

A Practical Guide for Demand Prediction in Retail

Including Detailed Implementations and Notebooks

Maxime C. Cohen, Paul-Emile Gras, Arthur Pentecoste, Renyu Zhang

I/ Introduction	5
Motivation	5
Dataset	7
Objective and Scope	12
Prediction Accuracy Metrics	12
Training and Test Data	13
Application	14
II/ Data Pre-Processing and Modeling Factors	15
Dealing with Missing Data	15
Testing for Outliers	18
Accounting for Time Effects	20
Price and Lag-Prices	22
Featured on Main Page	23
Item Descriptive Features	23
Additional Features	24
Scaling	24
Sorting and Exporting the Dataset	25
III/ Common Demand Prediction Methods	26
Primer: Basic Linear Regression for One SKU	26
Structuring the Dataset	28
Centralized Approach	28
Decentralized Approach	30
Feature Selection and Regularization	31
Subset Selection	32
Presentation of Subset Selection	32
AIC and BIC	33
Lasso Regularization	37
Lasso for One SKU	37
Decentralized Lasso	40
Ridge Regularization	40
Ridge Model	40
Decentralized Ridge	41
Elastic Net Regularization	42
Elastic Net Model	42
Decentralized Elastic Net	42
Log Transformations	43
Log-Transformation on the Price Variable	43
Log-Transformations on the Target Variable	46
Applying transformations and prediction accuracy	48
Centralized Approach with SKU-Fixed Effects	49
Centralized Approach with Price-Fixed Effects	52
Centralized Approach with SKU-Price-Fixed Effects	56
Decentralized Approach with Aggregated Seasonality	57

Summary and Next Steps	59
IV/ Tree-Based Methods	61
Decision Tree	62
Centralized Decision Tree	62
Selecting the Parameters	63
Focusing on the Best Model	67
Example of a Plotted Tree	67
Decentralized Decision Tree	69
Selecting the Parameters	70
Focusing on the Best Model	71
Random Forest	72
Centralized Random Forest	72
Selecting the Parameters	72
Focusing on the Best Model	74
Decentralized Random Forest	74
Selecting the Parameters	74
Focusing on the Best Model	76
Gradient Boosted Tree	76
Centralized Gradient Boosted Tree	77
Selecting the Parameters	77
Focusing on the Best Model	79
Decentralized Gradient Boosted Tree	79
Selecting the Parameters	79
Focusing on the Best Model	81
Methods Comparison	81
V/ Clustering Techniques	83
K-means Clustering	83
Description of K-means Clustering	83
Clustering using Average Predictive Variables	86
Different Choices of Features	88
DBSCAN Clustering	93
Description of DBSCAN Clustering	93
Clustering using Average Predictive Variables	97
Different Choices of Features	101
VII/ Evaluation and Visualization	102
Summary of Results	102
Prediction vs. Actual	104
Varying the Split Ratio	108
VIII/ More Advanced Methods	113
The Prophet Method	113
What is the Prophet Method?	113
How it Works	113

Illustration for One SKU	115
Forecasting with Prophet	118
Univariate Time-Series	118
Adding Features	121
Data Aggregation and Demand Prediction	125
Presentation of the DAC Model	125
Tuning the Hyperparameters	127
Interpretation of the Outcomes	129
VIII/ Conclusion and Advanced Topics	131
References	134