Project Planning & Management

Project Overview

The Sales Forecasting and Optimization project aims to develop a predictive model that enhances sales forecasting accuracy for retail and e-commerce businesses. By leveraging historical sales data and advanced machine learning techniques, the project seeks to optimize inventory, marketing, and sales strategies, ensuring data-driven decision-making.

Project Objectives

- Collect, clean, and preprocess historical sales data.
- Perform exploratory data analysis (EDA) to identify trends, seasonality, and correlations.
- Develop and optimize a time-series forecasting model for accurate sales predictions.
- Deploy the model as a web-based application for real-time or batch predictions.
- Implement monitoring strategies to maintain model performance and mitigate data drift.

Scope of Work

Milestone 1: Data Collection, Exploration, and Preprocessing

- Acquire a dataset with features such as sales amount, date, promotions, and holidays.
- Conduct EDA, identifying trends, outliers, and correlations.
- Preprocess the dataset by handling missing values, engineering time-based features, and applying scaling techniques.

Milestone 2: Data Analysis and Visualization

- Perform advanced statistical analysis on sales correlations with factors like promotions and holidays.
- Visualize trends through interactive dashboards using Plotly or Dash.
- Prepare an analysis report summarizing key findings.

Milestone 3: Forecasting Model Development and Optimization

- Experiment with models such as ARIMA, SARIMA, Facebook Prophet, and XGBoost.
- Tune hyperparameters using Grid Search or Bayesian Optimization.
- Compare models using RMSE, MAE, and MAPE to select the best-performing approach.

Milestone 4: MLOps, Deployment, and Monitoring

- Implement model tracking with MLflow and version control with DVC.
- Deploy the model via Flask or Streamlit for real-time predictions.
- Set up monitoring systems to detect model drift and track performance.

Milestone 5: Final Documentation and Presentation

- Summarize findings, methodologies, and challenges in a comprehensive report.
- Develop a stakeholder presentation showcasing the model's impact.
- Provide recommendations for future improvements and business applications.

Team Members & Roles

Each team member is responsible for all data science tasks, including:

- Data acquisition and preprocessing
- Exploratory data analysis and visualization
- Model development and optimization
- Deployment and monitoring (excluding managerial oversight)

Key Performance Indicators (KPIs)

- **Model Performance:** RMSE, MAE, and MAPE thresholds for forecast accuracy.

- **Business Impact:** Improvement in stock optimization and revenue forecasting accuracy.
- **Deployment Success**: Model uptime, response time, and API latency.
- **Adoption Rate:** Percentage of stakeholders utilizing the forecasting tool.

Project Timeline & Deliverables

Milestone	Deliverables	Estimated Completion
Data Collection & Preprocessing	Cleaned dataset, EDA report	Week 2
Data Analysis & Visualization	Analysis report, interactive visualizations	Week 4
Model Development & Optimization	Model evaluation report, final model	Week 6
Deployment & Monitoring	Deployed model, monitoring setup	Week 8
Documentation & Presentation	Final report, stakeholder presentation	Week 9

Risks & Mitigation Strategies

Risk	Mitigation Strategy
Data Quality Issues	Implement rigorous preprocessing and validation steps.
Model Overfitting	Use cross-validation techniques and regularization.
Deployment Challenges	Conduct incremental testing before full deployment.
Model Drift	Set up automated monitoring and retraining mechanisms.

Conclusion

This project provides a structured approach to sales forecasting, integrating machine learning techniques to support business optimization. The deployed model will serve as a valuable tool for decision-makers, enhancing efficiency and strategic planning in sales and inventory management.