

High-Level Design Document

Team Details:

Name : Siddharth Gupta (justcallmesidd@gmail.com)

Project Name: Customer Satisfaction Prediction using Python and Tkinter

Abstract - :

This document outlines the high-level design of a machine learning project that utilises an Artificial Neural Network (ANN) to predict whether a bank customer will exit or not based on various parameters including Credit Score, Geography, Gender, Age, Tenure, Balance, Num Of Products, HasCrCard, IsActiveMember, and Estimated Salary. The project also involves the use of the Tkinter library for the development of a user-friendly interface.

Introduction -:

1. Why this High-Level Document?

This document serves as a blueprint for the development of a machine learning application that aims to predict customer churn in a bank using ANN. It provides a structured overview of the project's goals, scope, and design considerations.

2. Scope

The project's scope includes the development of a predictive model, a graphical user interface (GUI) using Tkinter, and the integration of these components to create a user-friendly application. The model will use historical banking data to predict customer churn.

3. Definition

- **Customer Churn:** The act of a customer leaving or discontinuing their relationship with the bank.
- **ANN (Artificial Neural Network):** A machine learning model inspired by the human brain's neural network structure, used for pattern recognition and prediction.
- **Tkinter:** A Python library for creating graphical user interfaces.

General Description-:

1. Product Perspective

The project involves the creation of a standalone machine learning application that can be used by bank employees to predict customer churn. It will be user-friendly and have the ability to input customer data for prediction.

2. Problem Statement

Customer churn is a significant concern for banks. This project aims to address this issue by providing a predictive model that can identify potential churners based on historical data.

3. Problem Solution

The solution involves the development of an ANN-based model trained on a dataset containing customer information. The user interface will allow bank staff to input customer details, and the model will predict whether the customer is likely to exit the bank.

4. Proposed Methodology

- **Data Collection:** Gather a dataset containing relevant customer information, including features such as Credit Score, Geography, Gender, Age, etc.
- **Data Preprocessing:** Clean and preprocess the dataset, handle missing values, and encode categorical variables.
- **Model Development:** Create an ANN model using suitable libraries (e.g., TensorFlow or PyTorch).
- **User Interface:** Develop a GUI using Tkinter to input customer data and display predictions.
- **Model Training:** Train the ANN model on the preprocessed dataset.
- **Integration:** Connect the trained model with the user interface.
- **Testing:** Evaluate the model's performance using appropriate metrics.
- **Deployment:** Make the application accessible to bank employees for real-time predictions.

5. Further Improvements

Future enhancements may include:

- Continuous model retraining with new data.
- Integration with the bank's CRM system for automatic data retrieval.
- Enhanced visualisations of prediction results.

6. Data Required

The project requires a dataset containing customer attributes, such as Credit Score, Geography, Gender, Age, Tenure, Balance, Num Of Products, HasCrCard, IsActiveMember, and Estimated Salary, along with a target variable indicating whether the customer exited the bank or not.

7. Tools Used

- Programming Language: Python
- Machine Learning Libraries: TensorFlow or PyTorch
- Graphical User Interface: Tkinter
- Data Manipulation: Pandas
- Data Visualization: Matplotlib or Seaborn
- Data Preprocessing: Scikit-learn

8. Constraints

- Availability of a suitable dataset.
- Adequate computing resources for model training.
- Privacy and security considerations for handling customer data.

9. Assumptions

- The historical data used for training is representative of future customer behavior.
- The provided dataset is accurate and complete.
- The bank staff using the application are trained to interpret and act on the model's predictions.

Design Details-:

1. Process Workflow
 1. Data Collection and Preprocessing
 2. Model Training
 3. GUI Development
 4. Model Integration with GUI
 5. Testing and Validation

2. Error Handling

- Implement robust error handling mechanisms in the GUI to handle invalid inputs.
- Log and handle errors gracefully during model prediction.

Performance -:

1. Reusability

The machine learning model can be reused for similar prediction tasks in the future by retraining it with updated data.

2. Application Compatibility

The application will be compatible with Windows, macOS, and Linux operating systems, as Tkinter is a cross-platform library.

3. Resources Utilisation

Efficient use of computing resources will be ensured during model training to optimise speed and cost.

4. Deployment

The application will be deployed on the bank's internal network or cloud platform to make it accessible to authorised users.

Conclusion

This high-level design document provides a comprehensive overview of the machine learning project that aims to predict customer churn in a bank using an ANN-based model and a Tkinter-based GUI. It outlines the project's goals, scope, methodology, tools, and design considerations, laying the foundation for its successful implementation.