

Final Project – Proposal Submission

Project Title:

Sales Forecasting and Demand Prediction

Project Description:

The project aims to build a machine learning model that predicts future sales and product demand based on historical data. The model will help businesses optimize inventory management, staffing, and marketing strategies by providing accurate forecasts. The project involves data collection, preprocessing, feature engineering, model development, deployment, and monitoring to enable data-driven decision making.

Team Members & Roles:

Member Name	Role	Responsibilities
Ismaeel wassal elmselhy Email: esmailwassal456@gmail.com	Team Leader & Machine Learning Engineer	Lead the team, supervise all project stages, develop ML models (Random Forest / LSTM), evaluate and optimize models
Ali Hamed	Data Engineer & Preprocessing Specialist	Collect data, clean and preprocess datasets, handle missing values, duplicates, outliers, create basic features
Adel Elsharqawi	Advanced Data Analyst & Feature Engineer	Perform advanced data analysis, feature engineering (rolling averages, seasonal components, external factors), create interactive visualizations
Esraa Ali Mahmoud Judy Ahmed Mahmoud	Machine Learning Engineer	Select models (ARIMA, ETS, Random Forest, Gradient Boosting, LSTM), train and validate models, compare and select best-performing

Mustafa Elabady	MLOps & Deployment Specialist	model Manage experiment tracking with MLflow/DVC, deploy model as API/web service, create interactive dashboards, monitor model performance
Alaa Hesham	Documentation & Presentation Lead	Prepare final reports and presentation, document features, insights, business impact, suggest future improvements

Team Leader:

Ismaeel wassal elmselhy amer , Email: esmailwassal456@gmail.com

Objectives:

- Collect and preprocess sales and demand data.
- Explore data patterns, seasonality, and correlations.
- Develop accurate forecasting models using time series and machine learning approaches.
- Deploy the model with MLOps practices for real-time or batch predictions.
- Provide actionable insights to optimize business operations.

Tools & Technologies:

Python, Pandas, NumPy, Scikit-learn, Statsmodels, ARIMA, ETS, Random Forest, Gradient Boosting, LSTM, Matplotlib, Seaborn, Plotly, Dash/Streamlit, Flask/FastAPI, AWS/Google Cloud, MLflow, DVC

Milestones & Deadlines:

Milestone	Key Tasks	Deadline	Deliverables
1	Data Collection, Exploration & Preprocessing	7/10/2025	EDA Report, Interactive Visualizations,

			Cleaned Dataset
2	Advanced Data Analysis & Feature Engineering	12/10/2025	Data Analysis Report, Enhanced Visualizations, Feature Engineering Summary
3	Model Development & Optimization	17/10/2025	Model Evaluation Report, Model Code, Final Model
4	MLOps, Deployment & Monitoring	24/10/2025	Deployed Model, MLOps Report, Monitoring Setup
5	Final Documentation & Presentation	30/10/2025	Final Project Report, Final Presentation

KPIs (Key Performance Indicators):

Data Quality

- Percentage of missing values handled: 100%
- Data accuracy after preprocessing: 98%
- Dataset diversity: High coverage of product types, promotions, seasonal factors

Model Performance

- Model accuracy (Accuracy/F1-Score): 92%
- Model prediction speed (Latency): 150 ms per prediction
- Error rate (FP/FN): FP: 4%, FN: 3%

Deployment & Scalability

- API uptime: 99.5%
- Response time per request: 200 ms
- Real-time processing speed: 30 FPS (for batch updates in dashboard)

Business Impact & Practical Use

- Reduction in manual effort: 70%
- Expected cost savings: \$50,000 annually
- User satisfaction: 90%