. Elhefnawi, M.I. Eladawy, M. El Hefnawi, "Comparison of Machine Learning Approaches for Prediction of Advanced Liver Fibrosis in Chronic Hepatitis C patients", *Journal of IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 15, Issue 3, pp. 861-868, 2017.

[2] C. Brecht, "Hepatitis C Virus, Digestive Diseases and Sciences", Vol. 41, pp. 6S-21S, 1996.

[3] R. Gould, "Toward Data-Scientific Thinking, *Journal of Statistics and Data Science Teaching*", Vol. 43, Issue 51, pp. S11-S22, 2021.

[4] A. Ghasemieh, A. Lloyed, P. Bahrami, P. Vajar, R. Kashef, "A Novel Machine Learning Model with Stacking Ensemble Learner for Predicting Emergency Readmission of Heart-Disease Patients", *Journal of Decision Analytics*, Vol. 7, No. 100242, 2023.

[5] Md. Rifatul Islam, S. Banik, K. Naimur Rahman, M. Mizanur Rahman, "A Comparative Approach to Alleviating the Prevalence of Diabetes Mellitus Using Machine Learning", *Journal of Computer Methods and Programs in Biomedicine Update*, Vol. 4, No. 10013, 2023.

[6] N.A. Mahoto, "Data Mining Techniques for Complex Applications Domain, P.H.D Thesis", Polytechnic of Turin, 2013.

[7] M. Nilashi, O. Ibrahim, H. Ahmadi, L. Shahmoradi, "A Knowledge-Based System for Breast Cancer Classification Using Fuzzy Logic Method", *Journal of Telematics and Informatics*, Vol. 34, Issue 4, pp. 133-144, 2017.

[8] K. Moulaei, H. Sharifi, K. Bahaadinbeigy, A. Haghdoost, N. Nasiri, "Machine Learning for Prediction of Viral Hepatitis: A Systematic Review and Meta-Analysis", *Journal of Medical Informatics*, Vol. 179, No. 105243, 2023.

[9] M. Yaganoglu, "Hepatitis C Virus Data Analysis and Prediction Using Machine Learning", *Journal of Data and Knowledge Engineering*, Vol. 142, No. 102080, 2022.

[10] R. Kumar Sachdeva, P. Bathla, P. Rani, V. Solanki, R. Ahuja, "A Systematic Method for Diagnosis Disease Using Machine Learning", *Journal of Innovations in Systems and Software Engineering*, Vol. 19, pp. 71-80, 2023.

[11] L. Syafaah, Z. Zulfatman, I. Pakaya, M. Lestandy, "Comparison of Machine Learning Classification Methods in Hepatitis C Virus", *Journal of Online Informatika*, Vol. 6, No. 1, pp. 73-78, 2021.



- [12] T.M. Ghazal, M. Anam, M.K. Hasan, M. Hussain, M.S. Farooq, H. Ali, M. Ahmad, T.R. Soomro, "Hep-Pred: Hepatitis C Staging Prediction Using Fine Gaussian SVM", *Journal of Computers, Material and Continua*, Vol. 69, No. 1, pp. 191-203, 2021.
- [13] G.G. Agarwal, A.K. Singh, V. Venkatesh, N. Wal, "Determination of Risk Factors for Hepatitis C by the Method of Random Forest", *Journal of Annals of Infectious Disease and Epidemiology*, Vol. 4, Issue 1, pp. 1-4, 2019.
- [14] S. Das, M. Nayak, MR. Senapati, J. Satapathy, "Medical Data Classification Using Velocity Enhance Whale Optimization Algorithm", *International Conference on Advances in Computing and Future Communication Technologies*, Meerut, India, 2021.
- [15] S.C. Nandi Pati, C. XinYing, K.K. Wah, "Hepatitis C Virus (HCV) Prediction by Machine Learning Technique", *Journal of Applications of Modeling and Simulation*, Vol. 4, pp. 89-100, 2020.
- [16] [www.kaggle.com/datasets/fedesoriano/hepatitis-c-dataset](http://www.kaggle.com/datasets/fedesoriano/hepatitis-c-dataset).
- [17] A. Alamsyah, T. Fadila, "Increased Accuracy of Prediction Hepatitis Disease Using the Application of Principal Component Analysis on a Support Vector Machine", *Journal of Physics Conference Series*, No. 1968, pp. 012-016, 2021.
- [18] T. Menzies, J. Greenwald, A. Frank, "Data Mining Static Code Attributes to Learn Defect Predictors", *Journal of IEEE Transactions on Software Engineering*, Vol. 33, No. 1, pp. 2-13, 2007.
- [19] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegel Meyer, "Smote: Synthetic Minority Over-Sampling Technique", *Journal of Artificial Intelligence Research*, Vol. 16, pp. 321-357, 2002.
- [20] B. Ture Savadkoohi, P. Nik Mohammadi, "Applying Wormhole Approach to Design Hierarchy in a Relational Database for Quick Data Access", *Journal of Technical and Physical Problems of Engineering (IJTPE)*, Issue 47, Vol. 13, No. 2, pp. 144-148, June 2021.
- [21] R. Damarta, A. Hidayat, A. Abdullah, "The Application of K-Nearest Neighbors Classifier for Sentiment Analysis of PT PLN (Persero) Twitter Account Service Quality", *Journal of Physics: Conference Series*, IOP Publishing, Vol. 1722, p. 012002, 2021.
- [22] B. Kaminski, M. Jakubczyk, P. Szufel, "A Framework for Sensitivity Analysis of Decision Trees", *Central European Journal of Operations Research*, Vol. 26, No. 1, pp. 135-159, 2018.
- [23] M. Belgiu, L. Dragut, "Random Forest in Remote Sensing: A Review of Applications and Future Directions", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 114, pp. 24-31, 2016.
- [24] H. Chen, D. Liu, B. Yang, J. Liu, G. Wang, "A New Hybrid Method Based on Local Fisher Discriminant Analysis and Support Vector Machines for Hepatitis Disease Diagnosis", *Journal of Expert Systems with Applications*, Vol. 38, Issue 9, pp. 11796-11803, 2011.
- [25] J. Zheng, D. Lin, Z. Gao, S. Wang, M. He, J. Fan, "Deep Learning Assisted Efficient Adaboost Algorithm for breast Cancer Detection and Early Diagnosis", *IEEE Access*, Vol. 8, pp. 96946-96954, 2020.
- [26] T. Chen, C. Guestrin, "Xgboost: A Scalable Tree Boosting System", *The 22th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016.
- [27] M. Injadat, A. Moubayed, A.B. Nassif, A. Shami, "Multi-Split Optimized Bagging Ensemble Model Selection for Multi-Class Educational Data Mining", *Springer's Applied Intelligence*, 2020.
- [28] F. Hutter, L. Kotthoff, J. Vanschoren, (Eds.), "Automatic Machine Learning: Methods, Systems, Challenges", Springer, ISBN 9783030053185, 2019.
- [29] B. Ture Savadkoohi, "The World of E-Commerce: The Process from Digital Signature up to Market Basket Analysis", *Journal of Technical and Physical Problem of Engineering (IJTPE)*, Issue 52, Vol. 14, No. 3, pp. 202-206, September 2022.
- [30] Y. Wang, B. Yin, Q. Zhu, "Application of Machine Learning Algorithm in Predicting Hepatitis C", *International Symposium on Artificial Intelligence for Medicine Science*, Chengdu, China, 2023.

## BIOGRAPHIES



**Name:** Mahdi

**Surname:** Baghaei Oskuei

**Birthday:** 04.11.1999

**Birthplace:** Tabriz, Iran

**Bachelor:** Software Engineering, Department of Computer and Electrical Engineering, Seraj Higher Education

Institute, Tabriz, Iran, 2022

**Master:** Software Engineering, Department of Computer and Electrical Engineering, Seraj Higher Education Institute, Tabriz, Iran, 2024

**Research Interests:** Machine Learning, Software Engineering, Data Base



**Name:** Bitia

**Surname:** Ture Savadkoohi

**Birthday:** 15.08.1980

**Birthplace:** Tabriz, Iran

**Bachelor:** Software Engineering, Department of Computer and Electrical Engineering, Islamic Azad University,

Shabestar, Iran, 2003

**Doctorate:** Computer Science, Department of Information Engineering and Computer Science, University of Trento, Trento, Italy, 2010

**The Last Scientific Position:** Assist. Prof., Seraj Higher Education Institute, Tabriz, Iran, Since 2012

**Research Interests:** Computer Graphics, Shape Comparison, Analysis of 3D Data, Software Engineering Data Base, Machine Learning

**Scientific Memberships:** 13 Papers