Store Sales Prediction

Abstract:

The goal of this project, I will use machine learning to Store Sales Prediction to provide the stores with essential insights into the upcoming inventory and cash flow for the following months.

Data Description:

This dataset can be found at Kaggle. This dataset contains over 18000 rows with 8 features and after engineering features 16. The dataset that they provided contains the information of sales with:

- ID: Unique identifier for a row
- Store id: Unique id for each store
- Store type: Type of the store
- Location type: Type of the location where the store is located
- Region Code: Code of the region where the store is located
- Date: Information about the date
- Holiday: If there is a holiday on the given date
- Discount: If the store offers a discount on the given date

Data Source:

Supplement Sales Prediction | Kaggle

Design:

Sales forecasting is an essential task for the management of a store.
Machine learning can help us discover the factors that influence sales in a retail store and estimate the number of sales in the near future.
In this post, we use historical sales data of a drug store chain to predict its sales up to one week in advance.

Algorithms:

• feature engineering:
LabelEncoder

• models:
RandomForestRegressor
GradientBoostingRegressorDecisionTreeClassifier
XGBRegressor

Tools:

There are tools that will be used to achieve the goal of this project, such as:

- Numpy.
- Pandas.
- Matplotlib.
- The work will be done through Jupyter Notebook.