

Drug classification using machine learning

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

Problem Statement: The pharmaceutical industry struggles with accurate drug classification due to manual, error-prone methods. We propose a machine learning model to classify drugs based on chemical properties, mechanisms of action, and therapeutic uses. This model, using Python and libraries like scikit-learn and TensorFlow, aims to improve accuracy and efficiency.

Enhanced classification will support drug discovery, regulatory compliance, and personalized medicine. Ultimately, this solution will optimize drug development and enhance patient safety.

Ref template: [click here](#)

Drug classification using machine learning Problem Statement Report:

Activity 2: Project Proposal (Proposed Solution)

The project aims to leverage machine learning for accurate drug classification. Using a comprehensive dataset with features like Age, Sex, BP, Cholesterol, Na_to_K and Drug. The project seeks to develop a predictive model to classify drugs into different categories. This model assists in understanding drug efficacy and safety.

Ref. template: [click here](#)

Drug classification using machine learning project proposal Report:

Activity 3: Initial Project Planning

Initial Project Planning: This involves outlining key objectives, defining scope, and identifying stakeholders for the drug classification project. It encompasses setting timelines, allocating resources, and determining the overall project strategy. Effective initial planning lays the foundation for a systematic and well-executed project, ensuring successful outcomes.

Ref. template: [Click here](#)

Drug classification using machine learning Project Planning Report:

Milestone 2: Data Collection and Preprocessing Phase

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

Data Collection Plan: The dataset for "Drug Classification" is sourced from public repositories. It Age, Sex, BP, Cholesterol, Na_to_K and Drug includes data of drugs. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modeling.

Ref. template: [click here](#)

Drug classification using machine learning Data Collection Report:

Activity 2: Data Quality Report

The dataset includes Age, Sex, BP, Cholesterol, Na_to_K and Drug activity data of drugs. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modeling.

Ref. template: [click here](#)

Drug classification using machine learning Data Quality Report:

Activity 3: Data Exploration and Preprocessing

This involves analyzing the drug dataset to understand patterns, distributions, and outliers. Preprocessing includes handling missing values, scaling, and encoding categorical variables. These crucial steps enhance the dataset's quality, ensuring it is ready for model training and testing phases.

Ref. template: [click here](#)

Drug classification using machine learning Data Exploration and Preprocessing Report:

Milestone 3: Model Development Phase

The Model Development Phase entails crafting a predictive model. It encompasses strategic feature selection, evaluating and selecting models (Random Forest, Decision

Tree, KNeighbors, Xgboost), initiating training with code, and rigorously validating and assessing model performance for informed decision-making in the lending process.

Activity 1: Feature Selection Report

This report outlines the rationale behind choosing specific features (e.g., Age, Sex, BP, Cholesterol, Na_to_K and Drug) for the drug classification model. It evaluates relevance, importance, and impact on predictive accuracy, ensuring the inclusion of key factors influencing the model's performance.

Ref. template: [click here](#)

Drug classification using machine learning Feature Selection Report:

Activity 2: Model Selection Report

The Model Selection Report details the rational behind choosing models such as Random Forest, Decision Tree, KNN, and XGBoost for drug classification. It considers each model's strengths in handling complex relationships, interpretability, adaptability, and overall predictive performance, ensuring an informed choice aligned with project objectives.

Ref. template: [click here](#)

Drug classification using machine learning Model Selection Report:

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

The Initial Model Training Code employs selected algorithms on the drug classification dataset, setting the foundation for predictive modeling. The subsequent Model Validation and Evaluation Report rigorously assesses model performance, employing metrics like accuracy and precision to ensure reliability and effectiveness in drug classification.

Ref. template: [click here](#)

Drug classification using machine learning Model Development Phase Report:

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 Documentation: This phase involves refining the machine learning models for peak performance. It includes fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Activity 2: Performance Metrics Comparison Report

The Performance Metrics Comparison Report contrasts the baseline and optimized metrics for various models, specifically highlighting the enhanced performance of the model. This assessment provides a clear understanding of the refined predictive capabilities achieved through hyperparameter tuning.

Activity 3: Final Model Selection Justification

The Final Model Selection Justification articulates the rationale for choosing Random forest as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring optimal drug classification.

Ref. template: [click here](#)

Drug classification using machine learning Model Optimization and Tuning Phase Report:

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.