Evaluation based Approaches for Liver Disease Prediction using Machine Learning Algorithms

**Alternate Title:** Chronic Kidney Disease Prediction using Data Mining and Machine Learning

**Aim**:

 To apply machine learning techniques result in improving the accuracy in the liver disease prediction.

**Abstract:**

 The life of humans living without liver tumors is one of the fundamental care of human livelihood. Therefore, for better care, detection of liver disease at a primitive phase is necessary. For medical experts, predicting the illness in the early stages due to subtle signs is a very difficult task. Many, when it is too late, the signs become evident. The current work aims to augment the perceive nature of liver disease by means of machine learning methods to solve this epidemic. The key purpose of the present work focused on algorithms for classification of healthy people from liver datasets. Centered on their success variables, this research also aims to compare the classification algorithms and to provide prediction accuracy results.

**Introduction:**

 The scale of patient medical records increases day by day in the health care sector. Data mining is the method of using a computer-based information system (CBIS), using modern tactics, to uncover insights from data. The machine learning method is close to that of data mining. Algorithms in machine learning can be differentiated from either supervised or unsupervised methods of learning. For statistical modeling, supervised learning approaches are commonly used. Predictive modeling is a subset of the area of clinical and business intelligence that is used to identify health risks and also to forecast individuals' potential health status. In order to store large-scale information on patient outcomes, procedures, etc., electronic health records (EHR) are used. The data on the HER can be organized or unstructured. Electronic health records are stored in a standardized data format using managed language to log patient knowledge as written texts that is hyperlinked in existence. The EHR aims to streamline knowledge about the clinical workflow. Ensemble learning is a well-known method used for prediction by integrating multiple ensemble models of machine learning. Aggregations of various classifiers are. Naive Bayes, etc. Ensembles search for better outcomes than all of the simple classifiers. The proposed work aims to enhance the predictive and classification quality of healthcare data by developing a hybrid predictive classifier model using the classifier ensemble.

**Synopsis:**

Evaluation-based Approaches for Liver Disease Prediction using Machine Learning Algorithms improving the accuracy in the liver disease prediction.

**Existing System:**

 Previously, they are using LR and SVM algorithms and it was found a low accuracy level. This is chances to predict a wrong output. So, we trained the dataset and use the different types of machine learning algorithm.

**Problem Definition:**

The main problem addressed in this research is the prediction of liver disease. Traditional diagnostic methods may have limitations in terms of accuracy and efficiency. This research aims to leverage machine learning algorithms to improve the accuracy of liver disease prediction and provide a reliable tool for early detection.

**Proposed System**:

 Our Aim is to predict the Liver disease using the machine learning algorithm. The system is automation for predicting the output. We proposed KNN, DT and Random Forest machine learning technique for liver Disease prediction of significant features. ML process starts from a pre-processing data phase followed by feature selection based on data cleaning, classification of modeling, performance evaluation, and the results with improved accuracy.

**Advantage:**

Machine learning algorithms have the potential to identify patterns and relationships in complex datasets, enabling early detection of liver disease before symptoms manifest. By leveraging a wide range of features and applying advanced machine learning algorithms, the accuracy of liver disease prediction can be improved compared to traditional diagnostic methods.

**Module Description:**

* Data Pre-Processing
* Algorithm Implementation
* Prediction

**Data Pre-Processing:**

Our liver Disease project dataset are collected from kaggle.com. Liver Disease data is pre-processed after collection of various records. The dataset contains a more number of patient records, where some records are with some missing values. Those missing records have been removed from the dataset and the remaining patient records are used in pre-processing. After that we remove some columns based on feature selection.

**Algorithm Implementation:**

 The Classification Algorithms to produce the best results. We are using KNN, DT and Random Forest Algorithm to predict the liver disease using ML. On an analysis conducted within various algorithms, the DT was found to provide highest efficiency. Then, the classifiers are applied to each clustered dataset in order to estimate its performance. The best performing models are identified from the above results based on their low rate of error.

* Decision Tree
* Random Forest Classifier
* K- Nearest Neighbor

**Prediction:**

Several standard performance metrics such as accuracy, precision and error in classification have been considered for the computation of performance efficacy of this model. Preprocessed data are trained and input given by the user goes to the trained dataset.

**Software Requirements:**

* Operating System : Windows 10 (64 bit)
* Software : Python 3.7
* Tools : Anaconda (Jupyter Note Book IDE)

**Hardware Requirements:**

* Hard Disk : 500GB and Above
* RAM : 4GB and Above
* Processor : I3 and Above

**CONCLUSION:**
 In this article, using machine learning techniques, the methods for diagnosing liver disease in patients has been proposed and evaluated . SVM, Logistic Regression, comprises two main machine learning techniques used. Using all the models, the prediction analysis has been implemented and their performance has been assessed. The probability of liver disease prediction attained with an high accuracy

**Future work:**

 In future the present scenario can be compared with other techniques such as naïve bayes classification, Random forest etc. Also this work can be further focused on implementation of parametric classifications by bio-inspired optimization algorithm

**Architecture Diagram:**

Dataset

Preprocessing

User Input

Trained model

Machine Learning

No Liver Disease

Liver Disease