Twitter as an Alternative Review Site

By Sinéad Dickson.

University of Dublin, Trinity College 2019.

Supervised by Professor Séamus Lawless and Anirban Chakraborty.

The increasing amount of information available on the internet means that recommender systems are growing in importance. Recommender systems help users to overcome information overload. They aim to help users to quickly find information relevant to them.

Online reviews are an important source of consumer opinions. They provide large quantities of data on consumer preferences and opinions. Twitter is a widely used microblogging social media platform, with 126 million daily users and 500 million tweets posted per day. Tweets posted to the site can often take the form of a review. This project will focus on harnessing these reviews and using them to help generate recommendations in the CoRE recommender system.

This dissertation proposes a method of first extracting review-like tweets from Twitter and then incorporating the sentiment of those tweets into a recommender system. This research aims to evaluate to what extent Twitter can provide a suitable source of online reviews that can be used effectively in the generation of recommendations in a recommender system. The project focused on reviews about hotels in Dublin.

This research explores the use of various classification algorithms and feature representations to classify whether tweets contain reviews. The sentiment of these review-like tweets is calculated and used to re-rank the CoRE recommender system.

The classification results were promising, showing that text classification is a valid method of extracting tweets from reviews. The best performing classifier was the Support Vector Machine. It achieved a precision score of 74%, a recall score of 74%, an f1-score of 73% and an accuracy score of 74.4%.

Incorporating the sentiment score into the CoRE had the desired effect and adjusted the rankings of the hotels. However, in terms of mean percentile rank (MPR) SentiCoRE performed worse that CoRE.