COVID-19 Data Scientist Project Proposal

Problem Identification

Problem Statement Formation:
The goal is to develop an accurate time series forecasting model for predicting the spread of COVID-19 cases and deaths specifically within the USA.

Context:
The COVID-19 continues to impact the USA, necessitating a focused approach to forecasting within the country. The dataset includes information on confirmed cases and deaths across different states and regions, incorporating temporal and geographical attributes.

Criteria for Success:
Success will be determined by the forecasting model's accuracy in predicting COVID-19 cases and deaths within the USA. The model should offer interpretable insights into temporal trends, regional variations, and factors influencing the spread of the virus.

Scope of Solution Space:
The solution involves developing a time series forecasting model tailored to the USA, addressing temporal dynamics, spatial heterogeneity, and relevant influencing factors specific to the country. The solution space includes data preprocessing, model selection (e.g., ARIMA, SARIMA, LSTM), hyperparameter tuning, and spatial analysis within the U.S. context.

Constraints:
Maintaining ethical standards in handling sensitive health data and ensuring compliance with U.S. data privacy regulations are of utmost importance. Limited computational resources may influence model complexity, and continuous updates are required due to the dynamic nature of the pandemic within the USA.

Stakeholders:
Stakeholders include U.S. public health authorities, federal and state policymakers, and healthcare institutions relying on accurate predictions for informed decision-making. The general U.S. public is an indirect stakeholder, with the model's outcomes impacting public health measures in the country.

Data Sources:
The primary data source is the COVID-19 dataset specific to the USA, obtained from Johns Hopkins University of Medicine, Corona Virus resource center. This dataset provides information on confirmed cases and deaths across states.