SmartBridge-Sentiment Analysis of Commodity News (Gold)

Milestone1:ProjectInitializationandPlanningPhase

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 andmitigation planning. Successful initiation sets the foundation for a well-organized and efficientlyexecuted machine learning project, ensuring clarity, alignment, and proactive measures forpotential challenges.

Activity1:DefineProblemStatement

Problem Statement: A self-employed customer wants to know the price sentiment of commodity news like gold from the newspaper headlines.

ProblemStatementReport:Click Here

Activity2:ProjectProposal(ProposedSolution)

The proposed project, "Sentiment Analysis of Commodity News (Gold)" aims to leverage machinelearning for more accurate price sentiment predictions. Using a comprehensive datasetincluding news headlines, price and conditions the project seeks to develop apredictive model optimizing Sentiment Analysis of Commodity News (Gold). This initiative aligns with ourobjectivetoenhancedecision-making,reducerisks,andstreamlinelendingoperations,ultimatelyimprovingcustomersatisfactionandoperationalefficiency.

ProjectProposalReport: Click Here

Activity3:InitialProjectPlanning

Initial Project Planning involves outlining key objectives, defining scope, and identifyingstakeholdersforaestimating and predicting sentiment analysis of gold.Itencompassessettingtimelines,allocatingresources,and determining the overall project strategy. During this phase, the team establishes a clearunderstanding of the dataset, formulates goals for analysis, and plans the workflow for dataprocessing. Effective initial planning lays the foundation for a systematic and well-executedproject,ensuringsuccessfuloutcomes.

ProjectPlanningReport: Click Here

Milestone2:DataCollectionandPreprocessingPhase

TheDataCollectionandPreprocessingPhaseinvolvesexecutingaplantogatherrelevantloan

applicationdatafromKaggle,ensuringdataqualitythroughverificationandaddressingmissingvalues.Preprocessing tasks include cleaning, encoding, and organizing the dataset for subsequent exploratoryanalysisandmachinelearningmodeldevelopment.

Activity1:DataCollectionPlan,RawDataSourcesIdentified,DataQualityReport

The dataset for "Sentiment Analysis of Commodity News (Gold)" is sourcedfrom Kaggle. It includes applicant details and financial metrics. Data quality is ensured throughthoroughverification,addressingmissingvalues,andmaintainingadherencetoethicalguidelines,establishingareliablefoundationforpredictivemodeling.

DataCollectionReport: **Click Here**

Activity2:DataQualityReport

The dataset for "Sentiment Analysis of Commodity News (Gold)" is sourcedfrom Kaggle. It includes applicant details and prediction metrics. Data quality is ensured thoroughverification,addressingmissingvalues,andmaintainingadherencetoethicalguidelines,establishingareliablefoundationforpredictivemodeling.

DataQualityReport: **Click Here**

Activity3:DataExplorationandPreprocessing

Data Exploration involves analyzing the prediction dataset to understand patterns,distributions.Preprocessingincludeshandlingmissingvalues,scaling,andencodingcategorical variables. These crucial steps enhance data quality, ensuring the reliability andeffectivenessofsubsequentanalysesintheloanapprovalproject.

DataExplorationandPreprocessingReport: Click Here

Milestone 3: Model Development Phase

The Model Development Phase entails crafting a predictive model for loan approval. It encompassesstrategicfeatureselection,evaluatingandselectingmodels(RandomForest,DecisionTree,KNN, XGB),initiating training with code, and rigorously validating and assessing model performance for informeddecision-makinginthelendingprocess.

Activity1:FeatureSelectionReport

The Feature Selection Report outlines the rationale behind choosing specific features (e.g., news headline, status etc.,)fortheprice predictionmodel.Itevaluatesrelevance,importance,and impact on predictive accuracy, ensuring the inclusion of key factors influencing the model'sabilitytodiscerncredibleprice prediction.

FeatureSelectionReport: **Click Here**

Activity2:ModelSelectionReport

TheModelSelectionReportdetailstherationalebehindchoosingRandomForest,DecisionTree,KNN, and XGB models for price prediction. It considers each model's strengths inhandlingcomplexrelationships,interpretability,adaptability,andoverallpredictiveperformance,ensuringaninformedchoicealignedwithprojectobjectives.

ModelSelectionReport:Click Here

Activity3:InitialModelTrainingCode,ModelValidationandEvaluationReport

The Initial Model Training Code employs selected algorithms on the sentiment price prediction dataset,setting the foundation for predictive modeling. The subsequent Model Validation and EvaluationReportrigorouslyassessesmodelperformance,employingmetricslikeaccuracyandprecisiontoensurereliabilityandeffectivenessinpredictingloanoutcomes.

ModelDevelopmentPhaseTemplate: Click Here

Milestone4:ModelOptimizationandTuningPhase

The Model Optimization and Tuning Phase involves refining machine learning models for peakperformance.Itincludesoptimizedmodelcode,fine-tuninghyperparameters,comparingperformancemetrics,andjustifyingthefinalmodelselectionforenhancedpredictiveaccuracyandefficiency.

Activity1:HyperparameterTuningDocumentation

The logistic regression model was selected for its superior performance, exhibiting high accuracyduring hyperparameter tuning. Its ability to handle complex relationships, minimize overfitting,andoptimizepredictiveaccuracyalignswithprojectobjectives,justifyingitsselectionasthefinalmodel.

Activity2:PerformanceMetricsComparisonReport

ThePerformanceMetricsComparisonReportcontraststhebaselineandoptimizedmetricsforvarious models, specifically highlighting the enhanced performance of the logistic regressionmodel. This assessment provides a clear understanding of the refined predictive capabilitiesachievedthroughhyperparametertuning.

Activity3:FinalModelSelectionJustification

The Final Model Selection Justification articulates the rationale for choosing logistic regression as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successfulhyperparametertuningalignwithprojectobjectives,ensuringoptimalloanapprovalpredictions.

ModelOptimizationandTuningPhaseReport: Click Here

Milestone5:ProjectFilesSubmissionand Documentation

,Kindlyclickthelinkandrefertotheflow. **click Here**

Forthedocumentation,Kindlyrefertothelink. Click Here

Milestone6:ProjectDemonstration

IntheupcomingmodulecalledProjectDemonstration,individualswillberequiredtorecordavideobysharing their screens. They will need to explain their project and demonstrate its execution during thepresentation.