|  |  |
| --- | --- |
| Project Id:  **14341**  Project Title:  **Loan Risk Prediction Using Transaction Information** | |
| Name of the student: **Chitaranjan Pradhan** | Roll no:- **201540789** |
| Name of the student:  **Bikram Pradhan** | Roll no:- **201510088** |
| Name of the advisor: **Shom Prasad Das** | |
| **Short description of the project:**  **Objective of the project is to study the ability of neural network algorithms to handle the problem of predicting credit default that measures the creditworthiness of the loan application over a time period** | |
| **Aim/Objective of the project:**  **The objective of proposed work is to predict loan credit risk and determine the probability of non-payment of bank financial services example whether a person will pay back a loan or not. The other objective of the project is to study the ability of neural network algorithms to handle the problem of predicting credit default that measures the creditworthiness of the loan application over a time period. However, there are many risks related to bank loans, for the bank and for those who get the loans. Risk prediction and monitoring is critical for the success of the business model. Credit risk is the probability that a customer won’t be able to make a required payment, causing a loss for the bank or financial institute that provided the loan.** | |
| **Methodology (the necessary software/tool or process steps):**  **Data has been collected from kaggle.com (lending club loan data) that consists of more than 8.5 million records. A random sample data of 60,000 records have been pulled out from the dataset and appropriate attribute selection has been done from 80 attributes. Attribute selection includes numeric and integer attributes along with some factor attribute relevant to the problem this paper is dealing with. Dataset consists of combination of variables as follows:**   1. **Dependent Variable: loan\_status(0 and 1);if the borrower will default then the investment will be bad and if the borrower will not default then he or she will be able to repay the full loan amount. So, to differentiate in neural network 0 indicates borrower will default and 1 indicates borrower will not default.** 2. **Independent Variable: Following variables are considered as an independent variable, loan\_amnt, funded\_amnt, emp\_length, Grade, funded\_amnt\_inv, term, int\_rate, instalment, annual\_inc, issue\_d and application\_type.** 3. **linear regression have been applied to the dataset**   **TOOLS:- 1.RapidMiner**  **2.R** | |
| **Possible outcome of the project:**  **artificial neural network and linear regression models to predict credit default. Both the system has been trained on the loan lending data provided by kaggle.com. Results should be equal effect on the data set and thus are very effective .** | |
| **Importance of the project in engineering aspect:**  **Neural network algorithms have a wide range of applications that are not only essential for residential mortgages. Other applications can be rating bonds issued by companies commonly known as bond rating, rating short term investments that can last up to 1 year, long-term and short-term ratings of local and foreign currencies, sovereign, or country ratings. The prediction system can be further enhanced to assign a credit rating to an application by using appropriate algorithms and technologies** | |
|  | |
| **Innovativeness if any (why the project is different from others):**  **To predict the credit default, several methods have been created and proposed. The use of method depends on the complexity of banks and financial institutions, size and type of the loan. The commonly used method has been discrimination analysis. This method uses a score function that helps in decision making whereas some researchers have stated doubts on the validity of discriminates analysis because of its restrictive assumptions; normality and independence among variables . Artificial neural network models have created to overcome the shortcomings of other inefficient credit default models.**  **The objective of this paper is to study the ability of neural network algorithms to tackle the problem of predicting credit default, that measures the creditworthiness of the loan application over a time period. Feed forward neural network algorithm is applied to a small dataset of residential mortgages applications of a bank to predict the credit default. The output of the model will generate a binary value that can be used as a classifier that will help banks to identify whether the borrower will default or not default. This paper will follow an empirical approach which will discuss two neural network-based models and experimental results will be reported by training and validating the models on residential mortgage loan applications. As the final step in the direction, linear regression method is also performed on the dataset.** | |
| **Reference :**  **[1] Chen, N., Ribeiro, B., & Chen, A. (2016). Financial credit risk assessment: a recent review. Artificial Intelligence Review, 45(1), 1-23.**  <https://www.datasciencecentral.com/profiles/blogs/a-guide-for-applying-machine-learning-techniques-in-finance>  <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2967184> | |