

## Project Initialization and Planning Phase

Date	13 March 2024
Team ID	738220
Project Title	Walmart Sales Analysis for Retail Industry with Machine Learning
Maximum Marks	3 Marks

### Project Proposal (Proposed Solution)

This project proposal outlines a solution to address a specific problem. With a clear objective, defined scope, and a concise problem statement, the proposed solution details the approach, key features, and resource requirements, including hardware, software, and personnel.

Project Overview	
Objective	The objective of the project is to develop a comprehensive sales analysis tool for Walmart, a leading retail corporation, leveraging machine learning techniques. By analyzing past sales data, store information, and promotional markdown events, the project aims to forecast future sales accurately. Specifically, the project aims to determine the impact of holidays, including Christmas, Thanksgiving, Super Bowl, and Labor Day, on store sales. The ultimate goal is to equip Walmart with a robust forecasting model that aids in informed decision-making, enhances performance predictions, and optimizes resource allocation.
Scope	The project scope encompasses the analysis of sales data from 45 Walmart stores, including store information and monthly sales. It involves identifying patterns and trends in the data, particularly focusing on the impact of promotional markdown events and holidays on sales performance. The scope also includes the implementation of machine learning algorithms such as Random Forest, Decision Tree, XGBoost, and ARIMA for training and testing the forecasting model. Additionally, the project involves integrating the developed model into a Flask application and deploying it on IBM Cloud for accessibility.
Problem Statement	
Description	The project involves analyzing historical sales data from Walmart's 45 stores to forecast future sales accurately. This analysis includes

	examining store information, monthly sales data, and promotional markdown events. The project will assess the impact of holidays, such as Christmas, Thanksgiving, Super Bowl, and Labor Day, on store sales, with holiday weeks weighted five times higher in the evaluation. Machine learning algorithms such as Random Forest, Decision Tree, XGBoost, and ARIMA will be utilized to train and test the forecasting model. Furthermore, the project entails integrating the developed model into a Flask application for interactive use and deploying it on IBM Cloud for accessibility and scalability.
Impact	The project's impact lies in its ability to provide Walmart with a powerful sales analysis tool that enables informed business decisions and enhances performance predictions. By accurately forecasting sales and identifying the impact of holidays and promotional events, Walmart can optimize inventory management, marketing strategies, and resource allocation. This, in turn, can lead to improved profitability, customer satisfaction, and competitiveness in the retail industry
<b>Proposed Solution</b>	
Approach	The project will follow a structured approach, beginning with data preprocessing to handle outliers and ensure data quality. Next, machine learning algorithms such as Random Forest, Decision Tree, XGBoost, and ARIMA will be applied to train and test the forecasting model using the historical sales data. The impact of holidays on sales will be evaluated, with holiday weeks weighted higher in the analysis. The developed model will be integrated into a Flask application for user interaction, facilitating easy access and utilization. Finally, the Flask application will be deployed on IBM Cloud or a similar platform for accessibility and scalability.
Key Features	<ul style="list-style-type: none"> <li>• Analysis of historical sales data from 45 Walmart stores</li> <li>• Assessment of the impact of holidays and promotional markdown events on sales</li> <li>• Utilization of machine learning algorithms including Random Forest, Decision Tree, XGBoost, and ARIMA</li> <li>• Integration of the forecasting model into a Flask application for interactive use</li> <li>• Deployment of the Flask application on IBM Cloud or a similar platform for accessibility and scalability</li> </ul>

## Resource Requirements

Resource Type	Description	Specification/Allocation
<b>Hardware</b>		
Computing Resources	CPU/GPU specifications, number of cores	2 x NVIDIA V100 GPUs
Memory	RAM specifications	8 GB
Storage	Disk space for data, models, and logs	1 TB SSD
<b>Software</b>		
Frameworks	Python frameworks	Flask
Libraries	Additional libraries	scikit-learn, pandas, numpy
Development Environment	IDE, version control	Google colab
<b>Data</b>		
Data	Source, size, format	Kaggle dataset