Project Overview: Machine Learning Model Selection

Introduction & Project Overview

- In this presentation, we will introduce our project and provide an overview of our goals and objectives. We will briefly describe the problem we are solving and highlight the significance of our project.

Predictive Use Case

- Describe the predictive use case, why it's important, and how it benefits your organization.

Dataset Selection

- **How and where you obtained the dataset**
- The dataset was obtained from [source].
- **Dataset description**
- The dataset contains [number] instances and [number] attributes.
- **Data size and attributes**
- The dataset is in [format] format and contains the following attributes: [list of attributes].

IBM Cloud Watson Studio

- Introduction to IBM Cloud Watson Studio
- Its relevance to your project

Importing the Dataset

- The dataset was obtained from XYZ source and was in CSV format. It contains X number of rows and Y number of columns.

Data Preprocessing

- **Data preprocessing steps**
- Handling missing values
- Encoding categorical variables
- Scaling numeric features

Feature Selection

- In order to build an effective machine learning model, it is crucial to select the most relevant features from the dataset. We used various visualization and analysis tools to identify the most important features and eliminate any redundant or irrelevant ones.

Machine Learning Model Selection

- We chose the Random Forest algorithm as it is suitable for our use case and provides high accuracy in predicting outcomes.

Training the Model

- **Model Training**
- After data preprocessing and feature selection, we trained our machine learning model using the selected algorithm.
- We used 80% of the dataset for training and 20% for validation to ensure the model's accuracy and generalization.
- The model was trained using IBM Cloud Watson Studio's built-in machine learning tools and libraries.