1. Introduction

1.1 Background

Grocery stores are unquestionably an essential part of urban life. From sandwiches to ibuprofen, local grocery stores are perhaps the third mostly visited places in one's day, apart from his or her flat and office. With the popularity of mobile apps and experience-sharing websites, ratings from vast amount of users are shared and facilitate the exploring of other users to the venue. Under this backdrop, it would be of interest to a curious user or the management of the grocery stores to find out the driving forces behind the rating of the store. Which factors could provide explaining power to the different ratings across those seemingly identical stores? Although the level of services and the completeness of variety of goods could be valid answers to address this question, those data are generally difficult to for an outsider to obtain. Rather, the relevant exogenous factors, such as the location of the store, the house prices of the neighbourhood, the competitive environment, are available through location data providers and other free online resources, and these factors are the focus of this research paper.

1.2 Define the problem

The key question of this research is to uncover the exogenous factors that drive users' ratings of a grocery store. To narrow down the problem, London and Tesco are chosen as the city and the grocery store company of this analysis. It is constructed as a classification problem and the user ratings are the labels to be predicted. The final fine-tuned machine learning model will provide the best explanatory inputs that contribute to the ratings.