ESTIMATING STOCK KEEPING UNIT USING ML

Project Planning

Team Members: Hemant, Sparsh Patidar, Khwaish

Project Timeline and Key Milestones:

Day 1: Project Setup and EDA

- Load and inspect the dataset
- Conduct exploratory data analysis (EDA)
- Visualize distributions, correlations, and missing values
- Define the business and technical problem clearly

Day 2: Data Preprocessing and Feature Engineering

- Handle missing values and outliers
- Convert categorical data into numerical form

- Create lag, rolling, and expanding window features
- Apply target encoding for categorical identifiers

Day 3: Model Development and Evaluation

- Train Random Forest and XGBoost models
- Evaluate models using R^2 Score and MAE
- Use RandomizedSearchCV for hyperparameter tuning

Day 4: Deployment and Presentation

- Build and test Flask web app
- Integrate trained model into backend
- Design frontend using HTML/CSS
- Prepare project documentation and present results

Tools and Technologies Used:

- Python (Pandas, NumPy, Scikit-learn, XGBoost)
- Data Visualization: Seaborn, Matplotlib
- Web Development: Flask, HTML, CSS
- Environment: Google Colab, Google Drive,VS CODE

Deliverables:

- Finalized predictive model with evaluation report
- Fully functional web app for predictions
- Comprehensive documentation
- Final presentation slide deck