Determining Optimal Business Location by using Existing Customer Reviews

Ciarán Porter

Dublin City University, Ireland
ciaran.porter5@mail.dcu.ie

Abstract. This paper investigates how a business owner could decide where to locate a new venture through the use of data analytics techniques. Rather than rely solely on domain knowledge to make this decision, customer review data can be leveraged off to come to a conclusion. The key features which determine why some customer reviews score more highly than others are investigated, as well as the attributes that make particular areas of a city more attractive than others. Yelp reviews from the city of Edinburgh have been chosen as the test case for this analysis. Much of the existing work in the field has focused on North American cities, so a European city has been selected to determine if review behaviour is fundamentally different. The features examined include those inherent in the data, but also additional underlying latent variables. These features have been derived using a variety of techniques, including k-means clustering and sentiment analysis. A series of machine learning models have been developed and evaluated to determine the significance of such features. This in turn drives a scoring model that ranks prospective locations. The results generated from this model show the top 3 recommended areas for 30 different business types spanning an 11-year period. These locations are chosen for both their attractiveness, as expressed by customers and as areas where future growth opportunities exist.

Keywords: location modelling; sentiment analysis; machine learning

1 Introduction

Any business that wishes to expand its offerings in an urban area will need to address the question of "where is the best place for us to locate?" A number of factors are typically considered when making this decision. These include the cost of renting property, ease of access for deliveries and customers, and the demographics of the target market. While these are all market forces that are essentially outside of the control of business owners, customer review information can be used as a data-driven source to arrive at an optimal solution to this question. A number of organisations such as Amazon, TripAdvisor and Yelp have built successful business models around collecting customer review information. This paper looks to focus on how Yelp data, provided as part of its Yelp Dataset Challenge [1] can help determine the optimal location for a prospective business venture.

As of June 2017, the publicly available Yelp dataset provides information on over 4 million customer reviews across 11 cities. The full review text, star ratings, as well as over a million business-attribute permutations (e.g. parking facilities, Wi-Fi availability, accessibility, etc.) have been provided for approximately 144,000 unique businesses. This paper will focus on an analysis for the city of Edinburgh. Much prior research in this domain has concentrated on North American cities, so using a European city as a source represents a unique challenge in addressing the question of prospective business location. The development of European cities outwards from historic centres differs from the grid-based system that is common in North America. Therefore, it will be interesting to see if this impacts on the location distribution of business types. In terms of the focus of this paper, we will discuss methods of detecting user sentiment, and in particular, aspect-based sentiment in reviews; establish the key factors (including latent factors) that determine star ratings, and use this information to recommend future location for businesses in a variety of sectors to settle in.

2 Overview of Dataset and Preparation

Prior to the machine learning and modelling phase of this paper, a series of data preparation steps were performed. This ensures that an accurate dataset can be used for analysis and helps unearth hidden features present in the data. These new features can then be used as part of the later machine learning and modelling steps.

2.1 Overview of the Yelp Dataset

The primary data source to be used as part of this analysis comes from Yelp's Dataset Challenge. As the focus of this research is on determining optimal business location in a city, the scope of the work has been narrowed to include reviews from Edinburgh only between 2006-2017. The dataset sample comprises of 44,069 customer reviews for 3,617 distinct businesses, submitted by 8,341 users. Indeed, the review data contains feedback in 12 different languages, although English is the predominant language used. Where it was feasible, analysis was performed using the original language of the text. However, an English language version of each text was also derived using Google Translate where this approach failed. As well as the review texts and star ratings, the dataset contains information on the various business attributes and a number of derived features that are discussed over the course of this paper.

2.2 Geohashing of Business Locations

Each business is assigned a latitude and longitude co-ordinate in the Yelp dataset. As this paper looks to identify the most appropriate area for a business venture to locate in, a method of grouping these co-ordinates into a higher level of abstraction is required. *Geohashing* is a process of encoding latitude and longitude co-ordinates as strings by dividing a map into grid, with each quadrilateral in the grid representing a distinct area. The length of the string determines how large an area is represented. For example, a two-character string is enough to cover most of Ireland and Great Britain, whilst three characters could cover an area approximately the size of London [2]. For the purpose of this analysis, we have grouped the list of Edinburgh businesses present in the review dataset into their accompanying geohash strings of length 6 and 7. This equates to areas of 0.732 km^2 and 0.023 km^2 respectively. An R script was run to assign geohash values to each unique business location. For the remainder of this paper, we will reference the geohash strings of length 6 as *areas* and length 7 as *localities*. In this instance, a total of 193 distinct *areas* and 835 *localities* were derived.

2.3 Deriving Business Categories

Each business in the Yelp review dataset contains a number of keywords that describe the type of product or service being offered. The Edinburgh dataset contains 453 keywords that relate to different types of businesses. Obviously, having this many categories in the dataset is impractical for analysis. Thus, clustering techniques have been employed to assign each business into one of 30 distinct categories. Each business contains a number of keywords describing its activity. For example, a business containing the keywords "beauty & spas", "hair salons" and "barbers" would be grouped under the overall category of "Beauty".

The process for deriving a fixed set of business categories requires a two-phase approach. The first stage involves ranking the frequency of keywords and assigning relative ranks for each business-word pair. For example, a business with the keywords "restaurant" and "deli" would have "restaurant" assigned its number 1 ranked value, as this is a more common term generally across the dataset. Each business was assigned its list of keywords in order of significance using a pivot table in SQL. The second stage involves grouping similar lists of terms into distinct clusters using a k-means approach. A document term matrix was derived, with each business ID considered the document reference and the list of its keywords the text. The input to the cluster model was a normalised term frequency-inverse document frequency matrix. This was then clustered into 30 groups. Each business was then assigned a single value from the newly summarised category set.

2.4 Sentiment Analysis

2.4.1 Aspect-Based Sentiment Analysis

The second piece of work involving sentiment analysis that was investigated related to aspect-based sentiment. Rather than performing traditional sentiment analysis, whereby a piece of text is classified as having either positive or negative overall sentiment, this approach looks to identify individual topics present in the text and determine the level of sentiment polarity at a topic level [3]. The Sentiment Analysis API developed by *MeaningCloud* allows for aspect-based examination of batches of review text. An aspect-based review was performed on all English, French, German, Spanish, Italian and Portuguese texts using MeaningCloud's Excel add-in. The software is currently unable to perform such analysis on other languages, so these texts were converted to English before being run through the tool. The output provided for each review shows the key terms, their associated topics, and a sentiment polarity value. These values range from very negative to very positive and include instances of neutral or no sentiment.

An analysis was performed on the derived topics and it was found that each sentiment aspect could fall into one of 14 categories. These ranged from “Food” to “Location” to “People”. For each individual review, its list of topics and polarity values were summarised. The top three most common topics for each review were retained and polarity scores for each were derived. These polarity scores had a value of between -2 and 2, with negative values representing negative sentiment expressed by the customer. The three topics and their review scores were then incorporated into the main review dataset as additional features.

2.4.2 Word-Emotion Association Lexicons

The final piece of text analytics performed on the Yelp review dataset relates to determining sentiment emotion in the review texts. As well as investigating if overall sentiment is positive or negative, the *syuzhe*t library in R allows us to estimate various sentiment emotions and their significance within the text. The list of potential sentiment types includes anger, trust and disgust, for example. Having these additional metrics may act as indicators as to why a review scored particular well or badly in terms of star rating.

An R script was run comprising of two main steps related to sentiment emotion. The first calculated the overall word counts and the proportion of positive and negative words in the text. This was achieved by comparing each text to pre-populated lists of “positive” and “negative” sentiment words. The ratio of positive to negative words was also calculated. The second step calculates the average sentiment scores and the word emotion lexicon scores [4] in each review text. Each sentence in the review text was assigned a score for *“anger”*, *“anticipation”*, *“disgust”*, *“fear”*, *“joy”*, *“sadness”*, *“surprise”* and *“trust”*. These scores were then aggregated at a review level. In addition, the overall sentiment standard deviation was calculated, which looks to see how sentiment fluctuates from sentence to sentence. As per previous sections, these score values were incorporated into the main review dataset, which will form part of the analysis described in Section 3.

3 Methodology

3.1 Modelling Techniques - Purpose

The overarching goal of this paper is to produce a model that can indicate what the best location is for a business venture to set up at a given point in time. Having performed the necessary data preparation, exploratory analysis and feature generation, two related datasets have been derived. The first of these datasets contains items at a review level. This includes information such as the review date, business name, business type, geographical location, business attributes, review topics and a series of sentiment score attributes. In total, the cleansed review dataset contains 74 features. The second dataset in use relates to a summarised version of the original review dataset. This contains metrics for each combination of business type, geographic area and year. In total, 46 features are included in this dataset.

For each dataset, a series of machine learning models will be run to predict star ratings. This in turn will help provide an indication of which features are most significant in determining such ratings. It is unlikely that all 74 features will be useful (or indeed all 46 in the case of the summarised dataset) in making this prediction. Therefore, this investigation will help filter out redundant variables, and allow us to obtain an understanding of the important attributes that determine a customer's sentiment about a particular product or service.

Each machine learning model iteration will enable us to rank features in order of significance. By reviewing the output of more useful models, we will be able to come to conclusions as to which features are generally important, and their impact on the overall rating. Appropriate weightings will be applied to each feature, relative to their overall model importance. These weightings will then be used as inputs to a final scoring model. By taking the most significant features, and applying the required weightings to them, we will be able to rank prospective locations and determine the optimal location for each business type. Such rankings will look to incorporate the overall performance of each location in scope, and the overall competitiveness of each area. Ideally, a business owner will be looking to find an area which scores highly in customer reviews, but is not saturated with existing direct competitors [5].

3.2 Machine Learning Models Overview

3.2.1 Predict Review Rating

The first machine learning model developed was designed to predict the performance of individual reviews. Note that the source data was heavily skewed in favour of 4 and 5 star reviews. It was found that having these star ratings as separate classes impacted on model accuracy, as most training models could not differentiate between a 4 and 5 star rating with the given features. In any case, a business owner is more likely to be concerned about differentiating between "poor", "average" and "good" reviews rather than individual star ratings. As a result, the star ratings target feature was reclassified as one of: *"poor"* (1 star), *"average"* (2-3 stars) or *"good"* (4-5 stars). A new *"star_rating"* feature, containing these three distinct levels was used as the target feature, for the purpose of this investigation.

A series of candidate machine learning models were selected for testing. For each model, the review dataset was split into a training set and a test set, using a 70:30 split. The target feature in each instance was the *"star_rating"* category. In the training dataset, up-sampling techniques were applied so that there was a balanced number of "poor", "average" and "good" reviews in the sample (i.e. the same number of records in each category). This was done in order to prevent a bias towards "good" reviews, as the dataset contains a naturally higher proportion of these values. Each model was then trained using k-fold cross-validation with k value of 5 and the caret library [6] in R. The metrics used for validating the accuracy of the models are described in section 3.4.

3.2.2 Predict Location Rating

In addition, the second iteration of machine learning models focused on the summarised version of the review dataset. In this instance, the average localised rating (*"ave_local_rating"*) was used as the target feature. The objective of this exercise was to predict the average rating for each business type in a particular locality and point in time. As this is an average score, it was retained as a numeric target feature, meaning no up-sampling adjustments were required on the training dataset. The training/test split was again set at 70:30. Six separate machine learning models were trained and evaluated for accuracy. For both sets of models, feature importance scores were calculated. The most significance features for predicting both the updated *"star_rating"* and the average localised rating were retained and used as inputs to the scoring models.

3.3 Scoring Models Overview

3.3.1 Review Level Scoring Model

The first scoring model that has been developed takes the key features from the review level machine learning outputs, applies numerical scaling where appropriate and assigns weights to each feature based on its relative importance in determining star rating. The assigned weights can be either positive or negative based on the feature's impact on the rating. An overall *"review_score"* was calculated, which performs a sum-product calculation across the weighted features. Each *"review_score"* was ranked to see which review scored highest within its particular business type (*"business_rank"*), area (*"in_area_rank"*) and locality (*"in_location_rank").* Ranked percentile scores were also calculated using these three metrics. The outputs from this method were then validated against the location level scoring model.

3.3.2 Location Level Scoring Model

The location level scoring model was developed in a similar fashion to the review level equivalent. Again, the most important features were appropriately scaled and weightings applied according to their relative significance. A sum-product calculation was performed across the features to derive an overall *"location_score"*. A check was also performed to see if business types in certain areas did not have any reviews submitted in a particular year. Where this was the case, the most recent available year's score for that location along with a penalty factor was used as a proxy score. This was implemented so areas that did not have any reviews in a particular year could be assigned a score, but were penalised for their lack of activity over the period of time. The *"location_score"* values were then ranked for each business type and year to derive a final *"review_rank"*.

In addition to the *"review_rank"* score, a *"location_rank"* value was also derived. This was calculated by averaging the rankings of each area's *"local_pull"*, *"local_saturation"* and *"area_saturation"* factors (these rankings were grouped by business type and year). The combination of the *"review_rank"* and *"location_rank"* was then calculated as the *"total_score"*. For each business type and year, the top 3 scoring locations were retained. These results were then validated against the review level model outputs.

3.4 Model Validation and Evaluation

3.4.1 Machine Learning Models

Balanced accuracy and the *Root Mean Squared Error* (*RMSE*) are the key metrics used to evaluate the results of the review performance and location performance models respectively. In the case of the first measure, balanced accuracy is a more appropriate metric than the generic accuracy term. This is due to the fact that this allows us to see the level of precision in predictions made across each of the three potential outcomes ("good", "average" and "poor"). This is particularly important when evaluating the test set, as this dataset is not subject to up-sampling, and is likely to be skewed in favour of positive reviews. For instance, if 90% of the test set contains a star rating that is deemed "good", then a model that simply predicts every review as "good" will still have an accuracy score of 90%, even with no training undertaken. The overall accuracy can mask the fact that a model performs poorly when predicting smaller classes. However, the balanced accuracy values ensure that we can see how well a model is performing in each class. This metric will be applied to the test set outcomes, and used to evaluate which models and features are useful in predicting star ratings. In the case of the review performance model, multiple iterations of each model type will be derived, by adjusting some of the relevant input parameters and the number of features in scope.

Validating the output for the location performance machine learning model is less complex. In this instance, the RMSE metric will be used to compare the different models in use. A model with a low RMSE score on both its training and test sets will be preferred. However, it is desired that the error values on both the training and test sets be closely aligned. If a model achieves a low RMSE score on its training set, but this error rate increases significantly on the test set, then it is likely to be suffering from over-fitting. Any outputs taken from models of this nature will not be considered for further analysis. Once again, the training and testing process will be performed on a number of iterations of each model type, with varying degrees of parameter and feature scope adjustment.

3.4.2 Scoring Models

The scoring models have been designed to rank individual reviews and overall locations. In order to assess the outcomes derived in these models, a comparison needs to be done between the location level rankings and the more detailed review level scores. For each business type and year, the top three location options in the form of geohashed localities were selected. The *in_location*, *in_area* and *in_business* percentile scores taken from the review level models were compared with their review level counterparts. The average review percentiles were then calculated for each of the three metrics above. Any results that scored above 0.5 as a percentile were deemed satisfactory, as this means that the location generally performs above average in terms of review scoring. This can then be deemed a suitable prospective business location. Details around the results generated are provided in Section 4.

4 Results

4.1 Review Rating Model

Six variants of machine learning models were executed, which attempted to predict the star rating of individual review items (*Table 1*). Initially, the training and test sets of each model incorporated the full population of 74 features with the newly derived *star_rating* retained as the target feature in each model. The number of features used was scaled back by investigating the variable importance of the attributes in each model. This allowed for more simplified versions of the models to be produced without compromising the final results. For each model, a number of its distinct input parameters were also tuned. The summary table below shows the result of testing performed on each model and their balanced accuracy values.

The key models that will be focused on, based on the testing outcomes, are the Neural Network and Linear Discriminant Analysis (LDA) models. The balanced accuracy scores for poor and good outcomes are relatively high for these models. This is in spite of the fact that the overall accuracy score on the LDA model is lower than some of the other alternatives. It does not appear as skewed towards higher rated reviews as other models tested. Similarly, the kappa scores for these models are marginally higher than the other alternatives, which suggest a greater level of precision across the three outcome types. The only exception to this is the Bagged CART model. This model is reasonable enough in terms of its predictive power, although it appears to suffer from over-fitting as seen by its drop in accuracy when moving from training to test data. The fact that this is a more complex model (uses all features) but is not significantly more powerful than the Neural Network or LDA models also makes it less desirable for continued use.

As has been discussed, the models selected contain a subset of features that will be carried forward to the next stage of the process. The key features in each model have been assigned relative weightings according to their predictive significance. Interestingly, it was found that certain features were likely to have greater importance on the review when it is classified as "good" or "poor", but this is less pronounced for "average" reviews. In total, 12 features were retained and average weightings reapplied to each. It is evident that the majority of these features relate to sentiment scoring values. The categorical and business attribute type features appear to have less of an impact on the final outcome. Perhaps this could be explained by the fact that these features are business-specific and do not tend to vary between reviews. For example, the presence (or non-presence) of Wi-Fi appears to have little impact on a review score, but the customer's quantified level of anger about the product or service does.

Table 1. Review Rating Model Outputs

Model	*No. Features*	*Training Result*	*Test Result*	*Balanced Accuracy*
Decision Tree	9	Complexity Parameter = 0.003 Accuracy = 0.57	Accuracy = 0.55 Kappa = 0.18	Poor = 0.79 Average = 0.54 Good = 0.65
Neural Net-work	12	Decay = 0.15 Size = 6 Accuracy = 0.61	Accuracy = 0.60 Kappa = 0.25	Poor = 0.80 Average = 0.59 Good = 0.68
Naïve Bayes	13	Complexity Parameter = 0.003 Accuracy = 0.62	Accuracy = 0.62 Kappa = 0.22	Poor = 0.77 Average = 0.55 Good = 0.65
Random Forest	9	No. Variables Sampled = 4 Accuracy = 0.92	Accuracy = 0.70 Kappa = 0.23	Poor = 0.61 Average = 0.58 Good = 0.63
Bagged CART	All	Accuracy = 0.92	Accuracy = 0.70 Kappa = 0.27	Poor = 0.65 Average = 0.60 Good = 0.65
Linear Dis-criminant Analysis	18	Accuracy = 0.60	Accuracy = 0.60 Kappa = 0.24	Poor = 0.80 Average = 0.58 Good = 0.68

4.2 Location Rating Model

At a location level, a further six distinct machine learning models were executed. Once again, all models were run with the full set of features, and then simplified with only the most significant features retained. The target feature being predicted was the average localised rating (*ave_local_rating*). Each model was compared using the RMSE value on both its training and test sets (*Table 2*).

Judging by the RMSE score on the test dataset, we can see that the features included in the Random Forest and Bayes Recurrent Neural Network (BRNN) model generate the most accurate predictions. Note that the random forest model generates its lowest RMSE when sampling 22 features. The fact that the number of features scales up to 22 is as a result of the introduction of dummy variables for categorical features. This stems from the *business_type* feature being converted into a series of binary features for each possible outcome, rather than retaining its original structure of having multiple categories. For example, Italian restaurants, pubs and education are all transformed into distinct binary variables rather than a single categorical feature. Furthermore, it is evident that the introduction of a Bayesian approach helps with the accuracy of the neural network model. It has been noted that this performs poorly in its generic form, even with a full set of features to choose from. For the purpose of feature selection for the final scoring model, the features from the Random Forest and

BRNN were retained, along with the outputs from the ridge regression model (although this model will not be considered as significant when assigning overall weights).

Table 2. Location Rating Model Outputs

Model	*No. Features*	*Training Result*	*Test Result*
Decision Tree	5	Complexity Parameter = 0.003 RMSE = 0.83	RMSE = 0.82
Neural Network	All	Decay = 0.0001 Size = 1 RMSE = 2.99	RMSE = 2.99
Ridge Regression	6	Lambda = 0.01 RMSE = 0.80	RMSE = 0.80
Random Forest	13	No. Variables Sampled = 22 RMSE = 0.78	RMSE = 0.78
Bagged CART	All	RMSE = 0.83	RMSE = 0.83
Bayes Recurrent Neural Network	7	Nodes = 2 RMSE = 0.78	RMSE = 0.78

4.3 Final Scoring Model

The features plotted below (*Figure 1*) show the attributes that have been extracted from the machine learning models, and used as inputs to the final scoring models. The colour shading provided indicates whether or not the weighting is positive or negative. In the case of the review scoring model, there is an almost even split of positively and negatively weighted attributes. However, the negative features are all assigned much smaller weights. The disgust score feature is the largest negatively weighted feature; it seems natural enough that this is the one feature most closely associated with negative reviews. Similarly, positive features dominate the location scoring outputs. Only 6 features have been retained, as the variable importance of any additional features was negligible across the various models tested. Perhaps the only non-obvious feature listed here is the average number of words. Generally, reviews with higher word counts do not score as well in terms of star rating. However, it must be qualified that the impact of this feature is relatively minor.

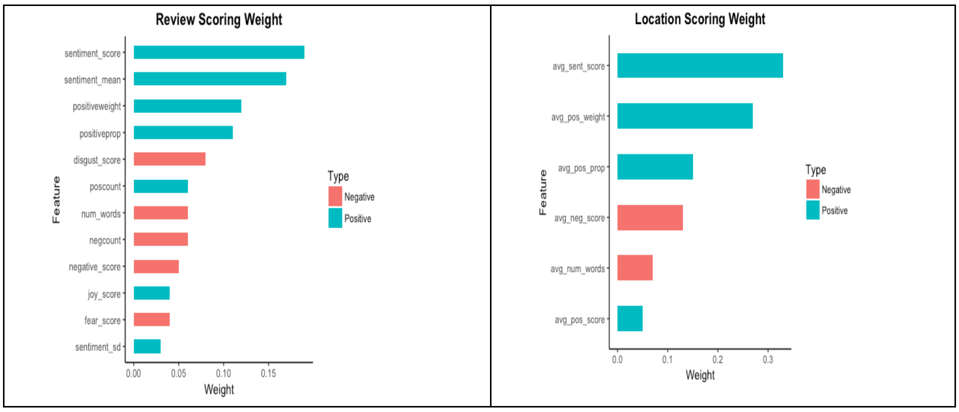

Fig. 1. Selected Features for Scoring

As has been outlined in the methodology, these metrics and their associated weightings were used as inputs to a scoring model, which generated a ranking for each locality. Using the location-summarised data, the top 3 ranked locations for each year and business type were retained. Each of these localities refers to a street level area. When comparing the top ranked locality values back to the corresponding review based scores, it was found that over 70% of the average business, area and locality percentiles were above the 0.5 mark. This means that the majority of ratings associated with these locations in the detailed scoring model are ranked in the top half of their respective business/year review groupings. Similarly, the equivalent maximum scores for each location are ranked in the top 5 percentile of outcomes. For example, the location with the *geohash* "gcvwr93" (St. Mary's) was selected as the most preferred overall location for Flowers & Gifts in 2009. This is supported by the fact that an individual review from this location was given the top ranking for the same business type and year. Furthermore, this process can be repeated for the minimum review scores. The selected values do not fall below the 0.35 percentile at an area and business level and 0.3 at a locality level. Thus, even the most negative reviews for these localities are still higher ranked than about a third of their direct competitors.

Therefore, by replicating this process across each business type for the period between 2006-2017, we can see which locations are most desirable in terms of setting up an enterprise. The top 3 most suitable locations have been retained, and are designed to fit the balance between having a target market that can be exploited but are not overburdened with existing competition. For example, in 2015, it was discovered that the top 3 selected locations are closely clustered together for their respective business types. Indeed, from analysing the recommended locations for each year in scope, it is clear that much of the attractive business areas are found in the city's "Old Town" district. However, over time, the attractiveness of this area begins to wane as the market appears to become more saturated. From 2010 onwards, there is a clear spreading out from the ”Old Town” district westwards towards the Princes Street area, and southwards towards the city's university district. This "sprawl" effect is characteristic of many historic cities, particularly in a European context. Of course, it is important to remember that the results generated are meant to provide guidance for prospective business owners. As has been mentioned, other external factors such as the price of property and access to transport facilities are also likely to be considered prior to selecting a location.

5 Conclusions and Recommendations

This paper has discussed the steps required for addressing the question of finding the optimal location for a new business venture. Machine learning techniques have been implemented to analyse the key factors that contribute to star ratings in Yelp user reviews, and validate their significance. This paper has incorporated a number of methods discussed in the existing literature, such as topic modelling, sentiment analysis and business location modelling. It has looked to build on prior work and incorporate these processes into developing a solution that can be used in a practical context. The source dataset comes with a degree of noise and a number of attributes, including hidden features. By stripping away this noise, and unearthing what the most significant features are, the results can give insights to Edinburgh business owners as to which areas of the city are best suited to certain types of activities. Of course, the techniques discussed in this paper could easily be replicated in another city. Indeed, repeating this process in a larger city, where there is a greater depth of information on user reviews would be a worthwhile exercise for a future investigation. Furthermore, it would be interesting to see the impact of additional external factors such as property rent prices or the quality of local transport on any future work in this area. This may help quantify the impact of inherent location-based features, in addition to the user-sentiment features that have had such significance on the results presented in this paper.

References

1. Yelp Inc, “Yelp Dataset Challenge”, Retrieved from Yelp: https://www.yelp.com/dataset_challenge/, 2017.
2. Elasticsearch, “Geohashes”, Retrieved from Elasticsearch: The Definitive Guide: https://www.elastic.co/guide/en/elasticsearch/guide/current/geohashes.html, 2017.
3. S. Moghaddam and M. Ester, “The FLDA model for aspect-based opinion mining: addressing the cold start problem”, International Conference on the World Wide Web, pp. 909-918, New York: ACM, 2013.
4. S. Mohammad and P. Turney, “Crowdsourcing a Word-Emotion Association Lexicon”, Computational Intelligence, Issue 29 (3), pp. 436-465, 2013.
5. F. Wang, L. Chen and W. Pan, “Where to Place Your Next Restaurant? Optimal Restaurant Placement via Leveraging User-Generated Reviews”, ACM International Conference on Information and Knowledge Management, Issue 25, pp. 2371-2376, New York: ACM, 2016.
6. M. Kuhn, “Caret Package”, Retrieved from Cran R Project: https://cran.r-project.org/web/packages/caret/caret.pdf, 2017 Author, F.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010).