**LOAN ELIGIBILITY PREDICTION**

**Submitted for**

**DATA VISULIZATION AND DASHBOARD**

Submitted by:

**(E23MCAG0024) HARSH KUMAR PANDEY**

Submitted to-

**Dr. Nitin Arvind Shelke**

**July-Dec 2023**

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**



**INDEX**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | Sr No | Content | Page No | | 1. | Abstract and Introduction | 1 | | 2. | Software Used and Methodology | 2 | | 3. | Experimental Result and Conclusions | 3 | | 4. | References | 4 | |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **LIST OF FIGURES**   |  |  |  | | --- | --- | --- | | **Figure No.** | **Title** | **Page No.** | | **1.** | **Flowchart Diagram** | 3 | | **2.** | **Countplot of Dependents** | **4** | | **3.** | **Countplot of Self Employed** | **4** | | **4.** | **Countplot of Credit History** | **5** | | **5.** | **Countplot of Loan Amount** | **5** | | **6.** | **Countplot of Gender** | **6** |      1. **ABSTRACT**   The smooth functioning in the growing economies like in India is inevitable. In the increasing number of customers applying for the loans in bank and also in the non-financial baking companies (NBFC), it’s become challenging for the bank and also for NBFC with their limited capital to device a standard resolutions and make a safe procedure to lend the money to the borrowers for the financial help. In the recent time the NBFC has faced downfall in their stock prices. The main focus is to determined whether is safe or not to allocate a loan to a particular person. The aim of the project is to predict that the particular person is eligible for the loan or not. |  |
| 1. **INTRODUCTION AND RELATED WORK**   **When you apply for the loan your eligibility depends on the income and repayment ability. There is some other factors that also determine the loan eligibility. We are going to use the training set to build the model and test set to validate. Both of the files are stored in the comma separated value (csv). The model will check that the person could repay the loan or not by checking its past credit history. The main focus of the paper is to determine that will be safe or not to give loan to the particular person or he/she is eligible for the loan or not. The following sections this paper contains:**   * **Collection of Data** * **Cleaning the Data** * **Evaluation of the Performance** * **Training the Data** * **Testing the Data** * **Model Building**  1. **SOFTWARE USED**    1. **Python**    2. **Pandas**    3. **Matplotlib**    4. **Seaborn**    5. **NumPy**    6. **Sci-kit learn**    7. **Jupyter Notebook** 2. **METHODOLOGY**   **In the dataset the various columns names that is used for the loan eligibility prediction are :- Gender, Marriage, Credit History, Loan Amount, Loan ID, Number of Dependents, Employment, Applicant Income, Co-Applicant Income, Loan Status, Area, Loan Term.**  **Data Cleaning:**  **In the real world the data are not so perfect, there are missing values which are get generated during the during the process of the data entry. Therefore the data is filtered, the missing values get filled, and the impropriate data are to be removed. The preprocessing of the data helps in the rendering the highly accurate values and the power of the prediction gets increased. The missing values are handled by:**   * **Calculating the total number of missing values present in the dataset.** * **Filling the categorical values with word.** * **Filling the numerical data with either mean or mode.** * **Checking the changes are applied.**   **Fig 1.**    Fig 2.  Fig 3.    Fig 4.  Fig 5.    Fig 6.  **EXPERIMENTAL RESULTS**  **After training all the models by using the train data and testing the test data, the results has come across by looking the values: RMSE value, Accuracy. WE conclude it by saying that the GaussianNB (Naïve Bayes) and Support Vector Machine (SVM) both have predict the eligibility of 82.9%.**  **GaussianNB (Naïve Bayes) = Support Vector Machine (SVM)**  **CONCLUSION**  **It’s observed that from the result GaussianNB and Support Vector Machine (SVM) predict the eligibility range highly and other such as Decision Tree and Random Forest predict in decreasing order in comparatively. The dataset is in the categorical in the nature. Hence, by using of the algorithms we are able to identify that the particular person is eligible for the loan or not. The predictions are of the great use for the formalities and it saves time and it produces the accuracy in result.**  **REFERENCES**  [1] Gupta, Anshika, et al. "Bank Loan Prediction System using Machine Learning." 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART). IEEE,2020.  [2] Supriya, P. Usha et al. “Loan Prediction by using Machine Learning Models.” 2019.  [3] Kumar Arun, Garg Ishan, Kaur Sanmeet, May-Jun. 2016. Loan Approval Prediction based on Machine Learning Approach, IOSR Journal of Computer Engineering (IOSR-JCE)  [4] K. Hanumantha Rao, G. Srinivas, A. Damodhar, M. Vikas Krishna: Implementation of Anomaly Detection Technique Using Machine Learning Algorithms: Internatinal Journal of Computer Science and Telecommunications |  |
|  |  |
|  |  |