

Determine: Loan from KIVA crowdfunding data

Milestone 1: Project Initialization and Planning Phase

The "Project Initialization and Planning Phase" marks the project's outset, defining goals, scope, and stakeholders. This crucial phase establishes project parameters, identifies key team members, allocates resources, and outlines a realistic timeline. It also involves risk assessment and mitigation planning. Successful initiation sets the foundation for a well-organized and efficiently executed machine learning project, ensuring clarity, alignment, and proactive measures for potential challenges.

Activity 1: Define Problem statement

The problem statement for Kiva crowdfunding data focuses on identifying factors that influence loan success rates, understanding borrower demographics and behaviors, and improving the platform's effectiveness in connecting lenders with borrowers to enhance global financial inclusion and support underserved communities.

Problem Statement Report: [Click Here](#)

Activity 2: Project Proposal (Proposed Solution)

The proposed solution for the Kiva crowdfunding project includes developing predictive models to identify key success factors, creating data visualizations to reveal trends, and implementing a recommendation system to match lenders with borrowers, enhancing loan success and financial inclusion.

Project Proposal Report: [Click Here](#)

Activity 3: Initial Project Planning

Initial project planning for Kiva crowdfunding data involves defining goals, identifying key stakeholders, outlining data sources, setting up a timeline, determining required tools and technologies, establishing data processing methods, and creating a communication plan for effective collaboration and progress tracking.

Project Planning Report: [Click Here](#)

Milestone 2: Data Collection and Preprocessing Phase

The Data Collection and Preprocessing Phase involves executing a plan to gather relevant Business

bankruptcy data from Kaggle, ensuring data quality through verification and addressing missing values. Preprocessing tasks include cleaning, encoding, and organizing the dataset for subsequent exploratory analysis and machine learning model development.

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

The dataset for "Smart bridge- Determine : Loan from Kiva crowdfunding data" is sourced from Kaggle. It includes Business details and financial metrics and business upturns and business downturns. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modeling.

Data Collection Report: [Click Here](#)

Activity 2: Data Quality Report

The dataset for "Smart bridge - Determine : Loan from Kiva crowdfunding data " is sourced from Kaggle. It includes applicant details and financial metrics. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modeling.

Data Quality Report: [Click Here](#)

Activity 3: Data Exploration and Preprocessing

Data exploration and preprocessing for Kiva crowdfunding data involve cleaning and normalizing the dataset, handling missing values, analyzing data distributions, identifying key features, and transforming data for analysis to ensure accurate and reliable insights for model development

Data Exploration and Preprocessing Report: [Click Here](#)

Milestone 3: Model Development Phase

The Model Development Phase entails crafting a predictive model for loan approval. It encompasses strategic feature selection, evaluating and selecting models (Random Forest, Decision Tree, KNN, Random Forest Classifier, Decision Tree Classifier), initiating training with code, and rigorously validating and assessing model performance for informed decision-making in the lending process.

Activity 1: Feature Selection Report

The feature selection report for Kiva crowdfunding data highlights key predictors of loan success, including borrower demographics, loan purpose, lender count, repayment term, and geographical location, based on statistical analysis and feature importance rankings to enhance predictive model accuracy

Feature Selection Report: [Click Here](#)

Activity 2: Model Selection Report

The model selection report for Kiva crowdfunding data evaluates various algorithms, including logistic regression, decision trees, random forests, and gradient boosting, comparing their performance using metrics like accuracy, precision, recall, and F1 score to determine the best model for predicting loan success.

Model Selection Report: [Click Here](#)

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

The initial model training code uses a random forest classifier on Kiva crowdfunding data, splitting data into training and test sets. The model validation and evaluation report assesses accuracy, precision, recall, and F1 score, confirming model reliability and identifying areas for improvement.

Model Development Phase Template: [Click Here](#)

Milestone 4: Model Optimization and Tuning Phase

The model optimization and tuning phase for Kiva crowdfunding data involves fine-tuning hyperparameters using grid search or random search, cross-validation to prevent overfitting, and feature engineering to enhance model performance, aiming to improve predictive accuracy and overall model robustness.

Activity 1: Hyperparameter Tuning Documentation

The hyperparameter tuning documentation for Kiva crowdfunding data loans details the process of optimizing model performance by adjusting parameters such as learning rate, tree depth, and regularization strength. It includes results from grid search or random search and their impact on model accuracy and stability.

Activity 2: Performance Metrics Comparison Report

The performance metrics comparison report for Kiva crowdfunding data loans evaluates models based on accuracy, precision, recall, and F1 score. It highlights the strengths and weaknesses of each model type, providing insights into their effectiveness in predicting loan success and guiding decision-making processes.

Activity 3: Final Model Selection Justification

The final model selection for Kiva crowdfunding data loans is justified based on the random forest classifier's superior performance in accuracy, precision, recall, and F1 score metrics during rigorous evaluation and validation. Its robustness and ability to handle complex relationships make

it ideal for predicting loan outcomes.

Model Optimization and Tuning Phase Report: [Click Here](#)

Milestone 5: Project Files Submission and Documentation

For project file submission in Git hub, Kindly click the link and refer to the flow. [Click Here](#)

For the documentation, Kindly refer to the link. [Click Here](#)

Milestone 6: Project Demonstration

In the upcoming module called Project Demonstration, individuals will be required to record a video by sharing their screens. They will need to explain their project and demonstrate its execution during the presentation.