**Analysis of customer Loan Approval prediction using Machine Learning Algorithms**

**Alternate Title:** CustomerLoan Approval Prediction Based On Machine Learning techniques.

**Aim**:

To determine the customer loan approval system using machine learning algorithms.

**Abstract:**

Loan approval is a very important process for banking organizations. The systems approved or reject the loan applications. Recovery of loans is a major contributing parameter in the financial statements of a bank. It is very difficult to predict the possibility of payment of loan by the customer. In recent years many researchers worked on loan approval prediction systems. Machine Learning (ML) techniques are very useful in predicting outcomes for large amount of data. In this paper different machine learning algorithms are applied to predict the loan approval of customers. In this paper, various machine learning algorithms that have been used in past are discussed and their accuracy is evaluated. The main focus of this paper is to determine whether the loan given to a particular person or an organization shall be approved or not.

**Synopsis:**

Customer loan approval process for a banking organization its main reason to save a time for customer and bank its output is approval or not approval.

**Existing System:**

The enhancement in the banking sector lots of people are applying for bank loans but the bank has its limited assets which it has to grant to limited people only, so finding out to whom the loan can be granted which will be a safer option for the bank is a typical process. In existing process, they are use LR algorithm in loan approval system. But the efficiency and accuracy was pretty low. Already banks are provide online transaction system, online bank account opening system, etc, But there is no loan approval system in the banking sector. Then now we create a new system for loan approval. So now we move on to the proposed system.

**Problem Definition:**

The data set is imbalance so we used smotetomke method for oversampling and then the dataset is balance its give correct prediction.

**Proposed System:**

The proposed model focuses on predicting the credibility of customers for loan repayment by analyzing their details. The input to the model is the customer details collected. On the output from the classifier, decision on whether to approve or reject the customer request can be made. Using different data analytics tools loan prediction and there severity can be forecasted. In this process it is required to train the data using different algorithms and then compare user data with trained data to predict the nature of loan. The training data set is now supplied to machine learning model; on the basis of this data set the model is trained. Every new applicant details filled at the time of application form acts as a test data set. After the operation of testing, model predict whether the new applicant is a fit case for approval of the loan or not based upon the inference it conclude on the basis of the training data sets. By providing real time input. That data passes to the algorithm. In our project, Decision tree gives high accuracy level compared with other algorithms. Finally, we are predicting the result.

**Advantage:**

The algorithms proposed for this work is Random Forest, Logistic Regression, Support Vector Machine (SVM), (K- NN) K- Nearest Neighbor those algorithms give best testing accuracy.

**Module Description:**

* **Dataset collection**
* **Machine Learning Algorithm**
* **Prediction**

**Dataset collection:**

Dataset is collected from the kaggle.com. That dataset have some value like gender, marital status, self-employed or not, monthly income, etc,. Dataset has the information, whether the previous loan is approved or not depends up on the customer information. That data well be preprocessed and proceed to the next step.

**Machine learning Algorithm:**

In this stage, the collected data will be given to the machine algorithm for training process. We use multiple algorithms to get high accuracy range of prediction. A preprocessed dataset are processed in different machine learning algorithms. Each algorithm gives some accuracy level. Each one is undergoes for the comparison.

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

**Prediction:**

Preprocessed data are trained and input given by the user goes to the trained dataset. The Decision Tree trained model is used to predict and determine whether the loan given to a particular person shall be approved or not.

**Software Requirement:**

* Operating System : Windows 10
* Software : python
* Tools :Anaconda (Jupyter Note Book IDE)

**Hardware Requirement:**

* Hard disk : 500 GB and above.
* Processor : i3 and above.
* Ram : 4GB and above.

**Technologies used:**

* Programming Language: Python.

**CONCLUSION**:   
This study proposes the use of machine learning algorithms to forecast loan acceptance. Customers' loan approval status is predicted using three machine learning algorithms. For loan prediction, the Random Forest algorithm has the uppermost accuracy. In the future, a thorough investigation of alternative machine learning methods for loan endorsement forecast

**Future work:**

In future I got real time dataset I will show high accuracy.

**Architecture Diagram:**

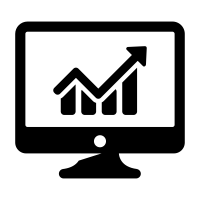
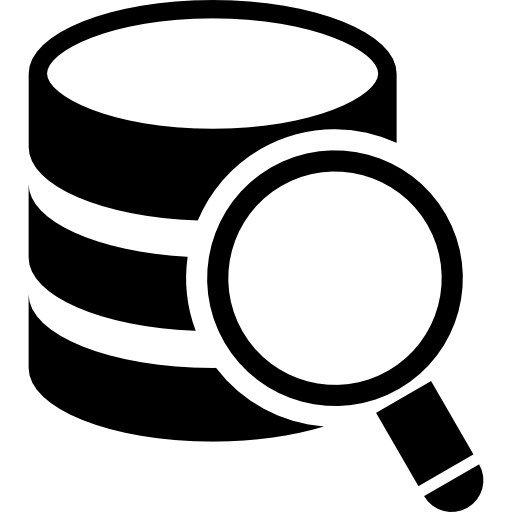
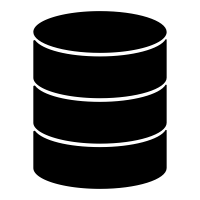
Data Collection

Machine Learning

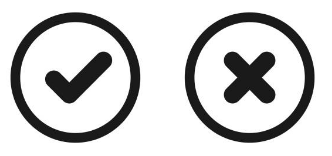
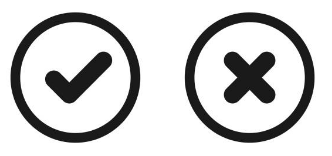
Algorithms

Analysis of Data

Prediction



Data Visualization



Approved

Rejected