

Implementation of Liver Disease Prediction Using Machine Learning

Boggarapu Sai Surya, Nitesh Kumar Singh, S Sasi Rekha

School of Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab, India

ABSTRACT

Article Info

Volume 7, Issue 2

Page Number: 606-616

Publication Issue :

March-April-2021

Article History

Accepted : 20 April 2021

Published : 28 April 2021

This work Liver Disease Prediction Using Machine Learning is a machine learning application. In this project, you predict whether the patient contain a liver disease or not using python Jupyter Notebook. To predict presence of liver disease we apply some of the classification techniques. It gives an idea of how machine learning helps in medical field and how classification techniques going to predict liver disease using liver disease data set.

Keywords : Chronic diseases, Classification schemes, Training datasets, Machine learning, Classifiers, Algorithms, Classification models

I. INTRODUCTION

The digital technological revolution is demonstrating its transformative innovation potential. There is no limit when it appears to be conceiving the various ways to use the immense potential of the digital marketing age, inadequate prognosis, diagnosis, treatment, and healthcare monitoring. With ongoing and upcoming technologies the upward surge in medical innovations, the limit is very high when it comes to conceiving the various methods to use the immense potential of the online marketing era, ineffective prognosis, diagnosis, treatment, and healthcare monitoring. on passing of any instance of a medical process running, a significant amount of data is dealt with on a daily basis. These data sets may be informational, referencable , or raw enough to be decisive of adding similar sets of meaningful medical knowledge. This data comes from a variety of sources and is used in a variety of ways. They can help with disease prognosis, diagnosis, and treatment. The analysis of the same could speed the completion of research projects in the same field. It could aid

statistical inferences when it comes to projecting various trends that could aid the overall process. Classification methods are widely used in data mining for medical diagnosis and disease prediction. The second largest internal organ is liver in our body, and it plays an crucial role in metabolism as well as many other essential functions, such as red blood cell decomposition[6]. It weighs approximately three pounds. The liver is responsible for many essential for body functioning, including digestion, metabolism, immunity growth, and storage of nutrients. These functions distinguish the liver as a essential organ, without it, body tissues will rapidly perish due to a lack of energy. The levels of enzymes in the blood have traditionally been used to diagnose liver disease. For prediction of liver disease, a combination of Nave Bayes, Logical Regression, Decision tree, Random Forest, SVM classifier algorithms is used in this study.

II. LITERATURE REVIEW

“Fuzzy Logic for Child-Pugh classification of **patients with Cirrhosis of the Liver**”: In the field of medicine,

survival analysis is a widely used technique. The ability to predict a subject's life expectancy is extremely valuable to both doctors and patients. Any medical care paradigm has three preliminary measures that serve as the fundamental basis. The diagnosis stage, classification stage, assessment stage, conclusion stage, and treatment stage are all stages of the treatment process. All of these stages should be precise to the criteria and reliable in their measurement to accurately represent the quantified magnitude and severity of the disease study the background. The Child-Pugh method is the used methodologies for a comprehensive assessment of liver diseases, especially cirrhosis. The life span of different patients suffering from different types of liver problem is understandable based on the detailed analysis of a large number of cases[1]. For example, Fuzzy Logic is a technique that fits the Context like a glove.

“Liver Disease Detection Due to Excessive Alcoholism Using Data Mining Techniques”: Liquor is consumed in excess by a notable number of people all over the world. Liquor use can cause dangerous liver diseases such as cirrhosis, which can lead to death. Faster noticing of liver diseases caused by excessive alcohol consumption will help to save the lives of many people. By detecting a liver ailment early on, it is very likely to be diagnosed and treated in a punctual manner, resulting in complete recovery in some patients. This paper suggests using knowledge mining calculations to identify as well as predict the likelihood of liver disease. For the dataset, we'll decide on a choice tree, and then the principles will be created. Following the determination of the principles, we will plan and analyze the dataset using various knowledge mining calculations in order to differentiate the liver disease. Our planning dataset was generated using data obtained from the UCI storehouse. It consists of 7 distinct attributes with a total of 345 occurrences[2]. Various classes of blood tests are considered in the dataset, all of which are

directly linked to liver diseases that may arise as a result of excessive alcohol consumption, as well as recurrence of alcohol consumption. The prediction could be made based on the type of liver disease that has been identified.

“Prediction of liver cancer using Conditional Probability Bayes theorem”: The most dangerous infection on the planet is malignant development. Malignant development can be found in the lungs, liver, bosom, bones, and other organs. The most dangerous type of cancer is liver cancer, which has a long-term prognosis. Jaundice, weight loss, yellow-colored urine, spewing, pain at the upper stomach region, sweats, fever, and an enlarged liver are all symptoms of a malignant liver development. Essential liver disease in the form of liver cancer that begins in the liver and helps in stopping the spread to another organs of the body. An auxiliary liver malignant growth is a disease that spreading another organs of the body before reaching the liver. The important organ of body is liver[3]. According to WHO reports, about 30 people out of 100,000 have had liver cancer, and it has mostly affected African and Asian countries in the past. It has since become a well-known ailment. Hepatocellular carcinoma is the most well-known in the form of cancer, and it affects men more than women. A majority of liver malignant development occurs as a result of increased alcohol consumption. To predict liver disease, a variety of data calculations and artificial insight ideas are used. Using the Bayes hypothesis and the WEKA tool, the probability of predicting liver malignant growth is calculated.

“Rule Optimization of Boosted C5.0 Classification Using Genetic Algorithm for Liver Disease Prediction,”: One of the most interesting and important topics among scientists working in the field of therapeutics and software engineering is diagnosing disease by considering highlights that have the greatest impact on acknowledgements. The topic discusses a different concept known as Medical Data

Mining (MDM). Without a question, knowledge mining techniques use a variety of methods, such as characterization and grouping, to organize diseases and their symptoms, which is useful for diagnosing. This paper describes a new method for analyzing liver disease that can assist physicians and their patients in identifying the symptoms of the disease and reducing the time it takes to diagnose and treat it. The proposed strategy would combine the principles of the Boosted C5.0 grouping technique with the Genetic Algorithm (GA) to improve the determination time and accuracy. Rather than making laws for a transformative calculation, the inherited calculation is used to improve and diminish the tenets of another calculation[4]. In comparison to other work in the field, we show that our proposed approach has better execution and throughput. In our work, precision has increased from 81 percent to 93 percent.

1.1 Findings

The existing mechanisms for predicting liver disease are subject to varying points of accuracy and effectiveness. The sense of importance, on the other side, is dictated by the current situation. Different diseases necessitate the precision of a different set of criteria and may not necessitate the same set of inferences in multiple cases. The study indicates that a reasonable level of precision would be reached in the immediate future. The aim of this paper, however, is to improve on those lines and develop more accurate standards. The recent cases' lack of consistency was addressed by designating various combinations to be considered while the case study was being considered. Certain problems with the trained dataset and data elements are also reflected in the current models. The following are some of the obvious weaknesses that have been identified in order to account for creativity in this article, which has resulted in the connotation of improvising along these lines. When it comes to the classification process, it is not necessary:

1. That a classifier's continuity with a specific

collection of data should be maintained for the remainder of the training set This is meant to suggest that certain classifiers are unfit for the data set in the background. Any of the machine learning techniques that are currently being used.

2. Do not appear to be feasible for a large volume of data. Due to fact that the technique fit the situations where there are smaller quantities of data, due consideration is given to the procedure. Certain methodologies are incompatible with one another, non-cohesive, while collection of original data and same implementation for prediction of liver disease.

III. ANALYSIS OF FACTORS AFFECTING ACCURACY

When it related to machine learning, the inferential information gathered are outcome of commonly analyzed observations that often result with the synonym of pattern in the group of data. There are various process for characterization, which is confirmed by the study[7], analysis that the learning of a volume of data that relates to the same context. One the same note, when the data is of limited volume it's abruptly considered in machine learning algorithm. It is difficult to come up with accurate data thus resulted in predictive analysis with inferential patterns.

There are issues that challenge the accuracy of the machine learning algorithms that are used for predictive analysis:

1. **Quality of Data:** Any data that is to serve as a template for critical studies, innovation, analysis and research purposes needs to be thoroughly checked on qualitative works. This is because a bit of inaccurate data can vandalize the integrity, expectancy and accuracy of the

process.

2. **Features of data:** It is imperative to any developer data analysts work with the raw data to project the standby in the very high feature space. Also accelerate the learning efficiency of machine learning system.
3. **The Quantity of data that is involved:** In short, the more the data, the accurate the results. With lesser data, the effectiveness of the process delayed.
4. **Parameters that are involved:** The study and analysis of the algorithms and functionality of system as whole should be feasible to understand and handy to execute by both non-technicians and newbies. In modern algorithms, the innovation is developed with involvement of more than a single parameter.
5. **Scope of the issues:** Algorithms paradigm in machine learning demands a huge collection of data for analysis, it is important to assign in-order with selectiveness features, that define the context in any given statements of problem.

IV. INFERENCE FROM LITERATURE REVIEW

Here we are having some logically inference we can be made from the literature review. Since the idea of using machine learning algorithm for prognosis, diagnosis, and study of liver diseases and their predictioning. The literature surveys proves and conclude the use of Naive Bayes and Support Vector Machine algorithms for the prediction of Liver Diseases. There are major two parameters involved in understanding the methodology.

1. Time taken to execute the prediction of disease.

2. The accuracy of the prediction results.

Through various studies and experimentations, it is clear that SVM (Support Vector Machine) classifier is best of all algorithms with extremely high accuracy

rates. In case of time taken to execute the process, Naive Bayes throughputs higher sustainability.

4.1 Objective

The objective of this paper is to be able to predict the occurrence of the liver disease in sample dataset. In order to calculate the predictability with the accurate accuracy using the appropriate machine learning algorithm.

4.2 Present System

Present system shares a same objective but make bigger with different methodologies to arrive at a relatively less accurate prediction. There are different situations of the data they are used in order to parametrically come to a definite conclusion over prediction of liver diseases. Fuzzy is for the classification of patients with liver cirrhosis. In gastroenterology, the child-push score is used to assess the prognosis of chronic liver diseases, mainly cirrhosis. It was originally made to predict the morality during surgery. Now we used to determine the prognosis which tells the required strength of treatment and the necessary of liver transplantation.

We can know the prognosis, treatment requires and requirement of liver transplantation.

The instrument score with the five clinical measures and each scoring between 1 to 3 and 3 indicates a serious condition of organ.

We use another version of child-pugh score where there are give back changes in these diseases and it has high conjugated bilirubin level. The higher limit for a single point is 4mg/dL and for 2 points is 10mg/dL.

Measure	1	2	3
Total bilirubin	<3 4(<2)	34-50 2-3	>5 0
Serum Albumin	>3 5	2.8 -	<2.8 3.5

Prothrombin Time	<4 <1.7	4-6 1.7-2.3	>6.0 >2.3
Ascites	None	Mild	High
Hepatic Encephalopathy	None	Grade 1-2	Grade 3-4

Liver diseases are classified into child-pugh class A to class C:

Points	Class	One year Survival	Two year
5-6	A	100%	85%
7-9	B	80%	60%
10-15	C	45%	35%

4.3 Proposed System

Machine learning is understandable and one of the most extensively used solutions of big data management where a high Input with raw data can be collated effectively to make appropriate inferences and eventually to come up with a usual appropriate inferences are knowingly to come up with collection of context useful collection of integrative information for a prediction.

4.4 Advantages of the proposed system

There are certain difference that have been adopted in current system the following distinct advantages are observed:

1. **The performance of classification of liver based diseases is further improved:** with the deep acknowledgment in medical field, proposed different set of parameters to distinctly determine the kind of liver disease. Its improvement has become a fear-less complicated task.

Time complexity an accuracy can be measured by various machine learning models, so that we can measure different **parameters, owing to the needs of**

the user: Every prediction system is based on the kind of parameters Expected to accept, compare and then finally come to predictive conclusion. Accordingly, there're different algorithms that are used to model the predictive system to suit the context.

2. **Different machine learning having high accuracy of the results:** comparison to other methodologies, the right machine learning algorithm can aptly increase the efficiency of the predictive system.
3. **Risk factors can be predictive early by machine learning models:** The machine learning algorithm predict the risk factors through simple methodologies

4.5 Advantages of Machine Learning Algorithms

Machine Learning is a functionality of a system to learn the extensive usage of examples that conditions that can be incorporated as a self-improvement process without being coded by a programmer. The results thus obtained is then used by the corporation, moving to an actionable decision making. The data plays a vital role in this formulation, its roots in data mining and closely associated with Bayesian predictive modelling.

4.6 Machine learning vs. Traditional Programming

In traditional programming, the programmer required to analyze, study and code all the rules in accordance to the user requirements with advisory capacity. These rules act as the logical foundation for machine. When a system grows, there is a rising need of complexity vows the system and the need to produce more rules. The learning system in the machine learning paradigm are essential for enabling a system to drive the functionality rules to inferentially use, a set of examples patterns to derive those essential rules and build a solid logical foundation in a system.

4.7 Working with machine learning algorithm

The machine learning component is considered as a brain for the system where all the learning aspects are controlled Centrally. The machine learning algorithms enables the systems to learn, similar to how a human brain works. However to make a machine to accurate prediction, data inputs are utilized. The core working of a machine learning system is proceeding with learning and inference. The discovery of pattern plays a major role. Features selection would be the follow-up procedure, where it decided which core values are used.

4.8 Inferences

The functionality of the proposed system has to be tested for the kind of limitations that could put up constraints on the operations of the system. The powerfulness of the system is tested by exploring the limits using data that are not introduced at each level. The new data that is incorporated into the system, is incorporated and transformed into feature vector, take inference of the model and come up with the prediction.

The life cycle of a machine learning programme is simple and can be summarized as follows:

1. Define a Question
2. Gather Data
3. Visualize Data
4. Train the Algorithm
5. Test
6. Gather Feedback
7. Use the Model to Predict

4.9 Supervised learning

A calculation is a method of determining the relationship between provided contributions and a given yield by gathering information and soliciting feedback from others. For instance, an expert may forecast jar sales using information such as showcasing costs and weather conditions. When the yield information is established, you can use

administered realizing. New knowledge would be predicted by the estimate. Controlled learn is divided into two categories:

1. Classification Task:

Consider the situation where you need to predict a client's sexual orientation for a company. You start by pulling details from your client database about their height, weight, job, pay, obtaining a crate, and so on. You know what each of your clients sexual preference is; it must be male or female. The classifier's goal will be to assign a probability of being a male or female (i.e., the name) based on the data (i.e., highlights you have gathered). You may use new knowledge to make an expectation after the model has worked out how to interpret male or female. For example, suppose you just received new data from an unknown client and need to know if the person is male or female. If the classifier predicts male = 70%, it means the equation is 70% certain that this client is a man and 30% certain that he is a woman. At least two groups can be represented by the label. The preceding example has only two classes, but if a classifier needs to predict an entity, it will have several more (e.g., glass, table, shoes, and so forth each article speaks to a class)

2. Regression task:

The task is a relapse when the result is recurrent esteem. For instance, a financial investigator may need to estimate the value of a stock based on a variety of factors such as its value, previous stock exhibits, and macroeconomic history. The method will be set up to estimate the value of stocks with the least amount of error possible.

V. IMPLEMENTATION

5.1 DATA GATHERING:

- 1) Information Gathering is the initial step of the AI life cycle. The objective of this progression is to

distinguish and acquire all information related issues. In this progression, we need to distinguish the distinctive information sources, as information can be gathered from different sources such as files, database, internet, or mobile gadgets. It is perhaps the main strides of the existence cycle. The amount and nature of the gathered information will decide the proficiency of the yield. The more will be the information, the more exact will be the expectation.

Information required for predicting liver disease.

1. Age
2. Gender
3. Total Bilirubin (mg/dl)
4. Direct Bilirubin (mg/dl)
5. Total Protein (g/dl)
6. Albumin (g/dl)
7. Albumin and Globulin Ratio (%)
8. Sgpt Alamine:Aminotransferase
9. Sgot Aspartate: Aminotransferase
10. Alkphos: Alkaline Phosphotase

Data set used in the project:

age	gender	Total Bilirubin	Direct Bilirubin	Alkphos Alkaline Phosphotase	Sgpt Alamine Aminotransferase	Sgot Aspartate Aminotransferase	Total Protiens	Album in	A/G	R
65	Female	0.7	0.1	187	16	18	6.8	3.3	0.9	1
62	Male	10.9	5.5	699	64	100	7.5	3.2	0.74	1
62	Male	7.3	4.1	490	60	68	7	3.3	0.89	1
58	Male	1	0.4	182	14	20	6.8	3.4	1	1
72	Male	3.9	2	195	27	59	7.3	2.4	0.4	1

5.2 DATA PREPARATION:

Subsequent to gathering the information, we need to set it up for additional means. Information readiness is a stage where we put our information into an appropriate place and set it up to use in our AI preparing. In this progression, first, we set up all information, and afterward randomize the requesting of information. This progression can be additionally partitioned into two cycles:

1. Data exploration:

It is utilized to comprehend the idea of information that we need to work with. We need to comprehend the attributes, organization, and nature of data. A better comprehension of information prompts a powerful result. In this, we discover Correlations, general patterns, and anomalies.

2. Data Pre-Processing:

Presently the subsequent stage is pre-handling of information for its investigation.

5.3 DATA WRANGLING:

It is utilized to comprehend the idea of information that we need to work with. We need to comprehend the attributes, organization, and nature of data. A better comprehension of information prompts a powerful result. In this, we discover Correlations, general patterns, and anomalies.

1. Missing Values
2. Duplicate data
3. Invalid data
4. Noise

So, we use various filtering techniques to clean the data. It is mandatory to detect and remove the above issues because it can negatively affect the quality of the outcome.

5.4 DATA ANALYSIS:

Presently the cleaned and arranged information is given to the examination step. This progression includes: Selection of analytical techniques:

1. Building models
2. Review the result

The point of this progression is to assemble an AI model to break down the information utilizing different insightful methods and survey the result. It begins with the assurance of the sort of the issues, where we select the AI strategies, for example,

Classification i.e., calculated relapse, choice tree, arbitrary timberland and so on, at that point fabricate the model utilizing arranged information, and assess the model. Consequently, in this progression, we take the information and use AI calculations to construct the model.

5.5 TRAIN MODEL:

Now the next step in the testing is to train the model, in this step we train our model to improve its performance for better outcome of the problem. We use datasets to train the model using various machine learning algorithms. Training a model is required so that it can understand the various patterns, rules, and, features.

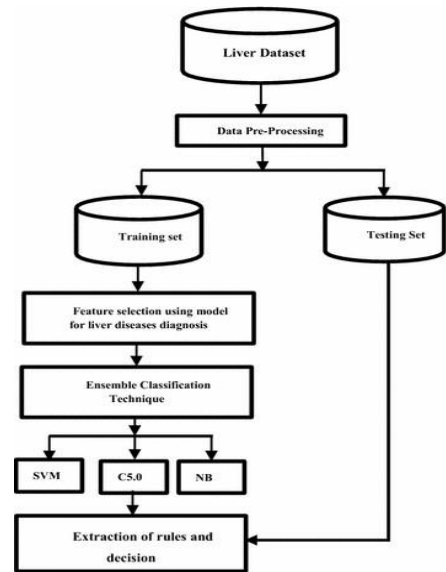
5.6 TEST MODEL:

Presently the subsequent stage is to prepare the model, in this progression we train our model to improve its exhibition for better result of the issue. We use datasets to prepare the model utilizing different AI calculations. Preparing a model is required so it can comprehend the different examples, rules, and, highlights.

5.7 DEPLOYMENT:

The last step of machine learning life cycle is deployment, where we deploy the model in the real-world system. If the above-prepared model is producing an accurate result as per our requirement with acceptable speed, then we deploy the model in the real system. But before deploying the project, we will check whether it is improving its performance using available data or not. The deployment phase is similar to making the final report for a project.

5.8 FLOW CHART:



Prediction of Liver Diseases Based on Machine Learning Technique

5.8 Implementation of classification Techniques:

1. LOGICAL REGRESSION
2. DECISION TREE
3. RANDOM FOREST
4. SUPPORT VECTOR MACHINE

VI. RESULTS

6.1 Logistic Regression:

Logistic regression is quite possibly the most well-known Machine Learning calculations, which goes under the Supervised Learning method. It is utilized for anticipating the absolute ward variable utilizing a given arrangement of autonomous factors. Strategic relapse predicts the yield of a downright reliant variable. Hence the result should be a straight out or discrete worth. It very well may be either Yes or No, 0 or 1, valid or False, and so on however as opposed to giving the specific worth as 0 and 1, it gives the probabilistic qualities which lie somewhere in the range of 0 and 1.

#Predicting the test set result [11]

Predicted_y= classifier.predict(x_test)

In the above code, we have created a predicted_y vector to predict the test set result. By executing the above code, a new vector (prediction_y) will be created under the variable explorer option. It can be seen as: [1 1 1 ... 1 1 1]

precision	recall	f1-score	support	
1	0.95	0.73	0.83	6262
2	0.15	0.56	0.24	528
accuracy			0.72	6790
macro avg	0.55	0.65	0.53	6790
weighted avg	0.89	0.72	0.78	6790

accuracy using LOGICAL REGRESSION

6.2 DECISION TREE:

Decision Tree is a Supervised learning technique that can be utilized for both order and Regression issues, however generally it is liked for taking care of Classification issues. It is a tree-organized classifier, where internal hubs address the highlights of a dataset, branches address the choice rules and each leaf hub addresses the result. In a Decision tree, there are two hubs, which are the Decision Node and Leaf Node. Decision hubs are utilized to used on any choice and have different branches, while Leaf hubs are the yield of those choices and don't contain any further branches. In request to construct a tree, we use the CART algorithm, which stands for Classification and Regression Tree calculation. A choice tree just poses an inquiry, and dependent on the appropriate response (Yes/No), it further split the tree into subtrees.

precision	recall	f1-score	support	
1	1.00	1.00	1.00	4827
2	1.00	1.00	1.00	1963
accuracy			1.00	6790
macro avg	1.00	1.00	1.00	6790
weighted avg	1.00	1.00	1.00	6790

accuracy using DECISION TREE

6.3 RANDOM FOREST:

Random Forest is a mainstream AI calculation that has a place with the regulated learning strategy. It tends to be utilized for both Classification and Regression issues in ML. It depends on the idea of ensemble learning, which is a cycle of combining different classifiers to tackle a mind boggling issue and to improve the exhibition of the model. As the name suggests, "Random Forest is a classifier that contains various choice trees on different subsets of the given dataset and takes the normal to improve the prescient exactness of that dataset." Instead of depending on one choice tree, the irregular woods takes the expectation from each tree and dependent on the dominant part votes of forecasts, and it predicts the last yield. The more prominent number of trees in the woods prompts higher precision and forestalls the issue of overfitting.

Predicting the test set result [11]

```
y_pred_r=clf_r.predict(X_test)ge : 27
```

In the above code, we have created a y_pred_r vector to predict the test set result.By executing the above code, a new vector (y_pred_r) will be created under the variable explorer option. It can be seen as:[1 1 1

... 1 1 1]

precision	recall	f1-score	support	
1	1.00	1.00	1.00	4825
2	1.00	1.00	1.00	1965
accuracy			1.00	6790
macro avg	1.00	1.00	1.00	6790
weighted avg	1.00	1.00	1.00	6790

accuracy using RANDOM FOREST

6.4 SUPPORT VECTOR MACHINE:

Random Forest is a mainstream AI calculation that has a place with the regulated learning strategy. It tends to be utilized for both Classification and Regression issues in ML. It depends on the idea of ensemble learning, which is a cycle of combining different classifiers to tackle a mind blowing issue and to improve the exhibition of the model. λ As the name suggests, "Random Forest is a classifier that contains various choice trees on different subsets of the given dataset and takes the normal to improve the prescient exactness of that dataset." Instead of depending on one choice tree, the irregular woods takes the expectation from each tree and dependent on the dominant part votes of forecasts, and it predicts the last yield. λ The more prominent number of trees in the woods prompts higher precision and forestalls the issue of overfitting.

```
#Predicting the test set result
[11]
y_pred_s=clf_s.predict(X_test)
```

In the above code, we have created a y_pred_s vector to predict the test set result. By executing the above code, a new vector (y_pred_s) will be created under the variable explorer option. It can be seen as : [1 1 1 ... 1 1 1]

```
precision recall f1-score support
1 1.00 1.00 1.00 4827
2 1.00 1.00 1.00 1963
accuracy 1.00 6790 macro avg 1.00 1.00 1.00 6790
weighted avg 1.00 1.00 1.00 6790
accuracy using SVM
```

3) Calculating the accuracy of algorithms:

Below is the accuracy is the percent of correct classifications and can be defined as:

$$\text{Accuracy} = (TP + TN / (TP + FP + TN + FN)) \times 100$$

Sensitivity is the rate of true positive and can be defined as:

$$\text{Precision} = (TP / (TP + FN)) \times 100$$

Specificity is the true negative rate and can be defined as:

$$\text{Recall} = (TN / (TN + FP)) \times 100$$

Where:

TP = the number of positive examples correctly classified.

FP = the number of positive examples misclassified as negative

FN = the number of negative examples misclassified as positive

TN = the number of negative examples correctly classified

VII. CONCLUSION

Information mining is a method for extracting an example from a large amount of data in the areas of AI, data bases, and insights. An knowledge mining technique for therapeutic discovery that includes grouping, ordering, and association. Significant order calculations, such as SVM, NB, and others, are used in the evaluation of liver problem infections forecast. There are 500 informational indexes of 10 characteristics in liver problem infections. Total Bilirubin, Direct Bilirubin, Total Proteins, Albumin, A/G proportion, SGPT (Alanine Aminotransferase), SGOT (Aspartate Aminotransferase), and Alkaline Phosphatase are the qualities that are measured. In the future, we will use a hybrid approach to increase execution precision for liver disease diseases forecasted with their reasonable informational sets.

VIII. REFERENCES

Cite this article as :

- [1]. Sebastian, Anu, and Surekha Mariam Varghese."Fuzzy logic for child-pugh classification of patients with cirrhosis of the liver." 2016 International Conference on Information Science (ICIS). IEEE, 2016
- [2]. Arshad, Insha, et al. "Liver disease detection due to excessive alcoholism using data mining techniques." 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE). IEEE, 2018.
- [3]Ramkumar, N., et al. "Prediction of live cancer using Conditional probability Bayes theorem." 2017 International Conference on Computer Communication and Informatics (ICCCI). IEEE, 2017.
- [4]. Hassoon, Mafazalyaqeen, et al. "Rule optimization of boosted c5. 0 classification using a genetic algorithm for liver disease prediction." 2017 International Conference on Computer and Applications (ICCA). IEEE, 2017.
- [5]. Karthik. S, Priyadarshini. A. Anuradha J. and Tripathi B. K, Classification and Rule Extraction using Rough Set for Diagnosis of Liver Disease and its Types, Ad.
- [6]. Sullivan, Tim. "Blitz scaling" Harvard business review 94.4 (2016): 15.
[7]https://www.academia.edu/39051667/Liver_disease_prediction_using_machine_learning
- [7]. [8]Jae-Young Lim, "The Prospect of the Fourth Industrial Revolution and Home Healthcare in SuperAgedSociety",https://www.researchgate.net/profile/Jae-Young_Lim2
- [9]<https://www.javatpoint.com/machine-learning>
- [8]. Hands on machine learning with scikit learn and tensor flow Introduction to machine learning with python
- Boggarapu Sai Surya, Nitesh Kumar Singh, S Sasi Rekha, "Implementation of Liver Disease Prediction Using Machine Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 2, pp. 606-616, March-April 2021. Available at doi : <https://doi.org/10.32628/CSEIT2172111> Journal URL : <https://ijsrcseit.com/CSEIT2172111>