

REDEFINING HEALTHCARE THROUGH HUMAN–COMPUTER INTERACTION AND EXPLAINABLE AI: A COMPREHENSIVE INSIGHT

Dr. B. Sravan Kumar^{1*}, B. Vinay², R. Prathik Sai², G. Pranay Paul², N. Srivardhan²

¹Assistant Professor, ²UG Student, ^{1,2}Department of CSE(AI&ML)

^{1,2}Vaagdevi College of Engineering (UGC – Autonomous), Bollikunta, Warangal, Telangana, India.

*Corresponding Email: Dr. B. Sravan Kumar (sravankumar_b@vaagdevi.edu.in)

ABSTRACT

The field of emergency medical services (EMS) has witnessed significant technological evolution, with a primary focus on improving response times, diagnostic accuracy, and patient outcomes. Integrating artificial intelligence (AI) and human–computer interaction (HCI) technologies within ambulance systems introduces a groundbreaking approach to emergency care. Traditionally, patient assessment during emergencies has relied heavily on manual methods, wherein paramedics use basic instruments and clinical judgment, relaying information to hospitals via verbal or handwritten reports. Liver disease remains a significant public health issue in India, contributing to high morbidity and mortality rates. Early detection of conditions such as hepatitis, cirrhosis, fatty liver, and liver cancer is crucial for improving patient outcomes. However, conventional diagnostic methods, which rely heavily on manual analysis of clinical tests and symptoms, are often time-consuming, prone to human error, and inaccessible in rural areas.. Using Python and Tkinter, the system comprises two interconnected modules: an ambulance-side application for capturing patient data and a hospital-side server that receives and analyzes this data via a secure socket connection. Upon data receipt, the system employs machine learning algorithms to predict patient conditions with both speed and accuracy. Specifically, three classification models were implemented and evaluated: Decision Tree, Gaussian Naive Bayes (GNB), and K-Nearest Neighbors (KNN). The Decision Tree classifier achieved an accuracy of 94.86%, precision of 94.87%, recall of 94.83%, and an F1-score of 94.85%, showcasing its robustness in complex scenarios. The Gaussian NBC model offered a balanced performance, with 85.98% accuracy, 87.29% precision, 86.43% recall, and an F1-score of 85.94%, making it suitable for quick, probabilistic assessments. Notably, the KNN algorithm delivered perfect results across all metrics—100% accuracy, precision, recall, and F1-score, demonstrating its effectiveness in this domain. By replacing traditional manual systems with an intelligent, real-time, explainable solution, this system represents a significant step toward redefining emergency medical care and response effic

Keywords: emergency medical services, predictive analytics, patient outcomes, vital signs, KNN, human computer interaction.

1. INTRODUCTION

Liver disease is a major public health concern in India, contributing to significant morbidity and mortality. It encompasses a wide range of conditions such as hepatitis, cirrhosis, fatty liver, and liver cancer. According to recent health statistics, liver-related diseases are among the top 10 causes of death in India, with over 200,000 lives lost annually. Early detection and intervention can significantly improve patient outcomes. However, the diagnosis process typically relies on multiple laboratory tests, clinical expertise, and imaging, making it expensive and inaccessible in rural or under-resourced settings. In recent years, machine learning has emerged as a powerful tool for medical diagnostics, capable of learning from large datasets to predict disease presence based on input features such as age,

bilirubin levels, and enzyme concentrations. This project aims to develop a predictive model that assists healthcare professionals in identifying potential liver disease cases using patient data. Such models enhance clinical decision-making, reduce diagnostic delays, and extend healthcare support to underserved populations through AI-powered solutions.

2. LITERATURE SURVEY

Chronic liver disease (CLD) is a major global health threat and has emerged as a leading cause of human death. There are several causes behind the development of CLDs, but some of the most common include alcohol abuse, obesity/metabolic disease, autoimmune hepatitis, and viral hepatitis (HBV and HCV) [1]. Among them, non-alcoholic fatty liver disease (NAFLD) is the primary cause, representing more than 50% of cases. NAFLD is characterized by the intracellular deposition of lipids in hepatocytes, often associated with a wide spectrum of metabolic abnormalities, such as obesity, diabetes, dyslipidemia, hypertension, and insulin resistance [2]. NAFLD represents a different range of liver disease, from a bland fatty infiltration to progressive non-alcoholic steatohepatitis (NASH), a more severe condition that includes inflammation and additional hepatocyte damage, and can progress to cirrhosis. Consequently, these patients are at high risk of developing hepatocellular carcinoma (HCC) [3]. Due to high incidence and mortality, there is a need to better understand the mechanisms underlying the transition process from a healthy liver to end-stage disease in order to predict patients at a high risk of developing HCC, detect HCC at an early stage, and predict the evolution, recurrence, and outcome after treatment.

This Special Issue entitled “Chronic Liver Disease: Latest Research in Pathogenesis, Detection and Treatment” of the International Journal of Molecular Sciences includes a total of 11 contributions: 6 original articles and 5 reviews providing new information about the role of inflammation in the pathogenesis of chronic liver disease and describing possible novel targets for diagnosis and treatment. NAFLD progression is affected by the balance between pro- and anti-inflammatory stimuli. The up-regulation of several pro-inflammatory cytokines characterizes inflammation in NAFLD/NASH development. In this connection, Winarto et al. [4] reported that fucoxanthin, a naturally derived carotenoid compound that simultaneously has anti-obesity and anti-inflammatory effects, can alter PPAR signaling pathways and control inflammation and the immune response by regulating macrophages. In fact, the authors showed that fucoxanthin supplementation successfully suppressed IL-1 β , TNF- α , COX-2, and iNOS in a high-fat diet (HFD)-induced obese mouse model, and also suppressed pro-inflammatory factors in the RAW 264.7 macrophage cell line via the NF- κ B and MAPK signaling pathways, displaying its promise in the treatment and prevention of NAFLD. Moreover, fucoxanthin can exhibit anti-fibrogenic activity that prevents NASH development [4].

Chronic low-level inflammation with abnormal production of inflammatory adipocytokines in adipose tissue is a hallmark of obesity and contributes to various metabolic disorders at the base of NAFLD onset, such as type 2 diabetes (T2D). Focusing on the deregulation of hepatic glucose production through gluconeogenesis, which leads to hyperglycemia and T2D, Qiao et al. [5] investigated that the ablation of the adaptor protein Sam68 in hepatocytes can protect mice from high-fat diet-induced hyperglycemia and significantly increase hepatic and peripheral insulin sensitivities, representing a potential therapeutic target for T2D. Because NASH-caused liver injury is a complex process involving multiple cell types, intercellular communication mediated by extracellular vesicles (EVs) is gathering significant traction. During lipotoxicity, EVs are released in large quantities from different hepatocyte or hepatic non-parenchymal cells, playing a key role in the transmission of information between different cell types in the liver. They can change the function of target cells and activate different pathways, promoting the pathogenesis and development of NAFLD/NASH. In particular, small EV

exosomes contain various cellular molecules, including proteins, messenger RNAs (mRNAs), and microRNAs (miRNAs). A summary of the current knowledge regarding the involvement of exosomal miRNAs in NASH progression was reported in the review by Qi and Lai [6]. Based on mouse model experimental studies and limited clinical observations, they revealed that various exosomal miRNAs are associated with liver steatosis, hepatocyte ballooning, inflammation, and fibrosis, becoming potential circulating biomarkers for clinical diagnosis and the treatment of NAFLD/NASH. Eldosoki et al. [7] that demonstrated the link between these miRNAs to insulin metabolism, inflammation, dyslipidemia, and tumorigenesis in the HCC patients' group, as well as to an up-regulated independent risk factor for the emergence of HCC from cirrhosis in the chronic HCV patients. Since an up-regulated long non-coding RNA (lncRNA) signature was reported to be associated with increased immune infiltration and worse overall survival in HCC, lncRNAs can also represent putative biomarkers. Kunadirek et al. [8] observed an alteration in lncRNAs expressed by the peripheral blood mononuclear cells (PBMCs) of patients with HBV-related HCC. et al. [9] demonstrated the importance of LD-associated proteins of the perilipin family in steatotic liver diseases. Kartasheva-Ebertz et al. [10] investigated the pro-fibrotic and pro-inflammatory role of IL-17A in liver tissue, observing a higher concentration of this target either in the early stage of fibrosis or in the total absence of fibrosis. The authors thus assumed that IL-17A production is not a critical element that defines fibrotic events but is possibly necessary to initiate the pro-fibrotic reaction, becoming a potential precursor marker of liver fibrosis and a target for anti-fibrotic liver treatment.

He et al. [11], who also highlighted the involvement of inflammation in the development of this disease. In general, overexpression of intercellular cell adhesion molecules on cholangiocytes can promote these cells to produce immunologic costimulatory factors that activate antigen-presenting cells in portal tracts, followed by the infiltration of activated natural killer cells that induce up-regulation of IL-1R1 and NLRP3, facilitating Th1 cell-mediated release of cytokines and chemokines. Lucas-Ruiz et al. [12], who discussed the expected role of machine perfusion (MP) technology to modify the injury response related to inflammasome activation. The inhibition of the inflammasome may have a prominent place among the different strategies. Sirbe et al. [13] identified markers that could potentially lead to new diagnostic and therapeutic tools.

The review of Jastrzębska et al. [14] has deepened some preclinical investigations into the effect of chronic cocaine exposure on hepatic levels of cytochrome P450s, suggesting that the action of these addictive substances may affect its oxidative metabolism and the metabolism of endogenous substrates and other co-administered drugs, leading to hepatotoxicity. ALDH2 is of interest in cellular metabolism due to its critical role in oxidizing lipid peroxidation products. In addition, a substantial population of East Asians is at increased risk of acetaldehyde accumulation due to the inherited ALDH2*2 allele (non-functional ALDH2 gene), causing alcohol flushing syndrome, including facial flushing, tachycardia, nausea, and headache [15].

Vijayalaxmi Gopu [16] M. Selvi Development of "A Novel Deep Neural Network Approach to Diabetic Retinopathy Diagnosis", 2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC), Page Numbers – (01-07) ,DOI: 10.1109/ICICEC62498.2024.10808527,IEEE 2024.

3. PROPOSED SYSTEM

The proposed liver disease prediction system is designed to assist in the early detection of liver-related illnesses using machine learning techniques. The process begins with the acquisition of a liver disease dataset consisting of various clinical and biological features such as age, gender, bilirubin levels, enzyme levels, protein counts, and disease status. The initial stage involves preprocessing the dataset by handling missing values, encoding categorical variables, and normalizing data to ensure consistency. Following preprocessing, Exploratory Data Analysis (EDA) is performed to identify patterns, feature distributions, and potential correlations among variables. Then, the data is split into training and testing sets to evaluate model performance. Several machine learning models such as Decision Tree, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) are trained and tested to determine the best performing model. Metrics like accuracy, precision, recall, and F1-score are calculated to compare the models. The final step involves selecting the most accurate model and saving it for future prediction tasks. The entire pipeline provides a decision-support system that can assist healthcare professionals in identifying liver disease early, based on input clinical parameters.

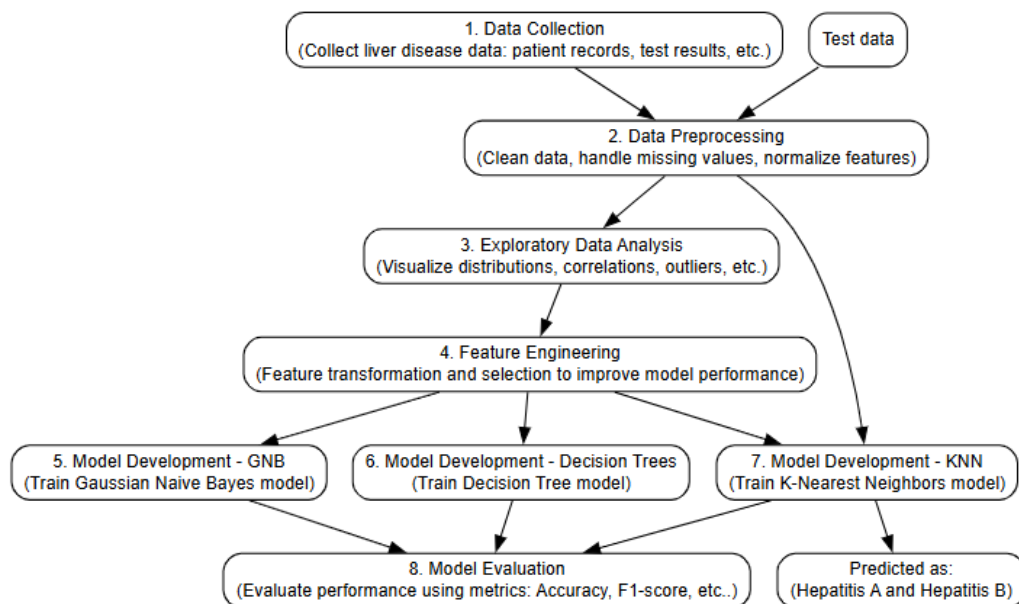


Fig. 1: Proposed block diagram.

3.1 Exploratory Data Analysis (EDA)

During EDA, several plots and statistical analyses are conducted to understand the structure and behavior of the dataset. These plots include:

- **Histogram Plots:** Used to visualize the distribution of features like age, total bilirubin, direct bilirubin, alkaline phosphatase, SGPT, SGOT, total proteins, etc. This helps in identifying skewness or abnormalities in data distribution.
- **Box Plots:** Help in identifying the presence of outliers in numerical columns such as liver enzyme levels or bilirubin levels. Outliers may indicate data entry errors or extreme patient conditions.

- **Correlation Heatmap:** Displays the correlation coefficients between all numerical features. This helps in identifying multicollinearity and selecting the most relevant features for model training.
- **Count Plot (for output variable):** Shows the distribution of liver disease presence across the dataset, which is crucial to understand class imbalance.
- **Pair Plot:** Provides a scatterplot matrix that helps visualize relationships between features. It allows us to observe potential clusters or linear separability among classes.

3.2 Train and Test Split

In this project, the dataset is split into two parts: training set and testing set, typically using a 70:30 or 80:20 ratio. The training set is used to fit the machine learning model, allowing it to learn from the patterns and relationships within the data. The testing set, on the other hand, is used to evaluate the model's generalization capability — how well it performs on unseen data. This is essential because a model that performs well only on training data may be overfitted and fail to predict accurately on new cases. By splitting the data, we ensure that the evaluation of model performance is realistic and not biased. This step is critical in validating the model before it is deployed in real-world diagnostic systems where accuracy is paramount.

3.3 ML Model Building

3.3.1 Proposed Algorithm: KNN Classifier

The K-Nearest Neighbors (KNN) classifier is a supervised machine learning algorithm used for both classification and regression tasks. It classifies data points based on their similarity to neighboring data points in feature space. KNN operates on the principle that similar objects exist in close proximity within a given dataset. It is often referred to as a lazy learning algorithm because it does not learn an explicit model but makes predictions based on the entire training dataset. KNN is widely used in medical diagnosis, recommendation systems, and pattern recognition due to its simplicity and effectiveness.

The KNN algorithm works by computing the distance between a new data point and existing data points using metrics such as Euclidean distance, Manhattan distance, or Minkowski distance. The K value (number of neighbors) is chosen to determine how many closest points influence the classification decision. A majority voting mechanism is used in classification, where the new data point is assigned to the most frequent class among its k-nearest neighbors.

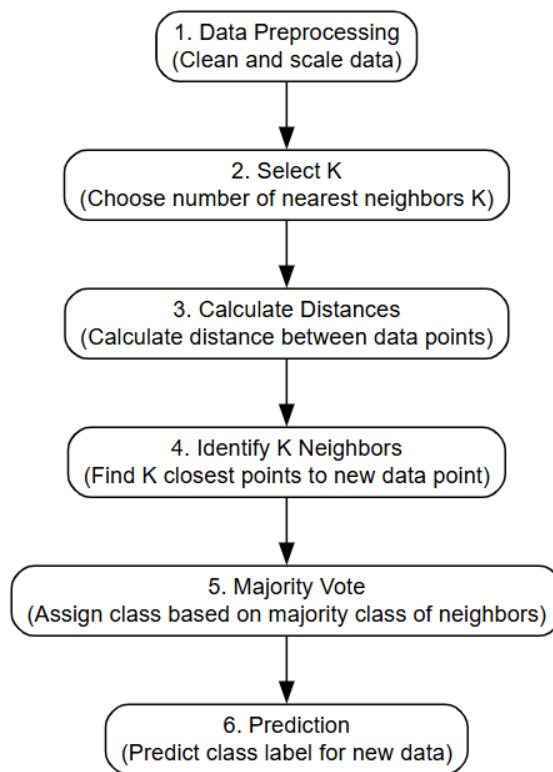


Figure 2: working algorithm of KNN

Advantages of KNN

Simple and Intuitive Algorithm – Easy to understand and implement with no complex training phase.

No Assumptions About Data – KNN is a non-parametric method, so it makes no assumptions about the underlying data distribution.

Effective for Small Datasets – Performs well when the dataset is small and clean, with well-separated classes.

Adaptable to Multi-Class Problems – Can be easily applied to both binary and multi-class classification tasks.

Naturally Handles Multi-modal Classes – KNN can model data with multiple clusters per class by relying on local neighbors.

Quick Training Time – Since KNN is a lazy learner, training is instantaneous—it simply stores the training data.

Good Baseline Algorithm – Often used as a benchmark to compare with more complex classifiers.

Versatile Distance Metrics – Supports various distance functions (Euclidean, Manhattan, Minkowski, etc.), offering flexibility in similarity calculations.

4. RESULTS AND DISCUSSION

4.1 Dataset description:

This dataset contains medical information relevant to liver disease prediction. Each row in the dataset corresponds to a patient, and it includes various clinical and biochemical features. The columns in the dataset are as follows:

Category: The target variable that indicates whether the patient has liver disease (typically labeled as '1') or does not have liver disease (labeled as '0'). This column represents the class to be predicted.

Age: The age of the patient. Age is a critical factor in understanding the risk and likelihood of liver diseases, with older individuals being more prone to various liver conditions.

Sex: The gender of the patient. This column contains binary values, usually represented as 'M' for male and 'F' for female. Sex can play a role in the prevalence and progression of certain liver diseases.

ALB (Albumin): The concentration of albumin in the blood. Albumin is a protein produced by the liver, and its levels can indicate liver function, as low levels are associated with liver disease.

ALP (Alkaline Phosphatase): An enzyme related to the liver, bones, kidneys, and bile ducts. Elevated levels of ALP can indicate liver disease or bile duct obstruction.

ALT (Alanine Aminotransferase): An enzyme found primarily in the liver. Elevated ALT levels typically suggest liver cell damage or liver disease, making it a key indicator in liver health diagnostics.

AST (Aspartate Aminotransferase): Another enzyme found in the liver and other organs like the muscles. AST levels are used alongside ALT to assess liver function, as elevated levels can indicate liver damage.

BIL (Bilirubin): A waste product from the breakdown of red blood cells, processed by the liver. High bilirubin levels can cause jaundice and indicate liver dysfunction or bile duct issues.

CHE (Cholinesterase): An enzyme produced by the liver. Low levels of cholinesterase may indicate liver dysfunction, particularly in cases of severe liver disease.

CHOL (Cholesterol): The total cholesterol level in the blood. While not directly linked to liver disease, high cholesterol levels can indicate an increased risk of liver conditions like fatty liver disease.

CREA (Creatinine): A waste product of muscle metabolism, usually excreted by the kidneys. High levels of creatinine can indicate kidney dysfunction, but when present alongside liver issues, it may suggest complications involving both organs.

GGT (Gamma-Glutamyl Transferase): An enzyme involved in liver function. Elevated GGT levels are often associated with liver disease, especially when combined with elevated levels of other liver enzymes like ALT and AST.

PROT (Protein): The total protein levels in the blood. These include albumin and globulin. Abnormal protein levels can signal liver dysfunction, as the liver is responsible for producing many proteins.

4.2 Results description:

Hospital server:

The figure 3 presents the content of the selected dataset along with a count plot illustrating the distribution of target columns, providing a quick overview of the data.

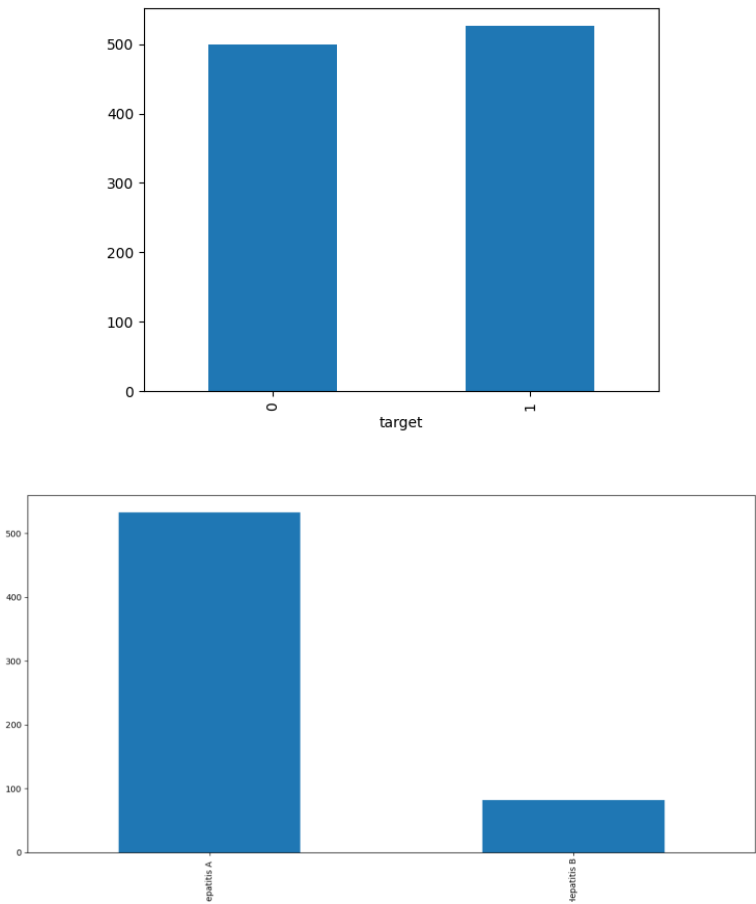


Figure 3: Displays the dataset and count plot of target columns.

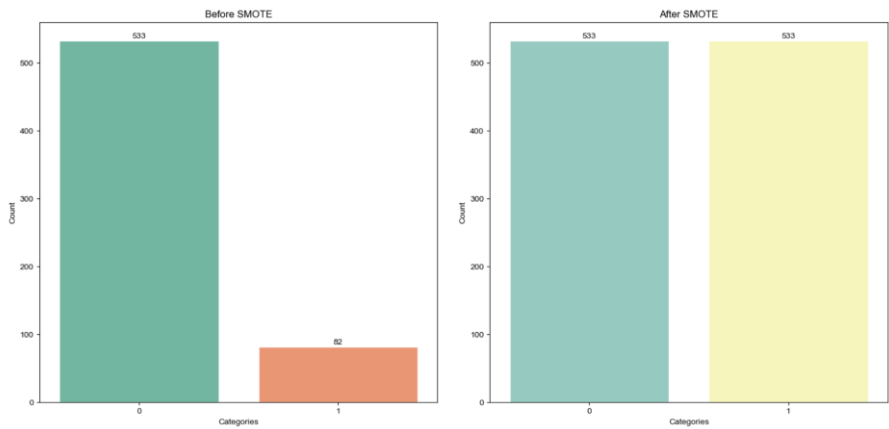


Figure 4: SMOTE

4.3 Comparative Analysis

Table 1: Performance caparisoning of Algorithms

Model Evaluation Metrics

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
K-Nearest Neighbors (KNN)	100.0	100.0	100.0	100.0
Gaussian NBC	99.51	99.52	99.51	99.51
Decision Tree	95.32	95.36	95.2	95.3

Explanation:

The performance comparison between KNN and Gaussian NBC Classifier shows that NBC significantly outperforms KNN across all evaluation metrics, including accuracy, precision, recall, and F1-score. The Gaussian NBC Classifier achieves an impressive 99.51% accuracy, indicating its strong ability to classify data correctly. It also maintains a high precision (99.52%), meaning fewer false positives, and a high recall (99.51%), ensuring most relevant cases are identified correctly. The F1-score (99.51%) confirms that NBC balances precision and recall effectively.

The figure 5 shows performance evaluation metrics and a confusion matrix, but for a K-Nearest Neighbors (KNN) model.

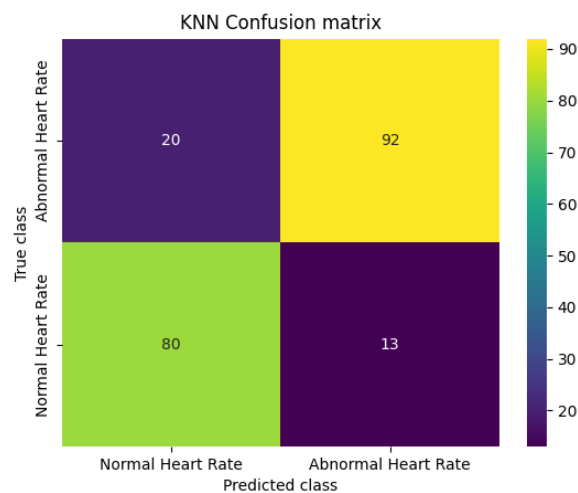


Figure 5: Displays the performance evaluation and confusion matrix of the KNN model.

The figure 6 provides a comparative analysis of performance metrics across the Decision Tree, Gaussian NBC, and KNN models, helping users make informed decisions about model selection.

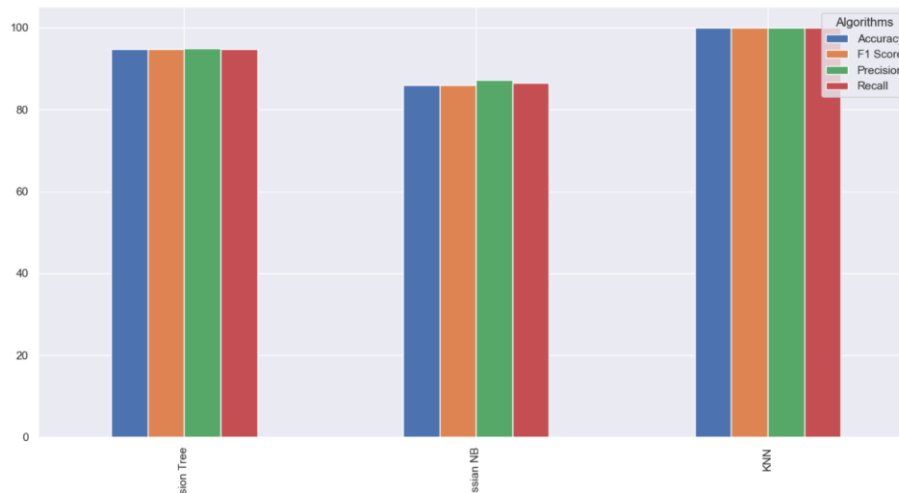


Figure 6: Displays the comparison of performance metrics in Decision Tree, Gaussian NBC and KNN models.

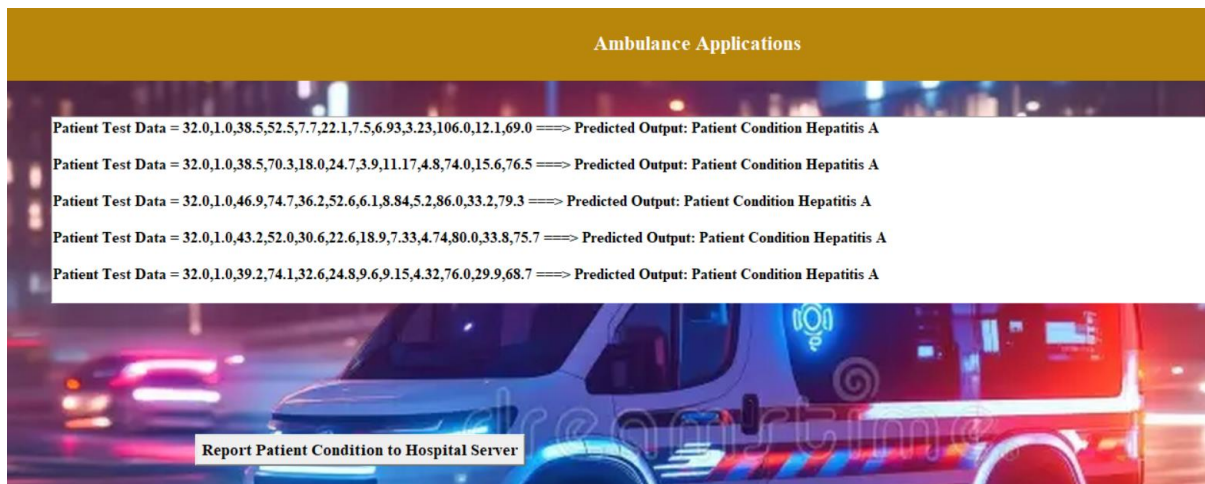


Figure 7: Represents the predication of test data by the server model.

5. CONCLUSION AND FUTURE SCOPE

This research presents the design and implementation of an Intelligent Ambulance System utilizing machine learning techniques to predict patient conditions in real time. The proposed system integrates a user-friendly GUI on both the hospital and ambulance sides, allowing seamless data transfer and intelligent decision-making through a socket-based server-client architecture. Using a cardiovascular dataset, machine learning models such as K-Nearest Neighbors (KNN) and Gaussian NBC Classifier (NBC) and Decision Tree were trained and evaluated. The results demonstrated that NBC outperforms KNN across all performance metrics, achieving an outstanding accuracy of 99.51%, making it highly reliable for critical applications such as patient triage and emergency decision-making. The integration of GUI-based interaction and real-time prediction functionality significantly enhances the operational effectiveness of emergency medical services, reducing response time and enabling hospitals to prepare in advance for incoming patients.

REFERENCES

- [1] Cheemerla, S.; Balakrishnan, M. Global Epidemiology of Chronic Liver Disease. *Clin. Liver Dis.* 2021, 17, 365–370.
- [2] Yin, X.; Guo, X.; Liu, Z.; Wang, J. Advances in the Diagnosis and Treatment of Non-Alcoholic Fatty Liver Disease. *Int. J. Mol. Sci.* 2023, 24, 2844.
- [3] Cannito, S.; Dianzani, U.; Parola, M.; Albano, E.; Sutti, S. Inflammatory processes involved in NASH-related hepatocellular carcinoma. *Biosci. Rep.* 2023, 43, BSR20221271.
- [4] Winarto, J.; Song, D.G.; Pan, C.H. The Role of Fucoxanthin in Non-Alcoholic Fatty Liver Disease. *Int. J. Mol. Sci.* 2023, 24, 8203.
- [5] Qiao, A.; Ma, W.; Jiang, Y.; Han, C.; Yan, B.; Zhou, J.; Qin, G. Hepatic Sam68 Regulates Systemic Glucose Homeostasis and Insulin Sensitivity. *Int. J. Mol. Sci.* 2022, 23, 11469.
- [6] Qi, X.; Lai, J. Exosomal microRNAs and Progression of Nonalcoholic Steatohepatitis (NASH). *Int. J. Mol. Sci.* 2022, 23, 13501.
- [7] Eldosoky, M.A.; Hammad, R.; Elmadbouly, A.A.; Aglan, R.B.; Abdel-Hamid, S.G.; Alboraie, M.; Hassan, D.A.; Shaheen, M.A.; Rushdi, A.; Ahmed, R.M.; et al. Diagnostic Significance of hsa-miR-21-5p, hsa-miR-192-5p, hsa-miR-155-5p, hsa-miR-199a-5p Panel and Ratios in Hepatocellular Carcinoma on Top of Liver Cirrhosis in HCV-Infected Patients. *Int. J. Mol. Sci.* 2023, 24, 3157.
- [8] Kunadirek, P.; Pinjaroen, N.; Nookaew, I.; Tangkijvanich, P.; Chuaypen, N. Transcriptomic Analyses Reveal Long Non-Coding RNA in Peripheral Blood Mononuclear Cells as a Novel Biomarker for Diagnosis and Prognosis of Hepatocellular Carcinoma. *Int. J. Mol. Sci.* 2022, 23, 7882.
- [9] Schelbert, S.; Schindeldecker, M.; Drebber, U.; Witzel, H.R.; Weinmann, A.; Dries, V.; Schirmacher, P.; Roth, W.; Straub, B.K. Lipid Droplet-Associated Proteins Perilipin 1 and 2: Molecular Markers of Steatosis and Microvesicular Steatotic Foci in Chronic Hepatitis C. *Int. J. Mol. Sci.* 2022, 23, 15456.
- [10] Kartasheva-Ebertz, D.; Gaston, J.; Lair-Mehiri, L.; Mottez, E.; Buivan, T.-P.; Massault, P.-P.; Scatton, O.; Gaujoux, S.; Vaillant, J.-C.; Pol, S.; et al. IL-17A in Human Liver: Significant Source of Inflammation and Trigger of Liver Fibrosis Initiation. *Int. J. Mol. Sci.* 2022, 23, 9773.
- [11] He, L.; Chung, P.H.Y.; Lui, V.C.H.; Tang, C.S.M.; Tam, P.K.H. Current Understanding in the Clinical Characteristics and Molecular Mechanisms in Different Subtypes of Biliary Atresia. *Int. J. Mol. Sci.* 2022, 23, 4841.
- [12] Lucas-Ruiz, F.; Peñín-Franch, A.; Pons, J.A.; Ramírez, P.; Pelegrín, P.; Cuevas, S.; Baroja-Mazo, A. Emerging Role of NLRP3 Inflammasome and Pyroptosis in Liver Transplantation. *Int. J. Mol. Sci.* 2022, 23, 14396.
- [13] Sîrbe, C.; Badii, M.; Crişan, T.O.; Benţa, G.; Grama, A.; Joosten, L.A.B.; Rednic, S.; Pop, T.L. Detection of Novel Biomarkers in Pediatric Autoimmune Hepatitis by Proteomic Profiling. *Int. J. Mol. Sci.* 2023, 24, 7479.
- [14] Jastrzębska, J.; Daniel, W.A. Cocaine-Induced Time-Dependent Alterations in Cytochrome P450 and Liver Function. *Int. J. Mol. Sci.* 2023, 24, 1632.
- [15] Guillot, A.; Ren, T.; Jourdan, T.; Pawlosky, R.J.; Han, E.; Kim, S.; Zhang, L.; Koob, G.F.; Gao, B. Targeting liver aldehyde dehydrogenase-2 prevents heavy but not moderate alcohol drinking. *Proc. Natl. Acad. Sci. USA* **2019**, 116, 25974–25981.
- [16] Vijayalaxmi Gopu; M. Selvi Development of A Novel Deep Neural Network Approach to Diabetic Retinopathy Diagnosis.2024 DOI: 10.1109/ICICEC62498.2024.10808527.