Forecasting Walmart Sales with Machine Learning

MGMT59000 Machine Learning

Seonkyu Kim
Yu-Hui Lin

Table of contents

01 Project Intro

Background description and project goal

Background

Project Goal

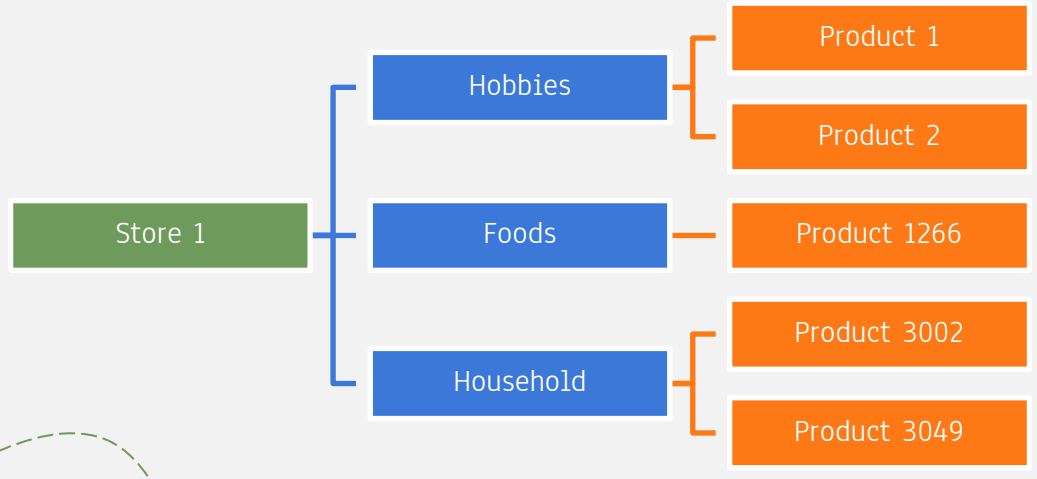
Apply machine learning practice to help Walmart better forecast future product sales

Scope

Future 28-day daily product sales of 10 stores in California, Texas, and Wisconsin

Datasets

- 10 stores across California, Texas, and Wisconsin
- 3,000+ products
- 5yr sales & price time series data (Jan., 2011 – Jun., 2016)
- Holidays and Events



02

Data Preparation

Data preprocessing and feature engineering

Supervised Machine Learning

Features

- 5-year historical prices and sales
- Events and context of products

Predicted Sales

- Forcasting horizon: 28 days ahead
- Loss function: Mean square error

Feature Processing

Historical Sales	Daily Price	Context	Calendar
Lagged sales	Price	State	Event
Sale std.	Max Price	Store	SNAP*
Sale moving average	Min Price	Category	
	Price std.	Product	
		Department	

* Supplemental Nutrition Assistance Program(SNAP), provides food benefits to low-income families to supplement their grocery budget.

Data Analysis

Model building, selection, and comparison

Model Comparison

	Gradient Boosting	Transformer	Ensemble
Structure	• 10 Models • 1-1913d Training • 1914-1941d Validation • Iterations=150 • Learning rate=0.01	• 10 Models • 1885-1913d Training • 1914-1941d Validation • Attention heads=12 • Dropout=0.2	Two models combined
MSE (validation)	6.23-9.38	0.47-3.92	-
Private Score*	2.5986	0.90539	MAX: 1.42651 AVG: 1.5558 MIN: 3.11345

* Private score is based on the submission score on Kaggle. A sample submission of "0" default value for all products scores 5.3907

Another Approach

Reduce Features

- Historical sales from day 351
- One day before event

Train

Various Neural Network Models

DNN

1D_CNN

LSTM

Transformer

Another Approach (Result)

DNN

Private Score = 1.5520

1D_CNN

Private Score = 1.0079

LSTM

Private Score = 0.6860

Transformer

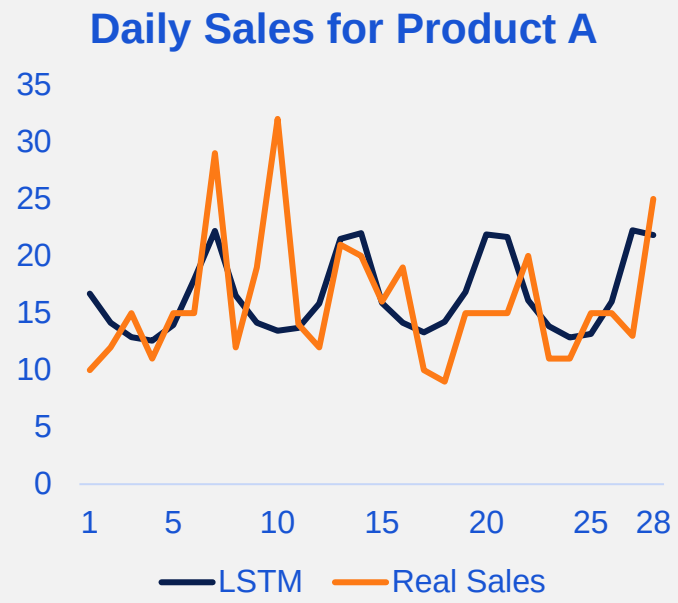
Private Score = 3.6327

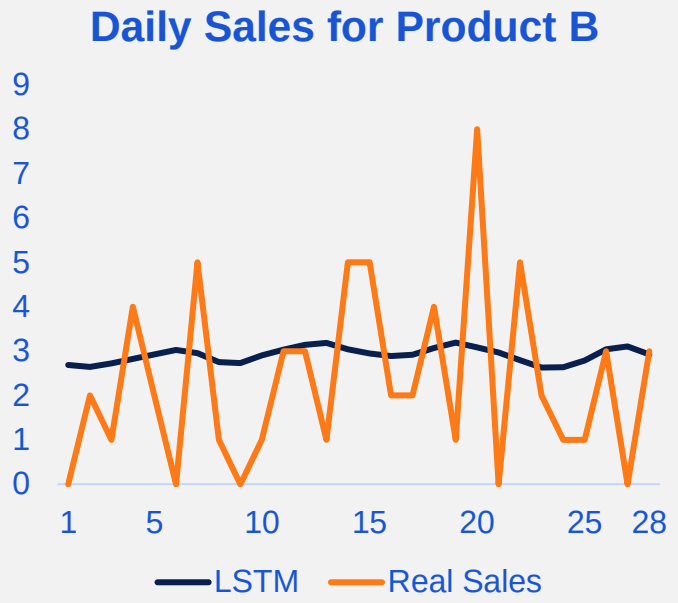
04

Performance Evaluation

Model prediction results

Forecasted Sales



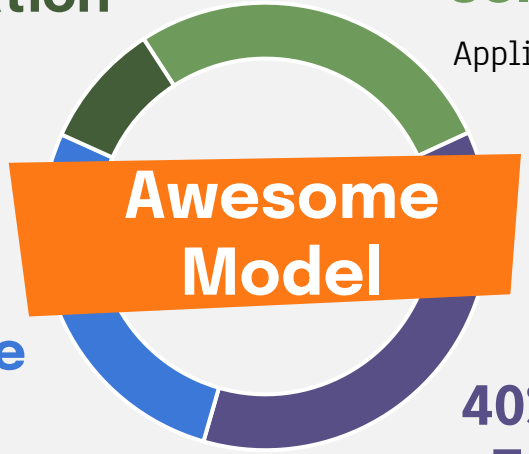


Conclusion

10% Data Manipulation

25% Model Selection

Applicable & Suitable



25% Feature Selection

Common sense vs. Bold guess

40% Model Training

Hardware matters !

Thanks!

Do you have any questions?