SmartLender - Applicant Credibility Prediction for Loan Approval

Milestone 1: Project Initialization and Planning Phase

In the Project Initialization and Planning Phase for our rainfall prediction using ML, we define objectives, gather requirements, create a roadmap, assemble the team, and identify resources and risks for successful execution.

Activity 1: Define Problem Statement

Our project aims to develop a machine learning model that accurately predicts rainfall patterns using historical weather data, enhancing forecasting accuracy for agriculture, disaster management, and resource planning.

SmartLender Problem Statement Report:

Activity 2: Project Proposal (Proposed Solution)

To address rainfall prediction, we propose developing a machine learning model leveraging historical weather data, including temperature, humidity, and precipitation patterns. Our solution involves data preprocessing, model selection, training, and evaluation, aiming to enhance forecast accuracy for better resource management and planning.

SmartLender Project Proposal Report:

Activity 3: Initial Project Planning

In initial planning, we set project objectives, gather historical weather data, assemble a skilled team, allocate resources, create a timeline, assess risks, and budget effectively for successful execution.

SmartLender Project Planning Report:

Milestone 2: Data Collection and Preprocessing Phase

In the Data Collection and Preprocessing Phase, we gather historical weather data, clean and handle missing values, normalize features, and split data into training and test sets for accurate model development.

Activity 1: Data Collection Plan, Raw Data Sources Identified, Data Quality Report

data Collection Plan: Gather historical weather data from NOAA and local stations.

Raw Data Sources: NOAA, NASA, and weather stations.

Data Quality Report: Ensure completeness and accuracy, addressing missing values and anomalies.

SmartLender Data Collection Report

Activity 2: Data Quality Report

Data Quality Report: The dataset is complete with minimal missing values. Accuracy is verified against multiple sources. Anomalies are addressed, ensuring consistency and reliability for effective machine learning model training.

SmartLender Data Quality Report

Activity 3: Data Exploration and Preprocessing

Analyze data distributions, identify patterns, and visualize relationships. Clean data by handling missing values, outliers, and normalizing features to ensure quality inputs for model training.

Activity 1: Feature Selection Report

Key features for rainfall prediction include temperature, humidity, and historical precipitation data. Irrelevant or redundant features are excluded, enhancing model performance and reducing overfitting risks.

SmartLender Feature Selection Report

Activity 2: Model Selection Report

Evaluated models like Linear Regression, Random Forest, and XGBoost. Chose Random Forest for its accuracy and robustness in handling complex relationships and variable interactions in rainfall prediction.

SmartLender Model Selection Report

Activity 3: Initial Model Training Code, Model Validation and Evaluation Report

Achieved high accuracy and low error rates; validated with cross-validation, and metrics like RMSE and R^2 confirmed robust performance for rainfall prediction.

SmartLender Model Development Phase Template:

Milestone 4: Model Optimization and Tuning Phase

Refined model parameters using grid search and cross-validation. Improved performance by adjusting hyperparameters, optimizing feature selection, and reducing overfitting to enhance prediction accuracy and generalization.

Activity 1: Hyperparameter Tuning Documentation

Applied grid search to optimize parameters such as tree depth and number of estimators. Evaluated performance using cross-validation, selecting the best combination for improved accuracy and model stability.

Activity 2: Performance Metrics Comparison Report

Compared models using metrics like RMSE, MAE, and R^2. Random Forest showed superior accuracy and lower error rates compared to Linear Regression and XGBoost, validating its effectiveness.

Activity 3: Final Model Selection Justification

Chose Random Forest for its superior accuracy, lower error rates, and robustness in handling complex patterns compared to other models, ensuring reliable and precise rainfall predictions.

SmartLender Model Optimization and Tuning Phase Report

Milestone 5: Project Files Submission and Documentation

For the documentation, Kindly refer to the link

Milestone 6: Project Demonstration

Showcase the model's functionality with real-time rainfall predictions using a user-friendly interface. Highlight accuracy, performance metrics, and practical applications to illustrate its effectiveness and impact on forecasting.