# Twitter Sentiment Analysis of New IKEA Stores Using Machine Learning

Yujiao Li

School of Technology and Business Studies

Dalarna University

Borlänge, Sweden

yli@du.se

Hasan Fleyeh
School of Technology and Business Studies
Dalarna University
Borlänge, Sweden
hfl@du.se

Abstract—This paper studied public emotion and opinion concerning the opening of new IKEA stores, specifically, how much attention are attracted, how much positive and negative emotion are aroused, what IKEA-related topics are talked due to this event. Emotion is difficult to measure in retail due to data availability and limited quantitative tools. Twitter texts, written by the public to express their opinion concerning this event, are used as a suitable data source to implement sentiment analysis. Around IKEA opening days, local people post IKEA related tweets to express their emotion and opinions on that. Such "IKEA" contained tweets are collected for opinion mining in this work. To compute sentiment polarity of tweets, lexiconbased approach is used for English tweets, and machine learning methods for Swedish tweets. The conclusion is new IKEA store are paid much attention indicated by significant increasing tweets frequency, most of them are positive emotions, and four studied cities have different topics and interests related IKEA. This paper extends knowledge of consumption emotion studies of prepurchase, provide empirical analysis of IKEA entry effect on emotion. Moreover, it develops a Swedish sentiment prediction model, elastic net method, to compute Swedish tweets' sentiment polarity which has been rarely conducted.

Index Terms—big-box effect, opinion analysis, customer emotion, elastic net model, text mining, natural language processing

#### I. INTRODUCTION

Both society and economy development determine factors for government investment decision making. Having a new IKEA store is costly in budget plan, so its predicted benefits and risks of both aspects are considered prudently in advance for local government. From the perspective of economy, IKEA entry can promote regional economic development because attracted customers improve durable-goods revenues and increase the number of employees in the regions [1].

However, its social effects has not been studied which is critical for regional government and IKEA store because positive response can improve governments' reputation and boost local business environment, while negative opinions cause adverse influence, such as direct loss of investment and potential less popular support of management. Consumer sentiment index has predictive power on its own for future changes in consumption spending [2]. Two possible interpretations: first, sentiment is an independent driving factor in the economy, and that changes in sentiment not only

forecast changes in spending but also cause them; second, sentiment reflects the overall outlook for the economy, i.e. when consumers are optimistic about the outlook for the economy, they give upbeat responses to interviewers. Hence, on average, those optimistic expectations are substantiated, and spending eventually increases as foreshadowed by sentiment.

Thus, this paper aims to evaluate public opinions on IKEA entry by studying various traditional and new approaches in natural language processing. Four cities are investigated, namely Umeå-Sweden, Halifax-Canada, Fishers-Ireland, and Sheffield-UK which sequentially have new IKEA stores in 2016 and 2017.

Data are collected from Twitter, a micro-blogging platform, which has been widely used for analysis of sentiment ("positive", "negative" or "neutral"). The crawled twitter data are those tweets include "IKEA" in investigated cities during individual IKEA entry period. Languages investigated in this work are English and Swedish. English-based tweets are state of the art in sentiment analysis, while there is very few studies in the Swedish-based sentiment analysis. The challenge is Swedish tweets' opinions cannot be identified directly due to lack of developed tools for sentiment analysis. While English has been studied well in natural language processing field as literature review shows. To identify opinions of those tweets in Swedish, the basic idea is to fit a sentiment prediction model for Swedish based on training data which are opinions labeled tweets in Swedish. Then apply fitted model to predict those "IKEA" tweets in Swedish. Hereby, the optimal fitted model is gained by examining and comparing several machine learning classifiers;

To collect sentiment labeled tweets in Swedish, this work followed the same procedure as [3] who demonstrated the validity of using emoticons. The authors query tweets for two types of emoticons: positive emoticons such as ":-)" and negative emoticons like ":(" etc. This approach would be biased if the text content express multiple conflicted opinions. Because each tweet is usually composed of a single sentence which is restricted to be less than 140 characters, it is reasonable to assume emoticon can represent the whole message. So far, the work can be quantitatively addressed as following questions:

• Does number of "IKEA" tweets increase significantly after entry day?

- How do positive and negative tweets evolved in scale?
- What do people talk about when they talk about IKEA?

Contribution of this work is two-fold: in terms of method, it presents the results of elastic net approach and compare its performance with other machine learning classifications in Swedish sentiment analysis. In terms of empirical studies, it detected IKEA entry causal effect. The rest of this article is organized as follows. Section II discussed related work focusing on Swedish tweets research and applied approaches in sentiment analysis. Section III describes procedure of analysis including data and methods.the proposed method. Section 4 presents the results and Section 5 is the conclusion.

## II. RELATED WORK

Prior studies of IKEA entry focused on its economic effects in retail industry. On one hand, sales and employment increases when Ikea enters a municipality [1]. On the other hand, IKEA entry causes spill-over effects on surrounding business areas [4]. However, the effect of IKEA entry on residents' sentiment has not been discussed which is meaningful for government decision, related investors and sociologists.

Literature validated using emoticons as emotion signal which provide referential distant supervision approaches for this work. There is a number of related work using emoticons in English-based twitter sentiment analysis. Wang et al. examined sentiment expression via emoticons on social media and find that a few emoticons are strong and reliable signals of sentiment polarity which are the emoticons applied in this work [5]. Go, Bhayani and Huang used distant supervision for Twitter Sentiment Classification where training data consisting of Twitter messages with emoticons [3]. They conclude that machine learning algorithms (Naïve Bayes, Maximum Entropy and Support vector machine) have accuracy about 80% when trained with emoticon data. Santos and Gatti used Neural Networks to predict emoji usage from Twitter Data and found that Recurrent Neural Network with Long-Term Short Memory and Convolution Neural Network model outperform general machine learning approaches [6]. Altawaier and Tiun used emoticons data finding that Naïve Bayes outperformed other machine learning approaches for classifying sentiments in Arabic tweets [7]. Montoyo, MartíNez-Barco and Balahur classified English and Spanish tweets with a small but precise lexicon to capture which sentiment is expressed in Twitter data [8]. As for text classification, Marafino, Boscardin and Dudley invoke elastic-net-regularized classifiers to predict ICU mortality risk based on the free text of nursing notes and they also discovered text features which are most strongly associated with mortality [9].

However, there is a lack of research regarding Twitter analysis in Swedish language. Ludovici and Weegar studied Swedish articles sentiment analysis which cannot be applied undistinguishably in twitter messages because they have many unique attributes including the maximum length of the message, data availability, language model and domain twitter users' characteristics [10]. In terms of features in machine

learning for sentiment analysis, unigrams and bigrams, partof-speech (POS) tags, and adjectives are applied in a movie review dataset [11].

## III. THE PROPOSED METHOD

In order to investigate public opinions about newly opened IKEA store, local inhabitancies' tweets containing the word "IKEA" are crawled (namely "IKEA dataset"). Then the sentiment of individual tweets are computed for deep understanding of their opinion. The framework of sentiment analysis to deal with English and Swedish tweets is shown in Fig. 1.

For English tweets, lexicon-based approaches is used hereby in view of its power and credibility in many current applications. Thelwall, Buckley and Paltoglou use a dictionary of sentiment words with their associated sentiment polarity value to compute sentiment of each sentence by incorporating negation and intensification [12]. However, due to lack of such sentiment dictionary in Swedish, its sentiment can be predicted by machine learning-based methods. The training data is sentiment labelled Swedish tweets where each word in tweets are features fed into classifiers and associated sentiment ("positive" and "negative") are expected to be fitted accurately from alternative machine learning classification techniques. The performance of these classifiers is evaluated and the best one is invoked to predict sentiment of Swedish IKEA tweets.

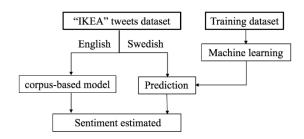


Fig. 1. The proposed model to predict IKEA sentiment.

# A. Data Collection

The above-mentioned two datasets, tweets containing "IKEA" and training dataset, are crawled separately as described in the following subsections.

- a query term "IKEA", location of twitter user, and a period of time. This dataset is filtered by constricting users' location from investigated IKEA-entry cities. Furthermore, the period of time under which dataset is collected should be long enough in order to catch how the trend of opinions varies. Twitter data over one month ranging from prior- to post- the entry date can be treated as a complete time-series data around IKEA opening dates. The examples of this dataset is presented in Table I and its summary is depicted in Table II.
- 2) Training dataset: As mentioned before, the challenge in Swedish sentiment analysis is the absence of Swedish sentiment dictionary and therefore the solution is to use machine learning to build a sentiment classifier. Thus, collecting

TABLE I SAMPLES OF IKEA TWEETS DATASET

| Place     | Time                   | Tweet   |  |  |  |
|-----------|------------------------|---|--|--|--|
| Sheffield | 2017-09-22<br>14:21:52 | @PeteWilcox1564 You have no idea how excited I am about the new Ikea!   |  |  |  |
| Sheffield | 2017-09-24<br>12:01:22 | I think @ChogRocker really enjoyed his first IKEA meal :') https://t.co/jaEACm9Thp                                |  |  |  |
| Sheffield | 2017-09-24<br>13:26:40 | I can confirm IKEA Sheffield restaurant (and frozen meatballs) accessible from the start                          |  |  |  |
| Halifax   | 2017-09-27<br>12:35:56 | Happy IKEA day! If you can't make it you can buy online and have it shipped to you                                |  |  |  |
| Halifax   | 2017-09-27<br>12:55:17 | I wish I was on my way to IKEA instead<br>of on my way to work  |  |  |  |
| Halifax   | 2017-09-27<br>12:01:38 | RT @hofoshore: Imagine falling in love with someone and then finding out they waited in that overnight ikea line. |  |  |  |
| Umeå      | 2016-02-26<br>09:31:42 | Härlig stämning när de första kunderna<br>kommer in i IKEA-varuhuset. Heja Umeå!                                  |  |  |  |

TABLE II CRAWLING THE IKEA TWEETS DATASET

|                | Umeå       | Halifax    | Sheffield  | Fishers    |
|----------------|------------|------------|------------|------------|
| Entry date     | 2016-02-26 | 2017-09-27 | 2017-09-28 | 2017-10-11 |
| Crawling start | 2016-02-11 | 2017-08-25 | 2017-08-25 | 2017-08-25 |
| Crawling end   | 2016-03-11 | 2017-10-31 | 2017-10-31 | 2017-10-31 |
| Tweets count   | 146        | 1882       | 2396       | 279        |

Swedish corpus with labelled sentiment as training data is necessary to achieve this goal. The training dataset consists of 15517 negative and 20000 positive Swedish tweets. Table III depicts a sample of this dataset. To avoid any bias prediction, identical tweets appeared in both IKEA datasets and training dataset are removed. The training dataset consists of Swedish tweets that contain two types of emoticons which is proposed by Wang et al. [13]:

- Positive emoticons: ":-)", ":)", "=)", ":D"
- Negative emoticons: ":( ", ":-(", "=(" and ";("

TABLE III
A SAMPLE OF THE TRAINING DATASETS

| Place Time                       |  | Tweet  |  |  |
|----------------------------------|--|--|--|--|
| Söderort 2017-10-28<br>12:45:01  |  | @kebabnekaise Drömde jag tog studenten idag.<br>Grät på riktigt när jag vaknade :( |  |  |
| Stockholm 2017-10-28<br>21:43:21 |  | @adamIynchs hmm :( drick kallt vatten?   |  |  |
| Stockholm 2017-09-14 09:53:58    |  | Dags för Noffans sista tur inför lördag :)   |  |  |

## B. Data Processing

The Twitter language model has many unique properties which can be considered as noise information and should be cleaned by the pre-processing step. Pre-processing the two datasets is achieved in the following manner:

 Remove the tweet user's names, which are marked with "@", URL, punctuation, numbers as well as the retweet mark "RT".

- In the establishment of corpus, English and Swedish stop words are discarded.
- The remaining words are transformed from upper case to lower case.
- Stemming, i.e. reduce words to their word stem

## C. Feature Extraction

1) Word Vector Representation: Strings or plain text cannot be processed in their raw form. Instead tokenization and word embedding are widely used techniques to quantify language processing problems. Each tweet is segmented into word terms by aid of 'space' character (tokenization). Then all word terms are used to build a feature vector to enable all raw tweets mapping words into a sparse matrix of terms (word embedding). One way to represent a tweet is by vectorizations, i.e. frequency of words. However, the main problem of term-frequency approach is that important informative words with low frequency are scaled down, but non-informative words are scaled up due to its high frequency. This problem is solved by Term Frequency – Inverse Document Frequency (TF-IDF) weight which consider not only the isolated word but also its relative importance to entire corpus. It is given by Eq. 1:

$$TF-IDF = \frac{n_t}{N} \cdot \log \frac{K}{k_n} \tag{1}$$

Where  $n_t$  is number of times a term t appears in a tweet and N is total number of terms in the tweet corpus. The first term of (1) stands for how frequently the term occurs in the document, while the second term is the offset by the frequency of the word in the corpus, which helps to adjust for the fact that some words appear more frequently in general, K is total number of documents and  $k_n$  is number of documents contain the term t.

The word vectors obtained from two datasets are processed separately. For "IKEA" tweets dataset, the word vector space generated from English tweets are used to compute sentiment by lexicon-based method, i.e. invoking semantic orientation (polarity value) of each word annotated in sentiment dictionary, while for Swedish ones, a machine learning approach which will be discussed in Section III-D is employed. For training dataset, those Swedish word vectors are used as candidate features in the machine learning approach.

2) Emoticons Replacement: To gain precise sentiment of Swedish corpus, state of the art showed that manually labelling the sentiments had been implemented [14]. However, this method is time consuming, expensive, and subjective in practice. Instead of this manual labelling, the current work applies an automatic labelling of the tweets. Since emoticons are commonly used in tweets, they can be invoked to label the corresponding word vector. To achieve this goal, a dummy variable as "positive" or "negative" is used to replace emoticons to represent the sentiment.

The training data set is created by mapping the word vector space to the corresponding emoticons replacement as the desired class. The resultant training dataset to fit Swedish sentiment prediction consists of 5263 terms (features) for 1-gram language model and 6597 terms for the combination of 1-gram and 2-gram model. Table IV depicts a sample of the whole training set.

TABLE IV
COUNT MATRIX OF TRAINING DATASET

| Tweet | $[X_1, X_2, \ldots, X_n]$ |     |     |       |  | Y      |            |
|-------|---------------------------|-----|-----|-------|--|--------|------------|
| ID    | Drömde                    | jag | tog | kallt |  | lördag | Sentiment  |
| 1     | 1                         | 1   | 1   | 0     |  | 0      | "negative" |
| 2     | 0                         | 0   | 0   | 1     |  | 0      | "negative" |
| 3     | 0                         | 0   | 0   | 0     |  | 1      | "positive" |
|       |                           |     |     |       |  |        | •          |

# D. Machine Learning Methods

Even though current machine learning approaches perform well in prediction of sentiment, interpretability cannot be achieved. Specifically, to discover and understand how specific text features are strongly associated with sentiment, therefore, interpretable statistical models are required. To solve that, logistic regression as a parametric model is considered but it is not good at dealing with the sparse matrix which is major problem in text mining.

The following classification techniques are trained and tested for the sake of gaining optimal sentiment prediction model for Swedish tweets: Generalized Linear Model with Elastic Net Penalty (EN), Logistic Regression (LR), Naïve Bayes (NB), Support Vector Machines (SVM), Neural Network (NN) and Random Forest (RF).

1) Elastic net model: Elastic-net penalized regression method is proposed due to following reasons. First, this statistical model can interpret and extract significant word terms in predicting sentiment; Second, it can solve variable selection and multi-collinearity problem in terms of sparsity of twitter data, especially, when the dimension of the input feature space is large and becomes even larger when bi-gram or higher-order n-gram features are extracted. Third, its application to twitter sentiment analysis has not been studied. Logistic regression with elastic-net penalties representing a compromise between the ridge and lasso penalties [15] The elastic net optimizes the objective function

$$\min_{\beta_0,\beta} \left[ \frac{1}{2N} \sum_{i=1}^{N} (y_i - \beta_0 - x_i^T \beta)^2 + \lambda P_{\alpha}(\beta) \right]$$
 (2)

Where

$$P_{\alpha}(\beta) = \sum_{j=1}^{p} \left[ \frac{1}{2} (1 - \alpha) \beta_j^2 + \alpha |\beta| \right]$$
 (3)

The compromise between the ridge-regression penalty ( $\alpha=0$ ) and lasso penalty ( $\alpha=1$ ) is the attractive advantage. This penalty is particularly useful in the  $p\gg N$  situation or situations where there are many correlated predictor variables. The parameter  $\beta$  is the coefficient of the word term and  $x_i$  is the feature vector. Two parameters  $\lambda$  and  $\alpha$  are tested in different settings.

- 2) Logistic Regression: In the current work, logistic regression with backward step wise elimination of the candidate word features was applied. The rule of removing feature is based on Akaike information criterion (AIC) difference between nested model.
- 3) Neural Network: A Feed-forward neural network consists of 2 hidden layers with 20 neurons is trained with the training dataset. The output layer has single neuron representing sentiment class with activation function of Hyperbolic tan. The weights were learned by using the error back propagation algorithm. Potential disadvantage of using neural network is it might be overzealous in sentiment analysis because it tries to fit the model to a numerable word terms even when there is no relationship. Moreover, such "black box" method cannot reveal the significant terms in sentiment classification.
- 4) Support Vector Machine: SVM train a hyperplane classifier  $f(x) = g(w^TX + b)$  with feature of words X, parameters w and b to optimally achieve maximum margin between two classes. It can be applied to both linearly and nonlinearly separable data. Text mining is more likely to be nonlinearly one, which can use kernel function to map data into a high-dimensional feature space. Kernel function and penalty parameter, C, which controls over-fitting are two key parameters in testing work and optimal one are presented in next section.
- 5) Random Forest: Random forest is ensemble strategy classifier which is bootstrapping algorithm with decision tree model. By replicating the process of subsample data and features, final prediction is build by voting for each iteration of fitting. The stop rule of bootstrapping takes place when low correlation is ensured between trees. In this analysis, the number of trees is fixed as 80 with all features tested while minimum of sample leaf are examined with several different values.
- 6) Naïve Bayes: Naïve Bayes classifier is based on Bayes rule which applied in a bag of words (corpus of Swedish twitter data). Individual tweet's sentiment class is decided by calculated conditional probability p("positive") is computed based on the frequency of each word in entire corpus.

The other models test in this work are very well described in the literature, and therefore, their training parameters are only presented, as given in Table V.

## E. Evaluation and Comparison of Classifiers

In each of above-mentioned classifiers, feature extractors adopt n-gram term which is a contiguous sequence of n items from a given sequence of text. Considering efficiency and accuracy, both 1-gram and a combination of 1-gram and 2-gram are examined as features in the current work.

To assess the performance of different methods,  $F_1$  – score and AUC are generally used as criteria which are calculated from confusion matrix.  $F_1 = \frac{2(\text{precision-recall})}{\text{precision+recall}}$ , where precision is the number of correct positive results divided by the number of all positive results, and recall is the number of correct positive results divided by the number of positive results that

TABLE V MODEL PARAMETERS

| Classifier              | Parameters  |  |  |
|-------------------------|---|--|--|
| Elastic Net             | $\lambda = \{0.1, 0.2, \dots, 1\}$<br>$\alpha = \{0, 0.005, 0.01, 0.1, 0.5, 1\}$  |  |  |
| Logistic Regression     | None  |  |  |
| Naïve Bayes             | None  |  |  |
| Support Vector Machines | $\label{eq:Kernel} \text{Kernel} = \{ \text{linear,polynomial, Radial basis function, sigmoid} \} \\ \text{C} = \{ 0.25, 0.5, 0.1 \}$ |  |  |
| Neural Network          | Activation function = Hyperbolic tan No. of layer = 2 No. of neurons = $\{5, 10, 20, 40\}$ Weight decay = $\{0, 10^{-4}, 10^{-1}\}$   |  |  |
| Random Forest           | No. of trees = 80<br>Min. sample leaf = {1,5,10}<br>Max. no. of features = all features   |  |  |

should have been returned.AUC is the area under the Receiver Operating Characteristic (ROC) curve which plots true positive rate versus false positive rate. Authors performed 5-fold cross-validation to estimate AUC for different models and  $F_1$ —score is the measurement of every model in testing dataset. When the data is imbalanced between positive and negative samples,  $F_1$ —score may change intensely while AUC curve was unaffected. Since the current twitter data is balanced, i.e. the number of positive and negative tweets are almost the same, the results from two criteria should be consistent. To avoid potential confliction of two criteria which indicates a bad classification, both standards are calculated and presented in Section 5.

#### IV. RESULT AND DISCUSSION

## A. English tweets sentiment analysis

For sentiment analysis in English tweets, the procedure of analysis is (1) Sentiment prediction of individual tweets (2) Frequently used positive and negative words. (3) decide which words are highly correlated with "IKEA" in tweets. Fig. 2(a), 2(b), 2(c) shows that there is a significant spike on the IKEA-entry for three investigated cities. Negative opinions about IKEA entry takes a very small proportion for all cities.

To identify which topics or words are likely to be referred when people talk about IKEA, authors calculated the associations of word "IKEA" with every word term in established document-term matrix. Fig. 2(d), 2(e), 2(f) shows "IKEA" highly correlated words whose Person correlation coefficient is larger than 0.1. Ignoring some function words, Halifax people are likely to talk about topics about online shopping, such as "website", "coupon" "amazon", Sheffield residents might have casual talk about life style like "village" "pots" and "plants". Since Fisher is a relative smaller city, their focus might be covered by new IKEA store, so words like "latest", "soon", "Saturday" and "opening day" indicates this event.

Fig. 2(g), 2(h), 2(i) presents frequently used positive and negative words for three cities. It shows that they share similar positive expressions like "good", "love", "happy" etc. But in terms of negative aspects, people have different focuses between three cities, Halifax people talk about "scam", "sneak",

"fake" and "hideous", while Sheffield mention "congestion", "chaotic" as a big city and Fishers' people talk about "drunk", "fear" and "dangerous" more about feelings. Inferring confidential concerned topics based on these words is not feasible, but these frequently used words are informative to know people's attention when they talk about IKEA.

## B. Swedish Tweets Sentiment Analysis

1) Model Comparison: The training dataset to fit Swedish sentiment prediction model is TF-IDF vectorized as 5263 terms (features) for 1-gram language model and 6597 terms for the combination of 1-gram and 2-gram model. In this work, 5-fold cross validation is implemented to assess each model's performance. Elastic Net outperform other approaches with relative higher AUC and  $F_1$ -score (Table VI). Because Elastic Net allows those relevant word or n-gram features associated with the outcome to be discovered far more easily, and hence has the potential to supersede existing "black-box" approaches [9].

TABLE VI
COMPARISON OF DIFFERENT APPROACHES IN PREDICTION SENTIMENT IN
SWEDISH TWEETS

|     | AU     | JC     | F-score |        |  |
|-----|--------|--------|---------|--------|--|
|     | 1-gram | 2-gram | 1-gram  | 2-gram |  |
| EN  | 0.8004 | 0.8171 | 0.7934  | 0.7930 |  |
| LR  | 0.7035 | 0.7068 | 0.7241  | 0.7261 |  |
| NB  | 0.7000 | 0.7015 | 0.7186  | 0.7208 |  |
| SVM | 0.7065 | 0.7204 | 0.7210  | 0.7217 |  |
| NN  | 0.7065 | 0.7204 | 0.6948  | 0.6998 |  |
| RF  | 0.7051 | 0.7115 | 0.7152  | 0.7146 |  |

2) Model Selection: To finalize optimal Elastic net model, further adjustment of parameters  $\lambda$  and  $\alpha$  are implemented. The experiment examines a series parameter  $\alpha$  given a fixed  $\lambda$  as shown in Fig. 3. The optimal setting are used for following prediction of Swedish "IKEA" tweets.

Some of typical extracted features and their coefficients in this approach are shown hereby,

- Negative words (tvingat: -4.13, smyga: -3.734544, storslam: -3.71, hotar: -3.66, tyvärr: -3.57, näck: -3.55, tillträde: -3.54, sjuktalsfrågan: -3.36, överfallen:-3.26, förort: -3.23, byggas:-3.13, studentfest: -3.07, weird:-3.06, strömlöst: -2.98)
- Positive words (tack: 3.11, SvenskDemokrater: 3.05, sparsamt: 2.74, nakenhet: 2.49, social: 2.26, slät: 2.23, organization: 2.23. absolut: 2.16, hehe: 2.10, bra: 1.95, beklagade: 1.93, glad: 1.72, Ekonomisk: 1.72, älskling: 1.69)
- 3) Sentiment Prediction: Based on the current optimal model, sentiment of every tweet in Umeå is estimated and sentiment scores for each tweet is their probability of positivity. Fig. 4 represent the sentiment of individual tweets during 2016-02-11 to 2016-03-11. Concentrated points are located in IKEA entry date 2016-02-26 and very little of them are less than 0.3 which indicates the IKEA entry attention and non-negative opinions in Umeå. To categorize them into three

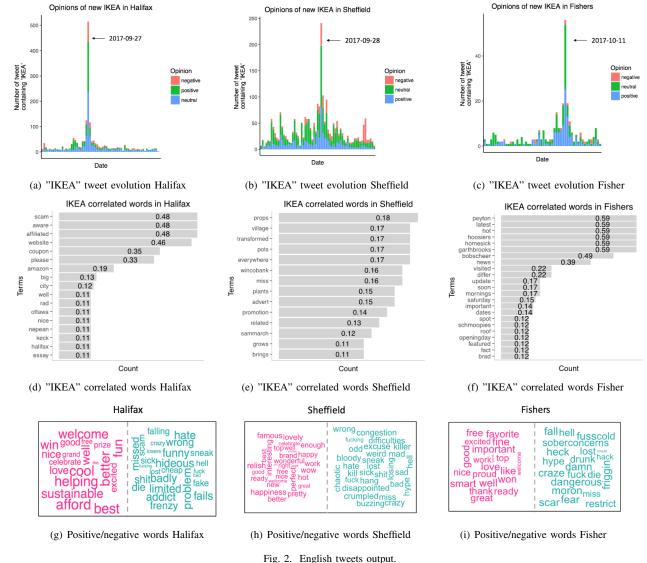


Fig. 2. English tweets output.

groups, all tweets are divided into three sets based on the probability: negative [0, 0.3], neutral [0.3,0.6], and positive [0.6,1]. The evolution of public opinions about IKEA is shown in Fig. 5. It shows that local people are welcome this new IKEA store and paid significant attention on IKEA entry date.

## V. CONCLUSION

This work studied people's opinion on new IKEA store based on Twitter data and a series of sentiment classification techniques. To compute sentiment polarity of tweets, lexicon based approach is used for English tweets, and machine learning methods for Swedish tweets. For English tweets, each word's associated sentiment polarity can be retrieved from existed English sentiment dictionary so every tweet sentence's sentiment can thus be computed by further incorporating negation and intensification.

For Swedish tweets, sentiment prediction requires to fit model due to lack of Swedish sentiment dictionary. The fitted sentiment prediction model build the relationship between word terms and the intrinsic sentiment so as to predict those "IKEA" tweets in Swedish.Comparison of different fitted models, Elastic-net approach outperformed other general classification models and provides the information about features in the current Swedish tweets corpus. To gain the optimal model, both 1-gram and the combination of 1-gram and 2-gram language modes are tested and parameters are adjusted in cross-validation.

Results show that all of the investigated cities payed significant attention on "IKEA" entry event which is drawn from the peak of tweets on their individual IKEA-entry date. Residents welcome new IKEA gladly since most of tweets relating "IKEA" are positive, very little negative tweets involved. However, the four cities have different negative topics due to the scale of city. To understand what do people talk about when they talk about IKEA, this work calculate word terms' correlation with "IKEA". It concludes that big city concerned

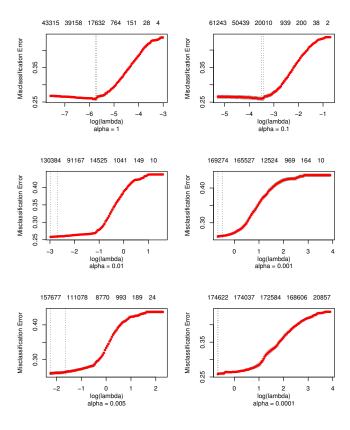


Fig. 3. Elastic-net model parameter curves for various  $\alpha$ .

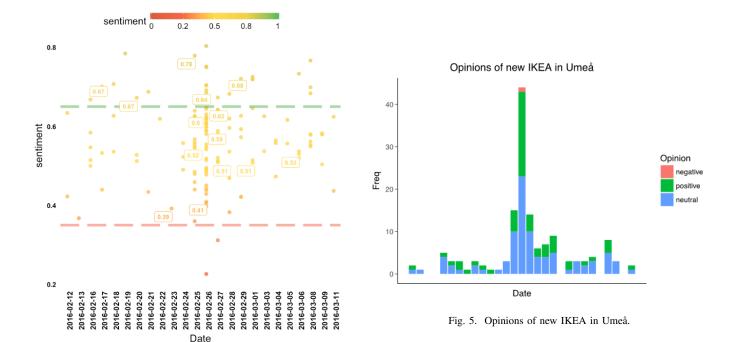


Fig. 4. Polarity of individual tweets in Umeå.

traffic and environment while small ones talk about life style and room renovation.

## REFERENCES

- S.-O. Daunfeldt, O. Mihaescu, H. Nilsson, and N. Rudholm, "What happens when ikea comes to town?" *Regional Studies*, vol. 51, no. 2, pp. 313–323, 2017.
- [2] J. C. F. Christopher D. Carroll and D. W. Wilcox, "Does consumer sentiment forecast household spending? if so, why?" *The American Economic Review*, vol. 84, no. 5, pp. 1397–1408, 1994.
- [3] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," CS224N Project Report, Stanford, vol. 1, no. 2009, p. 12, 2009.
- [4] M. Han, O. Mihaescu, Y. Li, and N. Rudholm, "Comparison and one-stop shopping after big-box retail entry: A spatial differencein-difference analysis," *Journal of Retailing and Consumer Services*, vol. 40, pp. 175–187, 2018.
- [5] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of twitter data," in *Proceedings of the workshop on languages in social media*. Association for Computational Linguistics, 2011, pp. 30–38.
- [6] C. N. Dos Santos and M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts." in COLING, 2014, pp. 69–78.
- [7] M. M. Altawaier and S. Tiun, "Comparison of machine learning approaches on arabic twitter sentiment analysis," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 6, no. 6, pp. 1067–1073, 2016.
- [8] A. Montoyo, P. MartíNez-Barco, and A. Balahur, "Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments," 2012.
- [9] B. J. Marafino, W. J. Boscardin, and R. A. Dudley, "Efficient and sparse feature selection for biomedical text classification via the elastic net: Application to icu risk stratification from nursing notes," *Journal of biomedical informatics*, vol. 54, pp. 114–120, 2015.
- [10] L. Michelle and R. Weegar, "A sentiment model for swedish with automatically created training data and handlers for language specific traits," 2016.
- [11] Y.-P. Ching, V. Y. Leong, M.-F. Lee, H.-T. Xu, D.-Y. Jin, and I. O.-L. Ng, "P21-activated protein kinase is overexpressed in hepatocellular carcinoma and enhances cancer metastasis involving c-jun nh2-terminal kinase activation and paxillin phosphorylation," *Cancer Research*, vol. 67, no. 8, pp. 3601–3608, 2007.
- [12] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment strength detection for the social web," *Journal of the Association for Information Science and Technology*, vol. 63, no. 1, pp. 163–173, 2012.
- [13] X. Wang, Y. Liu, C. Sun, B. Wang, and X. Wang, "Predicting polarities of tweets by composing word embeddings with long short-term memory." in ACL (1), 2015, pp. 1343–1353.
- [14] J. Wiebe, T. Wilson, and C. Cardie, "Annotating expressions of opinions and emotions in language," *Language resources and evaluation*, vol. 39, no. 2, pp. 165–210, 2005.
- [15] J. Friedman, T. Hastie, and R. Tibshirani, "glmnet: Lasso and elastic-net regularized generalized linear models," *R package version*, vol. 1, no. 4, 2009.