Predicting CO2 emissions by countries 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

Predictive Machine Learning (ML) models and the big amount of available data can be very useful to analyze the development of climate change trends or relevant contributors. In theory, the country emissions of greenhouse gases such as CO2 over a year could depend on certain country-specific aspects. In this context, I have developed a ML project aiming to analyze and predict CO2 emissions from country-specific parameters such as economic indicators, country,year, value of emission etc.

Ref. Template :click here **Problem Statement Report:** click here

Activity 2: Project Proposal (Proposed Solution)

The proposal report aims to transform sepsis survival prediction using Machine learning, boosting efficiency and accuracy. It tackles system inefficiencies , promising better clinical outcomes ,reduced risks, and happier patients. Key features include a machine learning -based decision-making.

Ref. Template: click here **Project Proposal Report:** click here

Activity 3: Initial Project Planning

The project will begin with a comprehensive literature review to identify key factors influencing CO2 emissions. Data collection will involve gathering historical emissions data and relevant socio-economic indicators from reliable sources. Data preprocessing will follow, including cleaning, normalization, and feature engineering. The next phase involves selecting and training machine learning models, such as Random forest regressor to predict CO2 emissions. Model evaluation will include cross-validation and performance metrics . Finally, the project will conclude with a detailed analysis of the results, visualization of predictions, and a report highlighting insights, limitations, and future work.

Ref. Template: click here **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 1application 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

Dataset variable will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling ,and forming a strong foundation for insights and predictions.

Ref. template: click here **Data Collection Report:** click here

Activity 2: Data Quality Report

The Data Quality Report will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Ref. template: click here **Data Quality Report:** click here

Activity 3: Data Exploration and Preprocessing

Dataset variable will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling ,and forming a strong foundation for insights and predictions.

Ref. template: click here**Data Exploration and Preprocessing Report:** click here

Milestone 3: Model Development Phase

The Model Development Phase entails crafting a predictive model for predicting co2 emission by countries. 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

In the forthcoming update, each feature will be accompanied by a brief description. Users will indicate whether it's selected or not , providing reasoning for their decision. This process will streamline decision-making and enhance transparency in future selection.

Ref. template:click here Feature Selection Report: click here

Activity 2: Model Selection Report

In the forthcoming Model Selection Report for predicting co2 emission by countries , various models will be outlined, detailing their descriptions, hyperparameters, including Accuracy or F1 score. This comprehensive report will provide insights into the chosen models and their effectiveness.

Ref. template: click here Model Selection Report: click here

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

This initial model training code will be showcased in the future through a screenshot. The model validation and evaluation report will include classification reports, accuracy, and confusion matrices for multiple models, presented through respective screenshots.

Ref. template: click here Model Development Phase Template:click here