

SmartLender – Gem Valuation Revolution: Predicting Diamond Prices With Artificial Neural Networks

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

Diamonds are a precious commodity, with their prices influenced by a multitude of factors such as carat, cut, color, clarity, and other attributes. Accurate pricing is crucial for various stakeholders, including jewelers, investors, and consumers. Traditional methods of price determination can be subjective and inconsistent. Hence, there is a need for a reliable and data-driven approach to predict diamond prices accurately.

Ref. template: [Click here](#)

SmartLender Problem Statement Report: [Click Here](#)

Activity 2: Project Proposal (Proposed Solution)

The proposed project aims to revolutionize the diamond pricing process by leveraging advanced machine learning techniques, specifically ANNs. By providing a reliable, consistent, and accurate pricing model, this project has the potential to greatly benefit various stakeholders in the gem industry, ensuring fair and transparent pricing mechanisms.

40

Ref. template: [Click Here](#)

SmartLender Project Proposal Report: [Click Here](#)

Activity 3: Initial Project Planning

Define detailed project requirements and success criteria. Assign project roles (e.g., Data Scientists, Data Engineers, Developers, Domain Experts). Estimate costs for data acquisition, software licenses, cloud services, and salaries.

Ref. template: [Click Here](#)

SmartLender 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 loan application 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

To collect comprehensive and high-quality data on diamonds, including various attributes that influence their pricing. Attributes: Carat, Cut, Color, Clarity, Depth, Table, Price, Dimensions (length, width, height). Uniform formatting and standardized attribute values (e.g., categorical values for Cut, Color, Clarity).

Ref. template: [Click Here](#)

SmartLender Data Collection Report: [Click Here](#)

Activity 2: Data Quality Report

The purpose of this data quality report is to evaluate the quality of the diamond dataset collected for the "Gem Valuation Revolution: Predicting Diamond Prices with ANN" project. This assessment covers various dimensions of data quality, including completeness, consistency, accuracy, timeliness, and validity. The report also outlines data cleaning and preprocessing steps undertaken to improve data quality.

Ref. template: [Click Here](#)

SmartLender Data Quality Report: [Click Here](#)

Activity 3: Data Exploration and Preprocessing

The data exploration and preprocessing steps have provided a deep understanding of the dataset and ensured it is clean, consistent, and ready for modeling. This foundation is crucial for developing an accurate and robust Artificial Neural Network (ANN) to predict diamond prices. The next steps will involve model development, training, and evaluation, building on the prepared dataset.

40

Ref. template: [Click Here](#)

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

Milestone 3: Model Development Phase

The Model Development Phase entails crafting a predicting diamond prices. It encompasses strategic feature selection, evaluating and selecting models (Random Forest, Decision Tree, KNN, XGB), 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 process is crucial for developing an accurate and efficient model for predicting diamond prices. This report outlines the methods and rationale used to select the most relevant features from the dataset. The goal is to identify features that contribute significantly to the prediction of diamond prices while removing redundant or irrelevant features.

Ref. template: [Click Here](#)

SmartLender Feature Selection Report: [Click Here](#)

Activity 2: Model Selection Report

The selected Multilayer Perceptron (MLP) model, optimized through rigorous evaluation and hyperparameter tuning, is expected to provide accurate predictions of diamond prices based on the selected features. Further validation and testing will confirm the model's robustness and applicability for deployment in predicting diamond prices effectively.

3.5

Ref. template: [Click Here](#)

SmartLender Model Selection Report: [Click Here](#)

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

The initial training and evaluation of the MLPRegressor model show promising results in predicting diamond prices based on the selected features. Further optimization and fine-tuning may be explored to improve model performance, such as adjusting hyperparameters, exploring different architectures, or incorporating additional feature engineering techniques.

Ref. template: [Click Here](#)

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

Milestone 4: Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Activity 1: Hyperparameter Tuning Documentation

Hyperparameter tuning is crucial for optimizing the performance of machine learning models, including Artificial Neural Networks (ANNs) such as the Multilayer Perceptron (MLP) used in your project. This documentation outlines the process and considerations for tuning hyperparameters to achieve better accuracy and generalization in predicting diamond prices.

Activity 2: Performance Metrics Comparison Report

Based on the performance metrics evaluated, the Multilayer Perceptron (MLP) Regressor shows promising results for predicting diamond prices, potentially outperforming traditional regression models like Linear Regression and offering competitive accuracy compared to ensemble methods like Random Forest.

Activity 3: Final Model Selection Justification

Based on comprehensive evaluation of model performance metrics and alignment with project goals, the Multilayer Perceptron (MLP) Regressor is selected as the final model for predicting diamond prices. Its ability to handle complex relationships, achieve competitive accuracy, and potential for further optimization make it well-suited for practical deployment in predicting diamond prices based on the selected features.

Ref. template:

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

Milestone 5: Project Files Submission and Documentation

For project file submission in Github, 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.