|  |  |
| --- | --- |
| **Project Title** | **Bank Risk Controller Systems** |
| **Skills take away From This Project** | **Python, Analytics, statistics, Ploting, streamlit , machine learning, Deep Learning, GenAI** |
| **Domain** | **Banking** |

**Problem Statement:**

**Please refer the given data and do the prediction of whether the customer will get default or not. The column which you are going to predict in dataset is “TARGET”.**

**Business Use Cases:**

Potential Business Scenarios

**Risk Management:** Banks and financial institutions can use the model to assess the risk of potential borrowers defaulting on loans, helping in decision-making for loan approvals.

**Customer Segmentation:** Segment customers based on their risk profile to tailor financial products and services.

**Credit Scoring:** Enhance traditional credit scoring models with predictive analytics to improve accuracy.

**Fraud Detection:** Identify patterns that may indicate fraudulent loan applications.

**Approach:**

Methodology

**Data Collection:** Gather historical loan data from various sources.

**Data Preprocessing:** Clean and preprocess the data to handle missing values, outliers, and categorical variables.

**Exploratory Data Analysis (EDA):** Perform EDA to understand data distributions and relationships between variables.

**Feature Engineering:** Create new features that could enhance the predictive power of the model.

**Model Selection:** Compare various machine learning models such as Logistic Regression, Decision Trees, Random Forest, and Gradient Boosting, etc. to find the best fit.

**Model Training:** Train the selected model(s) on the training dataset.

**Model Evaluation:** Evaluate the model performance using appropriate metrics.

**Hyperparameter Tuning:** Optimize the model parameters to improve performance.

**Model Deployment:** Deploy the model for real-time prediction and integrate it with business systems.

**Results:**

**Expected Outcomes:**

Create a streamlit application dashboard with multiple sidebar menus and do the following operations in respective sidebars.

|  |  |  |
| --- | --- | --- |
| Sidebar Number | Sidebar Name | Function to do |
| 1 | Data | 1. Show the dataset which you are using for model building.  2. Show the model performance metric dataset. |
| 2 | EDA - Visual | 1. Do all kind of EDA analysis for the given dataset.  2. Plot the EDA plots and Statistics plots and show them in this menu.  3. Plot preferred are Plotly or seaborn, try to plot in plotly for interaction. |
| 3 | Prediction | 1. Create a text or number input items in streamlit for selected featured which you have selected for model building.  2. Once the user has entered the value and hit the button of predict, you model should predict and show whether the customer is defaulted or not. |

A trained machine learning model capable of predicting loan defaulters with high accuracy.

**Project Evaluation metrics:**

Success Criteria

Accuracy, Precision, Recall, F1 Score, ROC-AUC

All should be in high range, minimum 0.87 or 87 %.

**Technical Tags:**

*Machine Learning*

*Data Preprocessing*

*Feature Engineering*

*Model Training*

*Model Evaluation*

*Hyperparameter Tuning*

*Deep Learning*

*GenerativeAI*

*Llama2*

*GPT*

*Deployment*

**Data Set:**

Dataset is available in CSV format.

|  |  |  |
| --- | --- | --- |
| Dataset | Loan\_data.csv | <https://drive.google.com/drive/folders/10O0mhM9FCBKeFpIGuJprgDtzJxpEArdx?usp=sharing> |
| Data Dictionary | Data\_dictionary.csv | <https://drive.google.com/drive/folders/10O0mhM9FCBKeFpIGuJprgDtzJxpEArdx?usp=sharing> |

**Data Set Explanation:**

Content and Context

Historical Data: The dataset contains information about past loan applications, including personal details of applicants and their loan status.

Preprocessing Steps:

Handle missing values by imputation.

Encode categorical variables using techniques like one-hot encoding.

Normalize or standardize numerical variables if required.

Create new features based on domain knowledge to improve model performance.

**Project Deliverables:**

Submission Requirements

Source Code: The complete code used for data preprocessing, model training, and evaluation.

Documentation: A report detailing the methodology, analysis, results, and insights.

Presentation: A slide deck summarizing the project and key findings.

Model Files: The trained model ready for deployment.

README: Instructions on how to run the code and reproduce the results.

**Project Guidelines:**

Best Practices

Coding Standards: Standard code standard for Python code.

Version Control: Use Git for version control and regularly commit changes.

Documentation: Comment your code and provide clear explanations for your logic.

Collaboration: Use collaborative tools like GitHub or GitLab for team projects.

**Timeline:**

|  |  |
| --- | --- |
| Analyse data  EDA  Ploting  Building Model  ML Model Selection | 2 weeks |
| Streamlit | 3 Days |
| Total | 2 weeks 3 Day |

**PROJECT DOUBT CLARIFICATION SESSION ( PROJECT AND CLASS DOUBTS)**

**About Session:** The Project Doubt Clarification Session is a helpful resource for resolving questions and concerns about projects and class topics. It provides support in understanding project requirements, addressing code issues, and clarifying class concepts. The session aims to enhance comprehension and provide guidance to overcome challenges effectively.

**Note: Book the slot at least before 12:00 Pm on the same day**

**Timing: Tuesday, Thursday, Saturday (5:00PM to 7:00PM)**

**Booking link :**[**https://forms.gle/XC553oSbMJ2Gcfug9**](https://forms.gle/XC553oSbMJ2Gcfug9)

**LIVE EVALUATION SESSION (CAPSTONE AND FINAL PROJECT)**

**About Session:** The Live Evaluation Session for Capstone and Final Projects allows participants to showcase their projects and receive real-time feedback for improvement. It assesses project quality and provides an opportunity for discussion and evaluation.

**Note: This form will Open on Saturday and Sunday Only on Every Week**

**Timing: Monday-Saturday (11:30PM to 12:30PM)**

**Booking link :** [**https://forms.gle/1m2Gsro41fLtZurRA**](https://forms.gle/1m2Gsro41fLtZurRA)