Consumer Data Research Centre

An ESRC Data Investment

CBRE

Understanding footfall: using GPS data to define four shopping areas in London

Runsang Yu[1], Michael Batty[1] and Nick Samson[2]
[1]University College London, [2]CBRE

Project Background

With the rapid development of wireless communications technology and the smart mobile terminal recording mobile object trajectories, opportunities are created to obtain novel insights into human behaviours. Traditional human behaviour analysis relies on manual collection of data including questionnaires, which leads to low efficiency and accuracy. However, with the spread of smartphones, individual activity chains can be studied in detail. Retailers are aware that using this new technology generates new insights in commercial industry, and every year they are trying to implement new technologies which could give them a deeper insight into where people visit and stay in retail areas.

Data and Methods

This study will explore GPS data obtained from the CBRE Group for four shopping centres – Oxford Street, Westfield Shopping Centre, Kensington High Street, and Stratford Westfield City – for 2017. To explore the critical factors which may contribute to the footfall count, machine learning methods including random forest, extreme gradient boost, artificial neural network and deep learning were adopted to predict the footfall in the four shopping areas by considering several temporal factors and weather conditions. Furthermore, to analyse the semantic trajectory data of each individual generated by GPS, the retail attractiveness score is defined according to the stay points detected from each individual's trajectory. The trajectory similarity is calculated taking space, time and semantics into consideration to figure out whether people who come from the same ward in London are more likely to have high similarity in their trajectories.

Key Findings

The influence of the variables is shown in Figure 1. December is shown to be the factor that contributes most to the footfall count in three of the shopping areas. However, the boost in consumption generally brought by December does not exert a great impact on Kensington High Street. Instead, individuals who go shopping in Kensington High Street are more affected by weather conditions, owing to the fact that it is an open-air shopping area. This is also true of Oxford Street to a lesser degree. Furthermore, among the four machine learning methods used in this study, deep learning performs the best which shows the highest R square and lowest relative mean square error of the comparison between predicted footfall count values and original footfall.

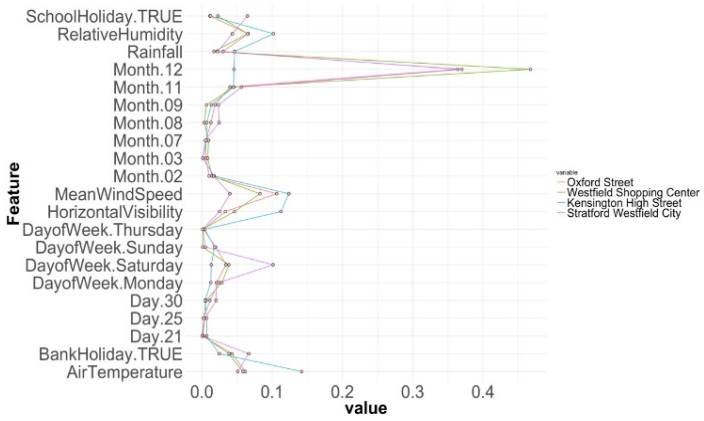

Figure 1: Feature importance

In respect of the semantic trajectory analysis, the attractiveness of each retail shop in the four studied areas in different time periods are defined using staying points. Using Oxford Street as an example, our research found that clothes and shoes stores account for nearly half of the overall attractiveness score during the daytime, but this gradually decreases from 6pm to 9pm. Meanwhile, the attractiveness score of coffee shops and supermarkets rise steadily from 9pm to 8am. This indicates that people tend to go to clothes or shoes stores during the day, and are more likely to visit coffee shops or supermarkets in the evenings.

The trajectory similarity values of the four studied shopping areas were calculated as of 1st January 2017. Oxford Street had the highest median similarity value (0.48), and Westfield Shopping Centre had the lowest median value (0.41). It is not unreasonable to infer that people that come from the same place (as a ward scale) share similar trajectories, but further analysis needs to be developed in the future to help explore this hypothesis.

Value of the Research

Traditional urban planning methods use rule-based logic that may not be able to adapt or benefit from modern technology, which captures data at a finer resolution in space and time. With the rapid development of GPS technology on handheld devices, it is possible to unlock the trajectories of large numbers of people at a finite scale. This study as indicated just one of the means through which it can be repurposed to harness better information on high street footfall.