
Online Sales Analysis and Prediction

Ruchitha M. Shanmugha Sundar, Meghna Venkatesha
Northeastern University
{midigarahallishanm.r, venkatesha.m}@husky.neu.edu

1 The Problem

In recent times Online Shopping is growing world wide. This has presented us with various oddities that we need to comprehend, to understand customer behaviour online and identify purchase pattern. This information can be used to come up with marketing strategies to improve sales and strategies for inventory management.

2 Proposed Analysis and Solution

We are looking to use the online retail data-sets available at <https://www.kaggle.com/vijayuv/onlineretail#OnlineRetail.csv> for our project. Our plan involves exploring the data-set to analyze all the attributes, identifying missing data and summarize the statistics based on our exploration as part of first step. Next use association rule to make suggestions for retailers. Followed by clustering to identify cluster of customers based on the transaction details and also perform time series analysis of the sales. If time permits, we will work on making few predictions using few Machine Learning models that can be used by retailers.

3 milestone

Table 1: Project Milestones

Weeks	Tasks
1-2	1) Data Selection, Prepossessing, Cleaning and Transformation.
2-4	1) Derivation of implied/calculated features, Selection of complete set of features. 2) Performing clustering using clustering models like k-means. 3) Association rules for business model and Time-series analysis
3-5	1) Applying Data Mining Models to make predictions 2)Project Write-Up.

References

- [1]<http://robotics.stanford.edu/~ronnyk/lessonsInDM.pdf>
- [2]<https://pdfs.semanticscholar.org/4879/a42015f48bda19ada8ae9707abef284253e6.pdf>
- [3]http://www.kdd.org/exploration_files/D MBA08 - Giering.pdf