Internship Projects

Number of Orders Prediction

Objective:

To develop a machine learning model that predicts the number of customer orders for a product or service based on historical data. The aim is to help businesses forecast demand, optimize inventory, and improve resource planning.

Tools & Environment:

Python, Pandas, NumPy, Scikit-learn, XGBoost or LightGBM, Matplotlib, Seaborn, Jupyter Notebook

---

Dataset:

- Historical order data including features such as:
- Date/time
- Product ID or category
- Price/discount
- Marketing campaign status

- User behavior data (optional)
- Public datasets like Instacart, Dunnhumby, or a simulated e-commerce dataset can be used.

Key Work Done:

- Cleaned and preprocessed raw order data: handled missing values, encoded categorical variables, and normalized numerical features
- Performed Exploratory Data Analysis (EDA) to uncover trends, seasonality, and key predictors of order volume
- Engineered features such as day of week, holidays, moving averages, and lag variables
- Trained multiple regression models (Linear Regression, Random Forest, XGBoost) to predict the number of orders
- Evaluated model performance using RMSE, MAE, and R^2
- Visualized predictions and compared actual vs. predicted values over time
- Deployed final model for batch or real-time inference (optional)

Outcome:

Built an accurate predictive model that forecasts the number of orders on a daily or weekly basis. The model helps in decision-making for logistics, supply chain, and marketing strategies. Demonstrated proficiency in data preprocessing, feature engineering, regression modeling, and business-focused ML application.