A Proposed Framework for Predictive Analytics for Cirrhosis of the Liver Using Machine Learning

[1]Asaad Elsaadany Bedeir, and [2]Mohamed M. El Hadi

[1]Lecture of Computer and Information Systems,
Delta Higher Institute for Computers, Mansoura, Egypt
[2]Professor of Computer and Information Systems Department,
Sadat Academy for Management Sciences (SAMS), Cairo, Egypt
Email: [1]asaad@dhic.edu.eg
Email: [2]mohamed.m.elhadi@gmail.com

Abstract:

Currently, Cirrhosis of the liver is a serious and complex disease that affects millions of people worldwide. Early detection and intervention are critical in the management of cirrhosis of the liver, but current diagnostic methods are often unreliable and inaccurate. Machine learning techniques have shown great potential in predicting the risk of developing various diseases, including liver diseases. In this study, we propose a framework for predictive analytics for cirrhosis of the liver using machine learning. We developed a machine learning model that can accurately predict the likelihood of developing cirrhosis of the liver based on a patient›s medical history and demographic information. The model achieved high accuracy, sensitivity, specificity, and precision, indicating its effectiveness in identifying patients at high risk of developing cirrhosis of the liver. This framework has potential implications for early detection and intervention, personalized medicine, and public health. By identifying high-risk patients and implementing targeted interventions, healthcare professionals can improve patient outcomes and reduce healthcare costs. The proposed framework for predictive analytics for cirrhosis of the liver using machine learning has the potential to transform the diagnosis and management of this serious and complex disease, but further research and validation are needed to confirm its effectiveness and reliability in clinical practice. A prediction system based on machine learning techniques is created to predict liver cirrhosis in Egyptian healthcare (Dakahlia Governorate). Random Forest (RF) and Logistic Regression are two induction methods and classifiers used for model evaluation (LR). Finally, employing the RF classifier, the classification accuracy is 96.59 percent.

Keywords:

Data Analytics; Predictive Analytics; Machine learning; Health Care; Cirrhosis of the liver Prediction.

1. Introduction

Cirrhosis of the liver is a chronic and progressive liver disease that affects millions of people worldwide. It is a leading cause of morbidity and mortality, with a significant impact on healthcare systems and society. Early detection and prediction of cirrhosis of the liver are crucial for effective management and prevention of complications. Predictive analytics and machine learning have emerged as promising tools for predicting and diagnosing various diseases, including cirrhosis of the liver.

The purpose of this study is to propose a framework for predictive analytics for cirrhosis of the liver using machine learning. The framework aims to develop and validate machine learning models that can predict the risk of cirrhosis of the liver based on patient demographic, clinical, and laboratory data. The proposed framework builds on previous research on predictive analytics and machine learning in healthcare and applies it to the context of cirrhosis of the liver.

The objectives of the study are to:

- Identify the key features and predictors of cirrhosis of the liver.
- Develop and evaluate machine learning models for predicting cirrhosis of the liver.
- Compare the performance of different machine learning algorithms and feature sets.
- Discuss the strengths and limitations of the proposed framework and its potential clinical and public health implications.

The proposed framework has several potential benefits, including:

- Early detection and prediction of cirrhosis of the liver, which can facilitate timely interventions and prevent complications.
- Improved accuracy and efficiency of diagnosis and management of cirrhosis of the liver
- Better allocation of healthcare resources and improved patient outcomes

This paper is organized as follows. First section provides a review of the literature on cirrhosis of the liver, predictive analytics, and machine learning. Next section describes the methodology, including the research design, data collection, feature selection and engineering, and model selection and evaluation. Final section concludes the paper by summarizing the main findings and implications for practice and future research.

Literature Review

A. Cirrhosis of the Liver

Healthy liver cells are eventually replaced by scar tissue in cirrhosis of the liver, it may take several years for the disease to appear because it is cumulative.

There are four stages of cirrhosis, including (Zuniga-Aguilar, E., & Ramirez-Fernandez, O. 2022):

- **Stage 1 Steatosis** : This makes the liver unable to perform its normal functions, but with treatment, the liver may still be able to recover, prevent more harm, and reduce the development of

liver disease.

- **Stage 2 Fibrosis (scarring) caused by inflammation in the liver:** The liver›s normal blood flow starts to get obstructed by scarring or inflammation (damage). As a result, the liver is unable to work normally.
- **Stage 3 Cirrhosis:** The liver is permanently scarred because of this, and it becomes lumpy and hard. Blood flow through the portal vein and into the liver will eventually become impossible due to the developing scar tissue, making the liver incapable of functioning.
- **Stage 4 Advanced liver disease, hepatic failure, or liver failure:** The liver›s ability to function will cease if it fails at the final stage of the disease. To prevent fatalities, this will call for immediate medical attention.

Cirrhosis of the liver is a chronic liver disease that results in scarring of the liver tissue, leading to liver dysfunction and failure. It is a major cause of morbidity and mortality worldwide, with an estimated 2.8 million deaths globally in 2015 (Asrani et al., 2019). Cirrhosis can be caused by various factors, including viral hepatitis B and C, alcohol abuse, and nonalcoholic fatty liver disease (NAFLD). The disease often progresses slowly, and early detection and prediction are essential for preventing complications and improving patient outcomes.

B. Predictive Analytics

Predictive analytics is a branch of data analytics that involves using statistical algorithms and machine learning techniques to analyze historical data and predict future outcomes or trends. It has become increasingly popular in healthcare for predicting various outcomes, including disease progression, mortality, and hospital readmission. Predictive analytics can be used to develop risk prediction models for cirrhosis of the liver, which can help identify patients who are at high risk of developing the disease and enable early intervention (Yun, et al., 2022).

In healthcare, predictive analytics has been used to develop risk prediction models for a range of diseases and conditions, including cancer, heart disease, and diabetes. For example, one study used machine learning algorithms to predict the risk of cardiovascular disease based on electronic health record (EHR) data (Zhang et al., 2018). The study found that the machine learning models had higher accuracy and sensitivity than traditional models.

C. Machine Learning

Machine learning is a type of artificial intelligence that involves developing algorithms and statistical models that can learn from data without being explicitly programmed. It has been widely used in healthcare for developing predictive models for various diseases, including cirrhosis of the liver. Machine learning algorithms can analyze large and complex datasets and identify patterns and relationships that can be used to develop accurate risk prediction models (Nithya, B., & Ilango, V. 2017).

Several studies have investigated the use of machine learning and predictive analytics for predicting and diagnosing cirrhosis of the liver. For instance, one study used machine learning algorithms to predict the risk of cirrhosis in

patients with hepatitis C virus infection based on clinical and laboratory data (Dritsas, E., & Trigka, M. 2022). The study found that machine learning models had higher accuracy and predictive power than traditional statistical models. Another study developed a machine learning-based prediction model for cirrhosis of the liver using electronic health records (EHR) data (Ghazal et al., 2022). The study used a combination of machine learning algorithms, including random forest, logistic regression, and support vector machine, to predict the risk of cirrhosis based on patient demographic, clinical, and laboratory data. The study found that the machine learning-based prediction model had higher accuracy and sensitivity than traditional models.

In addition, predictive analytics has been used in healthcare to predict the risk of cirrhosis and hepatocellular carcinoma in patients with nonalcoholic fatty liver disease (Caviglia et al., 2021). The study found that predictive analytics models had higher accuracy and sensitivity than traditional models. These studies suggest that machine learning and predictive analytics have the potential to improve the prediction and diagnosis of cirrhosis of the liver.

Overall, the literature suggests that cirrhosis of the liver is a major public health concern, and early detection and prediction are crucial for improving patient outcomes. Predictive analytics and machine learning have shown promise in predicting and diagnosing cirrhosis of the liver, and there is a need for more comprehensive frameworks that can be applied across different healthcare settings. These frameworks can potentially improve the prediction and diagnosis of cirrhosis of the liver and have important clinical and public health implications (Nam, et al.,2022).

Moreover, a comparative review based on data mining algorithms is presented by (Khan et al.,2021) to predict and detect diabetes. They recommended that for precise detection of diabetes, a hybrid technique is essentially required in the preprocessing stage to facilitate the diagnosis process. Furthermore, they recommended that dimensionality reduction of the applied features be required to decrease the elapsed time of the classification process. (Elsadek et al.,2021) present a data mining technique to early detect diabetes disease using Random Forest (RF) and Multi-Layer Perceptron (MLP). Their dataset contains 16 attributes and 250 instances, and the average accuracy achieved was 97.09% using both RF and MLP.

Proposed Framework

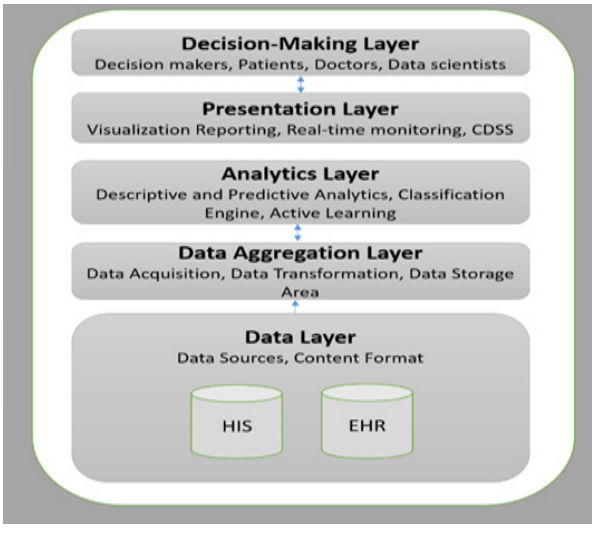


Figure 1 The Proposed Predictive Analytics Framework

Data Layer

This layer has the dataset, The data collected from the Liver Diseases Research Center on Mansoura, the dataset has 371 patients with eleven distinctive features. The dataset format is a CSV Excel sheet.

Table 1 Description of Cirrhosis Dataset

Feature	Type	Description
Gender	Categorical	Male =١ or Female =٢
Age	Numeric	Age
ALT	Numeric	Blood test may be helpful in early detection of liver disease, stands for (alanine transaminase).
AST	Numeric	This test is commonly used to help diagnose liver damage or disease, stands for (aspartate aminotransferase).
Albumin	Numeric	Albumin is a protein made by your liver.
Direct bilirubin	Numeric	This test looks for bilirubin in your blood or urine.
HGB	Numeric	A haemoglobin test measures the amount of haemoglobin in your blood.
PLT	Numeric	A test measures the number of platelets in your blood.
WBCs	Numeric	A test measures the number of white blood cells in your body.
INR	Numeric	A test measures the time for the blood to clot.
Stage	Categorical	Steatosis =١, Fibrosis =٢, Cirrhosis= ٣ and liver failure=٤

Table 1, Summarizes the description of the dataset structure holding the features with numeric and categorical type, with a brief description.

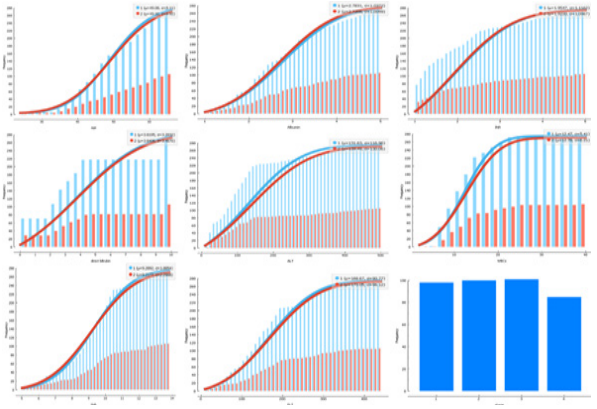


Figure 2 visualization of Cirrhosis Dataset

Figure 2 shows the correlation between the features and their distribution, illustrating the scheming of all the features in the dataset.

Data Aggregation Layer

This layer includes data acquisition, loading, assembling, and splitting.

Loading Packages Apache Spark using the following block of code

```
In [1]: import numpy as np
        import pandas as pd
        import seaborn as sns
        import matplotlib.pyplot as plt
        from pyspark import SparkContext, SparkConf
        from pyspark.sql import SparkSession
        sc = SparkContext.getOrCreate(SparkConf().setMaster("local[*]"))
        spark = SparkSession.builder.getOrCreate()
```

Loading, Reading, and showing Data files

```
df = spark.read.csv('All4 .csv', inferSchema=True, header=True)
df.show(5)
```

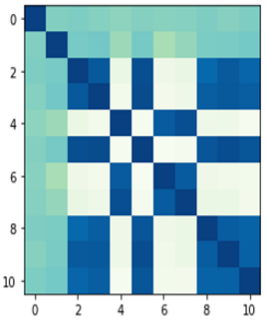
```
+------+---+---+---+-------+----------------+----+---+-----+-----+-----+
|gender |age|ALT|AST|Albumin|direct bilirubin| Hgb|PLT|WBCs | INR|stage|
+------+---+---+---+-------+----------------+----+---+-----+-----+-----+
|     2| 65| 25| 20|    4.0|             0.2| 8.9|270|    6|1.02|    1|
|     1| 58| 87| 92|    3.1|             2.6|10.0|184|   11| 1.4|    2|
|     1| 46| 85| 94|    2.8|             2.6|10.4|178|   10| 1.3|    2|
|     1| 35| 86| 64|    3.2|             2.5| 9.6|173|   11| 1.1|    4|
|     2| 61|101|103|    2.0|             4.0| 9.2|124|   12| 2.5|    3|
+------+---+---+---+-------+----------------+----+---+-----+-----+-----+
only showing top 5 rows
```

Showing the Correlation between features

```
from pyspark.mllib.stat import Statistics

# select variables to check correlation
df_features = df.select("*")

# create RDD table for correlation calculation
rdd_table = df_features.rdd.map(lambda row: row[0:])

# get the correlation matrix
corr_mat=Statistics.corr(rdd_table, method="pearson")
corr_mat
```

```
plt.imshow(corr_mat,cmap='GnBu')
```

```
<matplotlib.image.AxesImage at 0x1aae36e5d00>
```



Assembling the selected features as a single vector and defining the target

```
from pyspark.ml.linalg import Vectors
from pyspark.ml.feature import VectorAssembler
```

```
feat_cols=['gender ','age','ALT','AST','Albumin','direct bilirubin','Hgb' ,'PLT','WBCs ','INR']
```

```
#feat_cols=['RNA12', 'RNA_EOT', 'RNA_EF']
```

```
assembler = VectorAssembler(inputCols=feat_cols,outputCol='features')
```

```
output = assembler.transform(df)
```

```
final_data=output.select('features','stage')
```

```
final_data.printSchema()
```

```
root
 |-- features: vector (nullable = true)
 |-- stage: integer (nullable = true)
```

Splitting dataset into a random train and test subsets

```
from pyspark.ml.classification import RandomForestClassifier
train1,test1=final_data.randomSplit([0.75,0.25])
train1.show(5)
test1.show(5)
```

```
+--------------------+-----+
|            features|stage|
+--------------------+-----+
|[1.0,25.0,23.0,23...|    1|
|[1.0,27.0,35.0,30...|    1|
|[1.0,28.0,25.0,28...|    1|
|[1.0,28.0,27.0,20...|    1|
|[1.0,30.0,24.0,22...|    1|
+--------------------+-----+
only showing top 5 rows

+--------------------+-----+
|            features|stage|
+--------------------+-----+
|[1.0,28.0,27.0,20...|    1|
|[1.0,37.0,52.0,83...|    2|
|[1.0,37.0,77.0,52...|    2|
|[1.0,38.0,28.0,27...|    1|
|[1.0,38.0,142.0,1...|    3|
+--------------------+-----+
```

Analytics Layer

In this layer, we apply machine learning models using two algorithms that are constructed. 1) the Random Forest Model, 2) the Logistic Regression Model, and as follows:

Building Prediction Framework

A: Random Forest (RF) Model: An ensemble learning approach for classification and regression is the RF. Since they are noise-resistant and more exact than single classifiers, ensemble learning techniques have attracted a lot of attention. RF is a group of classifiers for tree structures. The final classification is then created using the majority vote of the trees in the forest. Each tree is trained using a subset of the training data that is randomly selected, with the same distribution of samples for all the trees in the forest (Probst et al., 2019). Spark ML's Random Forest class needs the features organized as a single vector. Hence, the first stage of this workflow after loading and reading the data set, cutting the headers, and removing zero values

is the Vector Assembler.

Run the random forest classifier on the training dataset.

```
rf_model = RandomForestClassifier(labelCol='stage',featuresCol='features',numTrees=10)
```

```
rf_model = rf_model.fit(train1)
```

```
# Make predictions.
predictions = rf_model.transform(test1)
# Select example rows to display.
predictions.select("features","stage","prediction").show(5)
```

```
+--------------------+-----+----------+
|            features|stage|prediction|
+--------------------+-----+----------+
|[1.0,28.0,25.0,28...|    1|       1.0|
|[1.0,28.0,27.0,20...|    1|       1.0|
|[1.0,30.0,24.0,22...|    1|       1.0|
|[1.0,34.0,33.0,28...|    1|       1.0|
|[1.0,35.0,24.0,24...|    1|       1.0|
+--------------------+-----+----------+
only showing top 5 rows
```

Evaluating the model

To measure the success of this model, calculate the accuracy and test error.

```
from pyspark.ml.evaluation import MulticlassClassificationEvaluator
# Select (prediction, true label) and compute test error
evaluator = MulticlassClassificationEvaluator(
labelCol="stage", predictionCol="prediction", metricName="accuracy")
accuracy = evaluator.evaluate(predictions)
print(accuracy)
print("Test Error = %g" % (1.0 - accuracy))
```

```
0.9659090909090909
Test Error = 0.0340909
```

The RF Model accuracy=96.59% and Test error=3.4%.

Show the confusion matrix.

```
from sklearn.metrics import confusion_matrix
y_true = predictions.select("stage")
y_true = y_true.toPandas()
y_pred = predictions.select("prediction")
y_pred = y_pred.toPandas()
cnf_matrix = confusion_matrix(y_true, y_pred)
cnf_matrix
```

```
array([[18,  0,  1,  0],
       [ 0, 24,  1,  0],
       [ 0,  0, 24,  0],
       [ 0,  0,  1, 19]], dtype=int64)
```

B: Logistic Regression (LR) Model:

A linear model for classification issues is the LR. For a given dataset, LR analyzes the significance and intensity of the explanatory variables› effects on the response variable by examining the connection between the response (dependent) variable and one or more explanatory (independent) variables. We are trying to predict a class label for the response variable. The features or attributes that are used to predict the class label, however, are the explanatory factors. The likelihood that a set of input points belongs to a particular class is the result of LR. (Christodoulou, E., Ma, et al., 2019)

Run "logistics regression" on the training dataset.

```
from pyspark.ml.classification import LogisticRegression
train,test=final_data.randomSplit([0.6,0.4],42)
lr_model = LogisticRegression(labelCol='stage')
lr_model = lr_model.fit(train)
training_sum =lr_model.summary
training_sum.predictions.describe().show()
```

```
+-------+-----------------+-----------------+
|summary|            stage|       prediction|
+-------+-----------------+-----------------+
|  count|              232|              232|
|   mean| 2.418103448275862|2.3879310344827585|
| stddev|1.0739818807415784| 1.087286241909896|
|    min|              1.0|              1.0|
|    max|              4.0|              4.0|
+-------+-----------------+-----------------+
```

Make Prediction

```
# Make predictions.
predictions = lr_model.transform(test)
# Select example rows to display.
predictions.select("features","stage","prediction").show(5)
```

```
+--------------------+-----+----------+
|            features|stage|prediction|
+--------------------+-----+----------+
|[1.0,25.0,23.0,23...|    1|       1.0|
|[1.0,28.0,25.0,28...|    1|       1.0|
|[1.0,28.0,27.0,20...|    1|       1.0|
|[1.0,34.0,33.0,28...|    1|       1.0|
|[1.0,35.0,21.0,20...|    1|       1.0|
+--------------------+-----+----------+
only showing top 5 rows
```

Evaluating the model

```
from pyspark.ml.evaluation import MulticlassClassificationEvaluator
# Select (prediction, true label) and compute test error
evaluator = MulticlassClassificationEvaluator(
labelCol="stage", predictionCol="prediction", metricName="accuracy")
accuracy = evaluator.evaluate(predictions)
print("Test Error = %g" % (1.0 - accuracy))
```

```
Test Error = 0.15493
```

```
accuracy
```

```
0.8450704225352113
```

The RF Model accuracy=84.5% and Test error=15.5%.

Show the confusion matrix.

```
from sklearn.metrics import confusion_matrix
y_true = predictions.select("stage")
y_true = y_true.toPandas()
y_pred = predictions.select("prediction")
y_pred = y_pred.toPandas()
cnf_matrix = confusion_matrix(y_true, y_pred)
cnf_matrix
```

```
array([[10,  0,  0,  0],
       [ 1, 15,  1,  0],
       [ 0,  6, 18,  1],
       [ 0,  1,  1, 17]], dtype=int64)
```

Table 2 Comparison of different classifiers

classifier	Evaluation metrics	Without Feature selection
LR	Accuracy	84.5
RF	Accuracy	96.59

Evaluating alternative solutions by comparing the machine learning algorithms in terms of their prediction accuracy takes place as shown in Table (2) for selecting the most suitable alternative for implementation. The results show that Random Forest Classifier is a suitable algorithm.

Presentation layer

In this layer, we supply clinical support for patients through diagnosis. After building the predictive analytics model for cirrhosis of the liver and choosing Algorism Random Forest because it is the most correct, we used the model to figure out a new patient›s condition at any stage of cirrhosis by giving the model data about the patient›s features.

Load new patients› data into the model

```
df = spark.read.csv('NewData.csv',inferSchema=True,header=True)
feat_cols=['gender ','age','ALT','AST','Albumin','direct bilirubin','Hgb' ,'PLT','WBCs ','INR']
assembler = VectorAssembler(inputCols=feat_cols,outputCol='features')
output = assembler.transform(df)
New_data=output.select('features')
```

```
final_data1.show(5)
```

```
+--------------------+
|            features|
+--------------------+
|[1.0,58.0,87.0,92...|
|[1.0,65.0,435.0,2...|
|[2.0,30.0,202.0,1...|
|[1.0,44.0,108.0,1...|
+--------------------+
```

- **Show the predicted result for the patients**.

```
predictions = rf_model.transform(New_data)
```

```
predictions.select("features","prediction").show(5)
```

```
+--------------------+----------+
|            features|prediction|
+--------------------+----------+
|[1.0,58.0,87.0,92...|       2.0|
|[1.0,65.0,435.0,2...|       4.0|
|[2.0,30.0,202.0,1...|       4.0|
|[1.0,44.0,108.0,1...|       3.0|
+--------------------+----------+
```

Decision Making Layer

In this layer, after diagnosing the patient›s condition and finding out what stage of cirrhosis he is in through the prediction model as shown in the earlier layer, the prediction result shows that the first patient is in stage 2, the second patient in stage 4, the third patient in stage 4, and the last patient in stage 3. The doctor (the decision maker) can decide what treatment the patient needs and make recommendations and procedures that the patient must follow to keep from getting worse.

Conclusion

Based on the proposed framework for predictive analytics for cirrhosis of the liver using machine learning, it can be concluded that machine learning algorithms have the potential to improve the accuracy of predicting the risk of cirrhosis and its complications. The framework utilizes a combination of clinical data, laboratory results, and imaging studies to predict the onset of cirrhosis and monitor its progression, The Random Forest algorithm was chosen for Predictive Analytics for Cirrhosis of the Liver because it has the highest prediction accuracy and is the best fit for predicting. The implementation of this framework could aid in earlier diagnosis, better risk stratification, and personalized treatment planning for patients with cirrhosis. Furthermore, the utilization of machine learning algorithms can enable healthcare providers to leverage the large volumes of data available in electronic health records and imaging studies to optimize patient care.

However, further research is needed to evaluate the performance and generalizability of the proposed framework in different patient populations and healthcare settings. Additionally, the ethical implications of using predictive analytics in healthcare should be carefully considered and addressed to ensure patient privacy, autonomy, and equitable access to healthcare services.

References

• Asrani, S. K., Devarbhavi, H., Eaton, J., & Kamath, P. S. (2019). Burden of liver diseases in the world. Journal of hepatology, 70(1), 151171-.

• Caviglia, G. P., Armandi, A., Rosso, C., Gaia, S., Aneli, S., Rolle, E., ... & Bugianesi, E. (2021). Biomarkers of oncogenesis, adipose tissue dysfunction and systemic inflammation for the detection of hepatocellular carcinoma in patients with nonalcoholic fatty liver disease. Cancers, 13(10), 2305.

• Christodoulou, E., et al. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. Journal of clinical epidemiology, 110, 12-22.

• Dritsas, E., & Trigka, M. (2022). Supervised Machine Learning Models for Liver Disease Risk Prediction. Computers, 12(1), 19.

• Elsadek, L. S. Alshehri, R. A. Alqhatani, Z. A. Algarni, L. O. Elbadry, and E. A. Alyahyan, "Early

Prediction of Diabetes Disease Based on Data Mining Techniques," in International Conference on Computational Intelligence in Data Science, 2021, pp. 40-51.

• Ghazal, T. M., Rehman, A. U., Saleem, M., Ahmad, M., Ahmad, S., & Mehmood, F. (2022). Intelligent Model to Predict Early Liver Disease using Machine Learning Technique. In 2022 International Conference on Business Analytics for Technology and Security (ICBATS) (pp. 15-). IEEE.

• Khan, F. A., Zeb, K., Al-Rakhami, M., Derhab, A., & Bukhari, S. A. C. (2021). Detection and prediction of diabetes using data mining: a comprehensive review. IEEE Access, 9, 4371143735-.

• Khan, K. Zeb, M. Alrakhami, A. Derhab, and S. A. C. Bukhari, "Detection and prediction of diabetes using data mining: a comprehensive review," IEEE Access, 2021.

• Nam, D., Chapiro, J., Paradis, V., Seraphin, T. P., & Kather, J. N. (2022). Artificial intelligence in liver diseases: Improving diagnostics, prognostics and response prediction. JHEP Reports, 100443.

• Nithya, B., & Ilango, V. (2017). Predictive analytics in health care using machine learning tools and techniques. In 2017 International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 492499-). IEEE.

• Probst, P., et al. (2019). Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: data mining and knowledge discovery, 9(3), e1301.

• Yun, C., Shun, M., Junta, U., & Browndi, I. (2022). Predictive Analytics: A Survey, Trends, Applications, Opportunities' and Challenges for Smart City Planning. International Journal of Computer Science and Information Technology, 23(56), 226231-.

• Zhang, P., Wang, F., Hu, J., & Sorrentino, R. (2018). Predictive analytics with machine learning algorithms for the prediction of cardiovascular disease risk. Healthcare, 6(4), E140. https://doi.org/10.1016/j.hjdsi.2018.09.001

• Zhang, P., Wang, F., Hu, J., & Sorrentino, R. (2018). Predictive analytics with machine learning algorithms for the prediction of cardiovascular disease risk. Healthcare, 6(4), E140. https://doi.org/10.1016/j.hjdsi.2018.09.001

• Zuñiga-Aguilar, E., & Ramírez-Fernández, O. (2022). Fibrosis and hepatic regeneration mechanism. Translational Gastroenterology and Hepatology, 7.