Predict 452 Section 55: Web and Network Analytics

Assignment 2: Website Analytics

Jack Doig

Management Problem

Understanding how a site is navigated allows for better understanding of the user, demand, and how to tune a site according to optimize the experience. Being able to offer intelligent recommendations to a user of a website can increase revenue as users can easily find what they are looking for.

Purpose of Study

The purpose of this study is to understand user behavior for an ecommerce site (a wine merchant registered in Australia). Specifically the goal is to understand the site in terms of pages visited and time spent on pages, to use a Network model to understand the user behavior types, and finally to predict which areas of the website a user might want to visit next, based on an Apriori association rule model.

Available Data

I was able to source an hour of traffic for a single-page ecommerce site, which sells wine to the Australian market. The data required significant cleaning so as to be usable, both removing bot data as well as cleaning the data.

Areas visited

The site consists of the ability to filter by type (white, red, sparkling), by grape variety, and then visit the wine pages themselves. The histograms below shows the visits broken down into areas visited, and average time spent on a page by page category
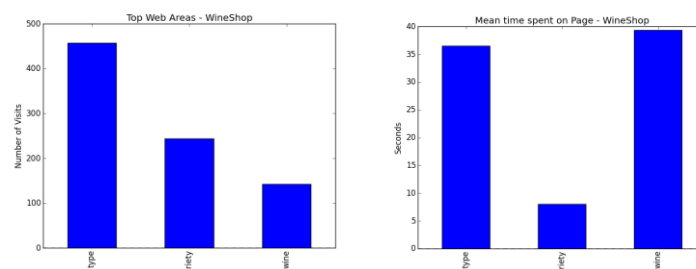
This shows most browsing is done through the 'type' pages, while shoppers spend the most time on the wine pages.

Modelling as a Network to understand user behaviors

The directed network diagrams below shows an individual user networks, as well as the degree distribution for the whole network.
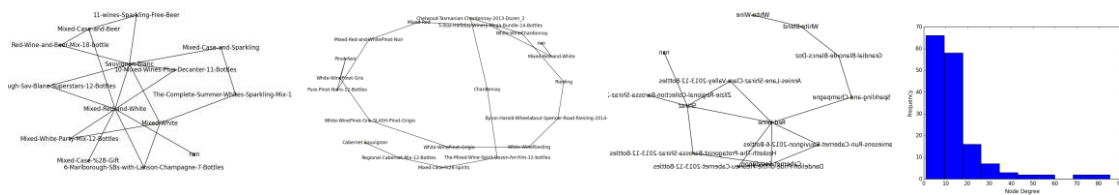
This shows there are a range of browsing behaviors; centralized (supported by the degree distribution graph), where users return to the search page, linear, where users don't click back, but browse organically, and a combination of the two.

Using the network to generate recommendations

I developed an Apriori algorithm in order to use the network of visitation data to predict likely pages that a visitor would visit. The way this works is through treating the pages visited by user as a set, and finds strong associations between item by comparing these sets. An example output of this, applied to the 'Shiraz' Page gave the following recommendations: 'Red-Wine','Biscay-Road-McLarenVale-Shiraz-2014-12-Bottles', 'Mixed-Red-and-White', 'Shiraz','White-Wine'. Although given the lack of data, I had to weaken the required support (alpha) parameter.

Next Steps/Recommendations

Although more data is needed, there seems to be little searching by grape variety, so it is questionable whether that mode of search should be retained. The site could rather be simplified. The typical user behavior tends to centralized around a type page which could influence UX design. Finally, there is value in an Apriori-based recommender system.