Deep Learning-Based Sentiment Analysis of Customer Reviews on Hotels Through Tweets

M Srisankar and K P Lochanambal

January 28, 2023

DEEP LEARNING-BASED SENTIMENT ANALYSIS OF CUSTOMER REVIEWS ON HOTELS THROUGH TWEETS

M.SRISANKAR[1]
Research Scholar,
Department of computer science
Government Arts college,udumalpet
Tamilnadu, India
Shankar276@gmail.com

Dr.K.P.LOCHANAMBAL[2]
Assistant Professor,
Department of computer science
Government Arts college,udumalpet
Tamilnadu, India
kplmca@rediffmail.com

ABSTRACT

Tweets are continually produced and short. Sentiments play a vital role in comprehending personal feelings. Multitudes show their opinions on any subject or object on social media. The public's sentiment is split into different categories namely, positive, negative and neutral. In this paper, data is acquired from hotel reviews on Twitter using Python's Tweepy library and pre-processed using Python scripts. Inconsistent data or noise, retweets, tags, URLs along with hashtag symbols and duplicate entries are removed. The tweets are upsampled and split by using scikit-learn in Python. The textual data is converted into vectors using Keras Tokenizer in Python. Bi-Sense Emoji Embedding (BSEE) is used for performing SA. The sentiments are categorized using Support Vector Machine (SVM) and Random forest (RF), and performance is compared with BSEE based on Accuracy, Recall, F-measure, Precision and Time period. It is seen that the proposed classifier offers better results.

Keywords

Tweets, Sentiment Analysis (SA), Support Vector Machine (SVM), Random forest (RF), Bi-Sense Emoji Embedding (BSEE)

I. INTRODUCTION

Sentiment detection and classification are currently the emphasis of social analytics on the online platform. Nowadays, Twitter has turned into a significant tool for procurement of peoples' feelings. Sentiment Analysis (SA) of tweets becomes a captivating exploration in the area of Artificial Intelligence (AI) [1]. SA is widely utilized in several fields like marketing, education, e-commerce, so on. The goal of SA is to analyse the online reviews and determine the scores for sentiments. Categorizing sentiments primarily depends on labelled data. English has more labelled texts in contrast to other languages [2].

SA is necessary for information alignment in social media as user engagement rates are significantly higher on Twitter [3]. Twitter has evolved into an important tool to collect information about people's emotions. Emotion recognition of tweets turns into a fascinating AI experiment.

Several methods are formulated for evaluating the SA of Twitter users, yet some improvisation is needed to improve accuracy and execution time. An efficient algorithm is needed to extract and classify tweet sentiments in Twitter [4]. A dataset for SA is a group of tweets about customer perceptions. Every tweet is given a score demonstrating whether it is positive, negative or neutral. For each tweet, SA techniques determine the sentiment. Using the sentiment score in the dataset, the calculated sentiment is then assessed as right or wrong. Accuracy identifies tweets that are accurately evaluated by algorithm [5]. Micro-blogging sites like Twitter involve infinite source of diverse types of information. People post opinions on varied topics, contend about present challenges, criticize as well as express emotions. Day-by-day, the number of users of platforms in addition to social networks gradually increase [6-8]. Data got from sources are employed in opinion mining, and sentiment along with emotion analysis. Sentiments, opinions besides emotions are included in opinion mining and SA. Data on social media are noisy involving errors in spelling, punctuation and grammar.Emoji is the general form of communication [9]. They are visual representations of emotions, symbols or objects. Ideograms as well as smileys are used in electronic messages in addition to web pages. They are of varying kinds comprising of facial expressions, weather, universal objects, places, animals and birds [10]. They look like emoticons, but Emojis are pictures and not typo-graphics. Emoticons are typographic exhibits of emotion, while Emojis are images that express facial expressions, emotions, activities, objects and animals [11-12]. Emojis and Emoticons are used in present applications for communication. They include text messaging and applications involved in social networking like Instagram, Facebook, Twitter and Snapchat [13]. As SA plays a major role in social media, in this paper, Bi-Sense Emoji

Embedding (BSEE) is proposed for performing SA of hotel reviews. Data is gathered using Python's Tweepy library and pre-processed using Python scripts [14]. Noise, duplicate, tags, URLs along with hashtag symbols are removed. Tweets are upsampled and split using scikit-learn in Python. It is seen that the proposed classification model offers improved results.

2. RELATED WORK

Salampasis et al (2014) [15] have used Twitter and also a system for gathering user produced data for a particular duration. An algorithm based on unsupervised and lexicon-based approach is used to perform SA to automatically handle the conveyed emotion of microblogs using ternary classification. It offers better performance in a variety of environments.Howells & Ertugan (2017) [16] have focussed on developing a model which analyses content of microblog and is capable of analysing customer feedback or perception. The model forms basis of application which an organisation uses for marketers for timely as well as strong illustration of customers' insights on any product as well as service. Initially, a Fuzzy Logic (FL)-based model is designed to convert linguistic variables into membership functions which define variables in a fuzzy manner. Every membership function signifies linguistic variable and forms Fuzzy Set (FS) which is related to specific factor. Once FS is formed, the ensuing step is defining if-then rules which describe FSs interaction. Operators are included to form connections amid FSs namely max (and) and min (or). Once connections are formed, propositions are combined into final FS. FS releases fuzzy number and is formed into crisp value using diverse methods.Al-Otaibi et al (2018) [17] have used Twitter data for gaining understanding from public opinion in data. SVM is employed for classifying sentiments obtained from tweets and unigram is applied for extracting features. Decision-making in Organizations as well as individuals by offering investigation of information about products, customer opinions and product reviews found in social media are handled. It aids in providing details about competitors and assess products available in market. This saves customer's time enabling them to know about the products and stimulate production of many products by knowing the customer viewpoints and other competitors' products. Textual data of products as well as services are considered in Twitter and features like age, education, followers, gender are considered. Experiments are carried out for large training datasets and it is seen that the algorithm offers improved accuracy. Smetanin & Komarov (2019) [18] offers a method for SA of Russian product reviews using Convolutional Neural Networks (CNNs). Vectors pre-trained with Word2Vec are given as inputs to NNs. This scheme does not use any hand-crafted attributes or lexicons. Training dataset is formed from reviews on high rated goods from main e-commerce sites, where user-rated scores are used as labels for classes. The proposed scheme offers improved F-measure. The training dataset as well as Word Embeddings (WEs) can be accessed by research community.Moussa (2019) [19] have introduced an emoji-based metric for observing emotions of consumers available on social media related to brands. Initial findings from study tests the metric based on 720 consumer tweets of 18 principal universal brands demonstrating 6 product groups or markets. This metric can be easily implemented and interpreted by every researcher as well as manager. Researchers and managers compute it using information that can be freely got by mining social media account. Huge amount of information seen in emoji-based User-Generated Content (UGC) regardless of language can be analysed. It is highly useful for researchers who perform multi-cultural study of emotions of consumers as well as managers of universal brands who are interested in observing consumers' feelings.Behera et al (2021) [20] have proposed a hybrid approach involving CNN as well as Long Short Term Memory (LSTM) for classification of sentiments in reviews related to numerous domains. Deep CNNs are highly useful in selection of local features, while Recurrent Neural Networks (RNNs) yield improved outcomes in sequential analysis of longer text. The propounded Co-LSTM model is highly flexible in observing huge volume of social data, managing scalability, and unlike conventional Machine Learning (ML) schemes, it is not restricted to any specific domain. The experiment is applied on 4 review datasets from varied domains to train model that can deal with every dependency which arises in post. The outcomes show that propounded ensemble model outdoes other ML schemes based on Accuracy and other factors.

3. CLASSIFICATION MODELS

Classification models incorporating Random Forest (RF) and SVM are discussed in detail. The proposed Bi-Sense Emoji Embedding (BSEE) for performing SA is detailed in the next section.

3.1 RANDOM FOREST (RF)

RF classifier, a supervised learning algorithm may be used for regression besides classification. It is a popular ML algorithm as it is highly flexible and easy to implement. It includes several Decision Trees (DTs) like a forest. Randomness is used for improving Accuracy and fight overfitting that may be an issue. DTs are created depending on arbitrary choice of data samples and obtain predictions from each tree. Best feasible solutions are obtained through votes. It is used in fraud detection, disease prediction etc., Let the amount of features in dataset be 'm'. RF chooses 'k' features arbitrarily, where k < m. The algorithm determines root node among 'k' features by choosing the node with increased information gain. The node is split into children and the process is repeated 'n' times. The results of DTs are combined to get the final result. This ensemble algorithm produces separate DTs by selecting attributes. The votes from every tree are taken and most prevalent class is taken as the final outcome. In a regression problem, the average of tree outputs is computed which is the end outcome.

3.2 SUPPORT VECTOR MACHINE (SVM)

SVM is also supervised learning algorithm that is used for classification in addition to regression problems. It is employed for classification in ML. It is used for generating decision boundary which can classify n-dimensional space into categories such that data point is put in right category. Hyperplane is the decision boundary. It selects extreme points or vectors which aid in producing hyperplane. The extreme cases are known as SVs. SVM determines a hyperplane which generates a boundary amid different kinds of data. Every data item in dataset is plotted in n-Dimensional space for 'n' features in data. Ideal hyperplanes are used to separate data. SVM performs binary classification. Nevertheless, several schemes that can be used for multi-class problems are available. By applying SVM to multi-class problems, a binary classifier can be created for every class of data.

4. PROPOSED METHODOLOGY

The propounded method involves the ensuing steps.

Data acquisition: Python's Tweepy library is employed for collecting tweets related to customer reviews on hotels in CSV format during the period 01 April 2022 to 30 May 2022. The tweets include features like Tweet ID, text and date and time when it was posted. An investigative analysis shows occurrence of tweets of diverse polarities in dataset.

Data Pre-processing: Pre-processing is performed using Python scripts. Raw data from Twitter is pre-processed to remove inconsistent data or noise before feeding to model. The dimensionality of data reduces which improves performance in SA. By using Python's Emoji module, Emojis are transformed to Common Locale Data Repository Project (CLDR) short names. Text is transformed to lowercase, followed by elimination of retweets, URLs, tags, and hashtag symbols. Incomplete words are rewritten. Duplicate entries are removed from text.

Data Upsampling along with Splitting: As the data is not balanced, most of the tweets fall under positive class. Neutral as well as negative tweets are upsampled in addition to split by using scikit-learn in Python. Hence, they are resampled and split into training (80%), validation and testing (20%).

Tokenisation, Padding and WEs: The textual data is converted to vectors called WEs using Keras Tokenizer in Python. The length of the sentence is fixed to 150. Hence, WEs of sentences with lesser lengths are padded with zeros.

4.1 Bi-sense Embedding

Word-Guide Attention-based LSTM (WGA-LSTM) as well as Multi-Level Attention-based LSTM (MLA-LSTM) is proposed. Bi-Sense Emoji Embedding (BSEE) is used for performing SA. The proposed system includes the ensuing steps:

Initialization of BEE

Generation of Senti-Emoji Embedding (SEE) depending on self-chosen attention

Classification of sentiments through attention-dependent LSTM networksCurrently, WEs like word2vec and fasttext play a major role [21, 22]. Fasttext aids initialisation of Emoji embeddings by considering emoji as word along with WEs. In contrast to traditional schemes where every emoji corresponds to an embedding vector which is Word-Emoji Embedding (WEE), emoji is embedded into 2 different vectors (BSEE). Two separate tokens are assigned to every emoji wherein, one is for a specific emoji employed in positive sentiments, and other is used in negative sentiments. Text sentiments are initialized using Vader [23]. Fasttext training helps in embedding tokens into diverse vectors. This aids in obtaining positive in addition to negative-sense embeddings for each emoji.

Word2vec is based on skip-gram model that is used for maximising log probability determined by totalling probabilities of current word incidences for a set of adjoining words. In case of fasttext model, the problem is formulated as binary classification for predicting incidence of every Context Word (CW) with negative samples being arbitrarily chosen from absent CWs. For a word structure $\{w_1, w_2 \ldots w_T\}$ given as input, and CW set (w_{c_t}) and collection of negative samples (w_{n_t}) of present word (w_t), objective function depends on Binary Logistic Loss (BLL).

$$\sum_{t=1}^{T}\left[\sum_{w_c \in w_{c_t}} \underset{\cdot}{\bar{L}}(S(w_t, w_c)) + \sum_{w_c \in w_{n_t}} \underset{\cdot}{\bar{L}}(S(w_t, w_n))\right] \tag{1}$$

where,

$\underset{\cdot}{\bar{L}}$ (s(·, ·)) - Score function's logistic loss

s(· , ·) - Determined by summing scalar products amid n-gram embeddings of present word and CW embedding that varies from word2vec wherein score is scalar product amid embeddings

Fasttext is chosen as it is computationally efficient. The models offer improved performance.

4.2 Word-Guide Attention-based LSTM (WGA-LSTM)

LSTM is widely used for encoding textual contents. Fundamental encoder model includes text embedding, LSTM and Fully-Connected (FC) layers for text classification depending on encoded feature. Functions in LSTM for time step (t) are shown below.

$$I_t = \sigma(S_i.X_t + T_i.H_{t-1} + B_i) \tag{2}$$

$$F_t = \sigma(S_f.X_t + T_f.H_{t-1} + B_f) \tag{3}$$

$$O_t = \sigma(S_o.X_t + T_o.H_{t-1} + B_o) \tag{4}$$

$$G_t = \tanh(S_c.X_t + T_c.H_{t-1} + B_c) \tag{5}$$

$$C_t = F_t \odot C_{t-1} + I_t \odot G_t \quad (6)$$

$$H_t = O_t \odot \tanh(C_t) \quad (7)$$

where,

H, H_{t-1} - Present and former hidden states

X_t - Present LSTM input

S and T - Weight matrices

σ - Sigmoid function

To make use of the benefits of BSEE, the input layer is modified into LSTM units. SEE is got as weighted average of BSEE depending on self-chosen attention scheme. Let $e_t^i, i \in (1, m)$ signifies 'i^{th}' sense EE (e_t) (m=2 in BSEE) and fatt $(., w_t)$ represents attention function on present WE. The attention weight (ρ_i) and SEE vector (V_t) are shown below.

$$u_t^i = f_{att}(e_t^i, W_t) \quad (8)$$

$$\rho_t^i = \frac{\exp(u_t, i)}{\sum_{i=1}^{n} \exp(u_t, i)} \quad (9)$$

$$v_t = \sum_{i=1}^{m} (\rho_t^i, e_t^i) \quad (10)$$

FC layer along with ReLU activation is used as attention function, and attention vector (V_t) is integrated with WE to LSTM. 'X_t' in Equations (2) to (7) becomes $[W_t, V_t]$. The LSTM unit's output is fed into FC layer with sigmoid activation to return sentiment. Binary Cross-Entropy (BCE) loss is employed as objection function with 'n' samples. The inspiration behind the proposed model is that every CW guides 'ρ_i' to impose model to choose embedding sense it attends. This model is denoted as WGA-LSTM with BSEE (WGA-BSEE-LSTM).

$$L(\theta) = -\frac{1}{n}\sum_{i=1}^{n} X_i . \log(P_i) + (1 - X_i). \log(1 - P_i) \quad (11)$$

4.3 Multi-level Attention-based LSTM (MLA-LSTM)

The attention scheme is formulated in a different way, where 'V_t' indicates how image information (emoji) is dispersed in CWs as seen in [24, 25]. The adapted SEE vector (V) is at tweet rather than word-level in Equations (8)-(10) by substituting 'W_t' with final state vector (H) received as output from final LSTM unit.

$$\rho_i = \frac{\exp(f_{att}(e_i, h))}{\sum_{i=1}^{m} \exp(f_{att}(e_i, h))} \quad (12)$$

$$V' = \sum_{i=1}^{m} (\rho_i, e_i) \quad (13)$$

The derived SEE (V') is used for determining an added layer of attention. For sequence $\{W_1, W_2, \dots W_T\}$, attention weight $\rho_t' \in (1, T)$ on SEE is expressed as follows:

$$\rho_t' = \frac{\exp(F_{att}(W_t, V'))}{\sum_{i=1}^{m} \exp(F_{att}(W_t, V'))} \quad (14)$$

In the MLA-LSTM with BSEE (MLA-BSEE-LSTM) model, new input (u_t) for every LSTM unit is constructed by integrating original WE and 'V_t' in Equation (15) to allocate SE information to every step. BCE is selected as loss function with comparable network configuration with WGA-BSEE-LSTM.

$$u = [W_t, \rho_t', V'] \quad (15)$$

5. RESULTS AND DISCUSSION

Data acquired from hotel reviews available on Twitter by using Python's Tweepy library. Data pre-processing is performed using Python scripts to remove inconsistent data, noise, retweets, tags, URLs, hashtag symbols and identical entries. Upsampling and splitting of samples is carried out by using scikit-learn. The textual data is converted into vectors using Keras Tokenizer. SA is performed using BSEE. Sentiments are categorized using SVM and RF. It is seen that the proposed scheme outperforms the existing schemes in terms of Accuracy, Recall, F-measure, Precision and Time period. BSEE offers 5.7% and 1.5% better Accuracy in contrast to SVM and RF respectively (Figure 1).

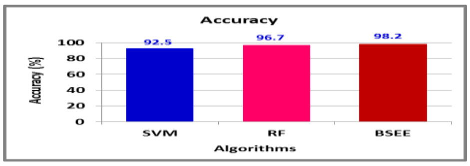

Figure1: Accuracy

BSEE offers 3.7% and 2.2% better Recall when compared to SVM and RF respectively (Figure 2).

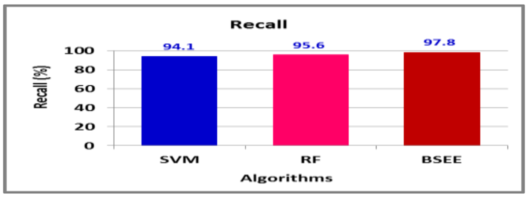

Figure2: Recall

BSEE offers 2.6% and 1.3% better F-Measure in contrast to SVM and RF respectively (Figure 3).

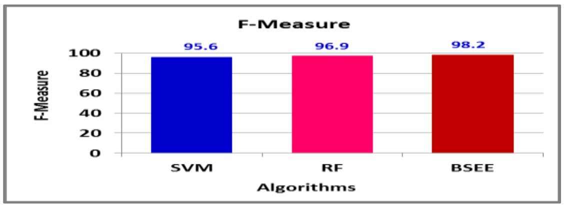

Figure3: F-Measure

BSEE offers 3.8% and 2.1% better Precision when compared to SVM and RF respectively (Figure 4).

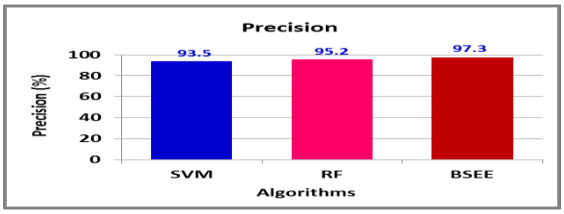

Figure4: Precision

BSEE offers 1.66 and 1.26 times reduced Time period in contrast to SVM and RF respectively (Figure 5).

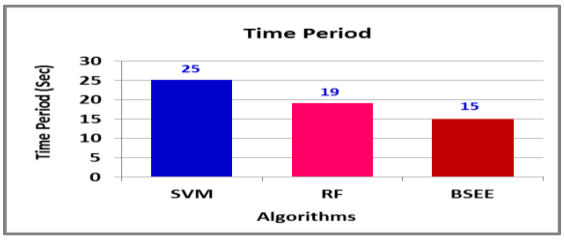

Figure5: Time Period

6. CONCLUSION

In this paper, Bi-Sense Emoji Embedding (BSEE) is used for performing SA. Sentiments are classified using RF and SVM and performance is compared with BSEE in terms of Accuracy, Recall, Precision, Time period and F-measure. Python's Tweepy library is used in data acquisition and Python scripts are used in pre-processing. Inconsistent data, noise, URLs, retweets, tags, hashtag symbols and duplicate entries are removed. Scikit-learn in Python is employed in upsampling and splitting of tweets. Keras Tokenizer is used for converting textual data into vectors. It is seen that the proposed classifier outperforms other models.

7. REFERENCES

1. Giachanou, A., & Crestani, F. (2016). Like it or not: A survey of twitter sentiment analysis methods. ACM Computing Surveys (CSUR), 49(2), 1-41.
2. Sarlan, A., Nadam, C., & Basri, S. (2014, November). Twitter sentiment analysis. In Proceedings of the 6th International conference on Information Technology and Multimedia (pp. 212-216). IEEE.
3. Zimbra, D., Abbasi, A., Zeng, D., & Chen, H. (2018). The state-of-the-art in Twitter sentiment analysis: A review and benchmark evaluation. ACM Transactions on Management Information Systems (TMIS), 9(2), 1-29.
4. Sahayak, V., Shete, V., & Pathan, A. (2015). Sentiment analysis on twitter data. International Journal of Innovative Research in Advanced Engineering (IJIRAE), 2(1), 178-183.
5. Neethu, M. S., & Rajasree, R. (2013, July). Sentiment analysis in twitter using machine learning techniques. In 2013 fourth international conference on computing, communications and networking technologies (ICCCNT) (pp. 1-5). IEEE.
6. Sebastian, T. (2012). Sentiment Analysis for Twitter (Doctoral dissertation).
7. Martínez-Cámara, E., Martín-Valdivia, M. T., Urena-López, L. A., & Montejo-Ráez, A. R. (2014). Sentiment analysis in Twitter. Natural language engineering, 20(1), 1-28.
8. Ramadhani, A. M., & Goo, H. S. (2017, August). Twitter sentiment analysis using deep learning methods. In 2017 7th International annual engineering seminar (InAES) (pp. 1-4). IEEE.
9. Shiha, M., & Ayvaz, S. (2017). The effects of emoji in sentiment analysis. Int. J. Comput. Electr. Eng.(IJCEE.), 9(1), 360-369.
10. Fernández-Gavilanes, M., Costa-Montenegro, E., García-Méndez, S., González-Castaño, F. J., & Juncal-Martínez, J. (2021). Evaluation of online emoji description resources for sentiment analysis purposes. Expert Systems with Applications, 184, 115279.
11. Guibon, G., Ochs, M., & Bellot, P. (2016, June). From emojis to sentiment analysis. In WACAI 2016.
12. Lou, Y., Zhang, Y., Li, F., Qian, T., & Ji, D. (2020). Emoji-based sentiment analysis using attention networks. ACM Transactions on asian and low-resource language information processing (TALLIP), 19(5), 1-13.
13. Lou, Y., Zhang, Y., Li, F., Qian, T., & Ji, D. (2020). Emoji-based sentiment analysis using attention networks. ACM Transactions on asian and low-resource language information processing (TALLIP), 19(5), 1-13.
14. Grover, V. (2022). Exploiting emojis in sentiment analysis: A survey. Journal of The Institution of Engineers (India): Series B, 103(1), 259-272.
15. Salampasis, M., Paltoglou, G., & Giachanou, A. (2014). Using Social Media for Continuous Monitoring and Mining of Consumer Behaviour. Int. J. Electron. Bus., 11(1), 85-96.
16. Howells, K., & Ertugan, A. (2017). Applying fuzzy logic for sentiment analysis of social media network data in marketing. Procedia computer science, 120, 664-670.
17. Al-Otaibi, S., Alnassar, A., Alshahrani, A., Al-Mubarak, A., Albugami, S., Almutiri, N., & Albugami, A. (2018). Customer satisfaction measurement using sentiment analysis. International Journal of Advanced Computer Science and Applications, 9(2).
18. Smetanin, S., & Komarov, M. (2019, July). Sentiment analysis of product reviews in Russian using convolutional neural networks. In 21st IEEE conference on business informatics (CBI), Vol. 1, pp. 482-486.
19. Moussa, S. (2019). An emoji-based metric for monitoring consumers' emotions toward brands on social media. Marketing Intelligence & Planning.

20. Behera, R. K., Jena, M., Rath, S. K., & Misra, S. (2021). Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data. Information Processing & Management, 58(1), 102435.
21. Tomas, M., Kai, C., Greg, C., & Jeffrey, D. (2013). Efficient Estimation of Word Representations in Vector Space. CoRR abs/1301.3781.
22. Piotr, B., Edouard, G., Armand, J., & Tomas, M. (2017). Enriching Word Vectors with Subword Information. TACL 5 (2017), 135-146.
23. Clayton, J.H. & Eric, G. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. In 8th International Conference on Weblogs and Social Media (ICWSM 2014).
24. Tianlang, C., Yuxiao, C., Han, G., & Jiebo, L. (2018). When E-commerce Meets Social Media: Identifying Business on WeChat Moment Using Bilateral-Attention LSTM. 27th International Conference on World Wide Web Companion, pp. 343-350.
25. Yequan, W., Minlie, H., Li, Z., Xiaoyan, Z. (2016). Attention-based lstm for aspect-level sentiment classification. Conference on Empirical Methods in Natural Language Processing, 606-615.