The Impacts of Primacy/Recency Effects on Item Review Sentiment Analysis

Besnik Gjergjizi¹, Thi Ngoc Trang Tran^{2,*} and Alexander Felfernig²

¹Graz University of Technology, Rechbauerstraße 12, 8010 Graz, Austria

²Institute of Software Technology, Graz University of Technology, Inffeldgasse 16b/II, 8010 Graz, Austria

Abstract

Primacy/recency effects, also known as *serial position effects*, are cognitive biases triggered when items are presented in the form of a list. Affected by these effects, users tend to recall items shown in the beginning or the end of the list more often than those in the middle. Although primacy/recency effects have been extensively analyzed within the field of psychology, they are not studied well in the context of sentiment analysis. In the literature, there are still missing studies that provide an in-depth analysis of the influences of these effects on machine learning algorithms for item review sentiment analysis. This paper bridges this gap by estimating the impacts of primacy/recency effects on sentiment analysis classifiers. We propose a *primacy/recency effects-aware* neural network of Bidirectional Long Short-Term Memory (so-called *PriRec-BiLSTM*) and compare the performance of this approach with the original neural network (*BiLSTM*). To sufficiently evaluate the classification accuracy of the proposed approach, we ran our approach in five datasets in different item domains, such as movies, Amazon smartphones, industry and science, and airlines Tweets. The experimental results show that considering primacy/recency effects helps increase sentiment classification accuracy.

Keywords

Machine Learning Algorithms, Decision Biases, Primacy/Recency Effects, Serial Position Effects, Item Review, Sentiment Analysis, Sentiment Classification, Neural Networks

1. Introduction

The rapid growth of the Internet and technology has led to the evolution of e-commerce and social networks. Through these online platforms, plenty of the daily-life activities of users have been done, which therefore causes a significant increase in the amount of shared information on the Internet. Various types of information can be collected, in which information about product/item reviews has been found to be helpful for user decision making and therefore have positive impacts on customer engagement and purchase intentions [1, 2]. In this context, *sentiment analysis* (known as one area of *Natural Language Processing* [3]) has emerged as a research area that helps classify user reviews into positive, negative, or neutral classes/salience.

IntRS'22: Joint Workshop on Interfaces and Human Decision Making for Recommender Systems, September 22, 2022, Seattle, US (hybrid event)

^{*}Corresponding author.

[🛆] besnik.gjergjizi@student.tugraz.at (B. Gjergjizi); ttrang@ist.tugraz.at (T. N. T. Tran);

alexander.felfernig@ist.tugraz.at (A. Felfernig)

D 0000-0002-3550-8352 (T. N. T. Tran)

^{© 0 2022} Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

On the other hand, *cognitive biases* are systematic errors in thinking that occur when people are processing and interpreting information in daily life and, therefore, affect their decision-making processes. *Primacy/recency effects*, also known as *serial position effects*, are cognitive biases triggered when items are presented in the form of a list. Unconsciously influenced by these effects, users tend to recall items shown at the beginning or the end of the list more often than those in the middle [4, 5]. Although these biases have been analyzed within the field of psychology and customer decision-making [6, 7, 8, 9], the literature is still missing studies that provide in-depth analyses of the influence of these psychological effects on item review sentiment analysis. To the best of our knowledge, no studies take into account primacy and recency effects in sentiment analysis.

The classification of sentiments can be achieved based on three approaches: *lexicon-based approach, machine learning approach*, and *hybrid approach* [10, 11, 12, 13, 14]. The lexicon-based approach is based on the construction of a lexicon or a dictionary containing words to be evaluated. A piece of text is a bag-of-words [13] and the text is evaluated according to the values of the words found in the dictionary or lexicon. Machine learning approaches are usually based on supervised learning, which does not consider primacy/recency effects. Hybrid approaches combining machine learning and lexicon approaches [12] in some cases would not perform well. In this paper, we focus on machine learning approaches. Different from previous studies in the same research line, we consider primacy/recency effects when designing our machine learning model to improve the accuracy of classification algorithms. The contribution of this work is to analyze the impacts of primacy/recency effects on a neural network. We run our approach in various datasets to sufficiently evaluate our algorithm's accuracy. The experimental results show that the primacy/recency-aware neural network achieves a higher classification accuracy compared to the original one.

The remainder of the paper is structured as follows. *Section 2* presents an overview of related work. *Section 3* summarizes the methods used for our analysis and presents our neural network model. The experimental results regarding the accuracy of our proposed neural network approach are provided in *Section 4*. Finally, in *Section 5*, we conclude the paper and discuss open issues for future work.

2. Related work

2.1. Primacy/Recency Effects

Primacy/recency effects are a well-established concept within the field of psychology, which was first mentioned in 1878 [15] and further examined in later research [16, 17, 18, 19]. More recently, these biases have been analyzed in the context of *recommender systems*¹, where users tend to focus on evaluating items shown *at the beginning* and *at the end* of a list [6, 7, 21, 22]. Felfernig et al. [6, 7] show that items at these positions are more likely to be evaluated than others. Primacy/recency effects can change the selection behavior of users when interacting with recommender systems. For instance, in personnel decision-making, Highhouse and Gallo [21]

¹Recommender systems are efficient tools that help to cope with information overload issues in many application domains [20].

find that candidates interviewed at the end of a recruitment process have a higher probability of being selected. Tran et al. [9] investigate the influence of these effects when the same group of users has to continuously make a sequence of decisions in different item domains. The authors show that the order of decision tasks causes changes in the decision-making strategies of group members. However, in the context of *multi-attribute items*, Tran et al. [23] show that the selection of a recommended item from a list of candidate items is immune to primacy/recency effects. These effects can also be exploited to increase the user interaction with the system. For instance, in an e-learning system, questions that have been answered wrongly by the students will be shown at the beginning or the end of the question list to increase the probability of being accessed by the students [24].

Based on this idea, we assume that primacy/recency effects can also be utilized in the context of item review sentiment analysis, where user reviews are tailored by a list of arguments. The arguments in the beginning and the end of a review are assumed to strongly reflect the overall evaluation of a review. To the best of our knowledge, there are no studies analyzing the impacts of these effects in the context of item review sentiment analysis.

2.2. Deep Learning Approaches for Sentiment Analysis

There exist plenty of traditional machine learning approaches that have been used for sentiment analysis, such as support vector machines, Naive Bayes classifier, logistic regression, and multi-layer perceptron [25, 26, 27, 28]. However, these approaches require extensive domain knowledge of how users use social media platforms to express their moods and emotions [29]. To deal with the mentioned limitation, deep learning methods have been proposed, which allow to automatically extract the features without the dependence on extensive manual feature engineering [29]. Besides, some related work has proven that deep learning approaches achieve the best results in sentiment analysis in different item formats, such as text, images, sound, and video [30]. There are three conventional deep learning techniques that can be applied to sentiment analysis, Recursive Neural Network (RNN) [31], Convolutional Neural Network (CNN) [32, 33], and Long Short-Term Memory (LSTM) neural network [34, 35]. RNN maps phrases through word embeddings and a parse tree. Afterward, vectors for higher nodes in the tree are computed, and a tensor-based composition function for all nodes is used [31]. This approach depends on the syntactic structure of the text as input which needs to be generated beforehand [29]. In contrast, CNN does not belong to the additional inputs. Its input is rather extracted directly from the reviews, which can be in the form of images, assigned weights, and biases. Although CNN applies to sentiment analysis [36], it is hard to adapt the weights of the neural network at the user level. In other words, there is no way to apply domain knowledge (e.g., primacy/recency effects) via hyperparameters. Finally, LSTM neural network uses word embedding as input and generates hidden states sequentially where a given hidden state is dependent on the previous one. This allows the network to model long-range dependencies [29].

In this work, we select the *LSTM* (in particular, *Bidirectional LSTM - BiLSTM*) approach to analyze the impacts of primacy/recency effects on sentiment analysis. The reason for this selection is that *LSTM* has been widely used in the field of *Natural Language Processing* and proven to be effectively applied to sentiment analysis [34, 37]. Our idea is to propose an extended version of *BiLSTM* (so-called *PriRec-BiLSTM*) that takes into primary/recency effects and to

estimate the classification accuracy in comparison with *BiLSTM* (the neural network without primacy/recency effects).

3. Method and Neural Network Models

3.1. Method

As mentioned earlier, our general idea is to consider primary/recency effects in the sentiment classification. To address this, given an entry (review) *E* of a dataset, we split it into a list of *n* sentences $(s_1, s_2, ..., s_{n-1}, s_n)$ and then created sub-lists by taking a specific *percentage amount* (*X*%) of the sentences in the beginning and the end of the list (see *Figure 1*). To achieve a sufficient analysis, we generated sub-lists with different percentage amounts to estimate the primacy/recency effects (X = 10%, X = 20%, and X = 30%). The sub-list generated by X% of the sentences in the beginning is the so-called *primacy sub-list P*. The sub-list generated by X% of the sentences in the end is the so-called *recency sub-list R*. These two sub-lists are then sent as inputs to a *logistic regression classifier* trained on the entire dataset. Two logit values corresponding to two sub-lists are generated for the review *E* (see *Fig. 1*) and added to the output layer of our network.



Figure 1: Sentiment splitting and primacy/recency calculation for a sample entry.



Figure 2: The structure of our neural network model (*PriRec-BiLSTM*), extending the *BiLSTM* neural network by adding primacy/recency logit values into the output layer of the neural network.



Figure 3: The architecture of neural network *Model 1 (BiLSTM)* and the extended neural network *Model 2 (PriRecBiLSTM)*

3.2. Neural Network Models

To analyze primary/recency effects, we used two neural network models. Model 1 is BiLSTM [38, 39, 40], a neural network of Bi-directional Long Short Term Memory. Model 2 is the so-called PriRec-BiLSTM that extends Model 1 by integrating an addition layer that adds primacy/recency values before sending to the output layer of the model (see the structure of Model 2 in Figure 2). The architecture of the two mentioned models is depicted in Figure 3, where Model 1 contains three layers: an embedding layer, a BiLSTM layer, and a dense layer. Model 2 extends Model 1 by inserting one further input layer and an addition layer in the end of the model. The aim of these neural networks is to update the weights of the embedding layer matrix. These weights are updated during the training process. At each time, they pass each of the layers to the output sigmoid layer. The value from the sigmoid dense layer ranges between 0 and 1, which indicates a probability output. If the probability value is less than 0.5 then the review is classified to a 'negative' class. Otherwise, the review belongs to a 'positive' class. For Model 2, further steps should be done after getting a probability value. In particular, this value needs to be inverted into a value in the range of $[-\infty, \infty]$, which then can be used to add primacy and recency values generated in the additional input layer. These values are then reconverted into a sigmoid value ($\in [0, 1]$) in order to predict the output class. The sigmoid value can be calculated using Formula 1.

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

4. Primacy/Recency Effects Evaluation

In order to estimate the influences of primacy/recency effects, we compare the performance (in terms of classification accuracy) between the two mentioned neural network models (*PriRec-BiLSTM* vs. *BiLSTM*) in different datasets. In the following subsections, we present in further details on the datasets as well as the evaluation results.

4.1. Datasets

We selected *five* available datasets to evaluate our neural network model: *IMDB*², *Amazon Smartphones*³, *Amazon Industrial and Scientific*⁴, *Tweets Sentiment 140*⁵, and *US Airline tweets*⁶. These datasets have different rating attributes. For instance, *Amazon* datasets use *5-score* rating scales (e.g., 5-star rating scale) [41, 42], whereas *IMDB* dataset uses *2-score* ones (positive/negative). Due to these differences, we had to normalize the datasets to a standard score with only two values - *'positive'* and *'negative'*. For 5-score Amazon datasets, entries rated with 1 or 2 stars were put into a *'negative'* class and those rated from 3 to 5 stars were assigned to a *'positive'* class. In case a dataset contains *'neutral'* entries, we counted them as *'positive'* entries. Details of the datasets with regard to the number of entries, distribution of positive/negative reviews, number of sentences/words in each entry are summarized in *Table 1* and *Table 2*.

Table 1

Sentiment distribution of the datasets.

| Dataset | #pos. entries | #neg. entries | %pos. entries | %neg. entries |
|-----------------------|---------------|---------------|---------------|---------------|
| IMDB | 25000 | 25000 | 50.0% | 50.0% |
| Amazon Smartphones | 170096 | 24343 | 87.48% | 12.51% |
| Amazon Industrial and | 72653 | 4418 | 94.26% | 5.73% |
| Scientific | | | | |
| Tweets Sentiment 140 | 800000 | 800000 | 50% | 50% |
| US Airline Tweets | 5462 | 9178 | 37.30% | 62.69% |

For each dataset, we split it with the ratio 80:20, i.e., 80% for training and 20% for testing. Since most of the datasets are relatively large, not the entire dataset is used immediately in an instance. A random selection of 50000 samples, equally split between positive and negative entries (i.e., 25000 for each), was chosen for a dataset. If the number of samples for either class, positive or negative, is smaller than 25000, then that number was chosen as a baseline and used for dataset splitting. In each dataset, we ran *three iterations* for the random selection. The results of the datasets are the averages of these three runs.

²https://www.kaggle.com/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews

³https://jmcauley.ucsd.edu/data/amazon

⁴https://jmcauley.ucsd.edu/data/amazon

⁵https://www.kaggle.com/kazanova/sentiment140

⁶https://www.kaggle.com/crowdflower/twitter-airline-sentiment

| Dataset | max. no. of sen- tences | avg. no. of sen- tences | max. no. words | of | avg. no. of words |
|-----------------------|----------------------------|----------------------------|-------------------|----|-------------------|
| IMDB | 148 | 12.044 | 2366 | | 219.75 |
| Amazon Smartphones | 255 | 5.03 | 5077 | | 86.63 |
| Amazon Industrial and | 231 | 3.30 | 5621 | | 41.38 |
| Scientific | | | | | |
| Tweets Sentiment 140 | 26 | 1.67 | 52 | | 12.05 |
| US Airline Tweets | 9 | 1.95 | 30 | | 15.64 |

 Table 2

 The average and maximum number of sentences/words in the entries of the datasets.

4.2. Dataset Pre-processing

Our dataset pre-processing includes some typical tasks such as word tokenization of dataset entries, removing stop words, removing single alphanumeric characters, removing links, removing multiple spaces, lowercase words, and simplifying words via lemmatization. We used Natural Language Toolkit library to perform these tasks. Two additional tasks needed for sentiment analysis are sequencing and padding. Sequencing allows to store all the words of an entry into a list and then sequences them into numbers so that each word has a unique number attributable to it. Padding increases the length of the list to specific number that reflects the maximum size of the padding. For instance, let N = 6 be the maximum size of the padding. An entry contains words that have been sequenced in a list (1,3,2,4) where each of the numbers represents a word. Padding increases the length of this list to N by padding it with zeroes. If the sequence is longer than N, then the sequence is truncated. The embedding layer requires all individual entries to be in the same length. In our evaluation, we chose the number N as the maximum estimated number of words detected in the entire dataset.



Figure 4: Classification accuracy between our approach (*PriRec-BiLSTM*) and neural network *BiLSTM* in the **IMDB** dataset.

4.3. Evaluation Results

In all datasets, we used a batch size of 128, the embedding matrix with 300 dimensions, and the regularization value of 0.00001. All the evaluations were run on the computer with Windows 10 Education 64bit, RAM 16GB DDR4-3200, and CPU Ryzen 5 5600H - 6 cores. The performance evaluations (in terms of *classification accuracy*) of our neural network model (*PriRec-BiLSTM* -



Figure 5: Classification accuracy between our approach (*PriRec-BiLSTM*) and neural network *BiLSTM* in the **Amazon Smartphones** dataset.



Figure 6: Classification accuracy between our approach (*PriRec-BiLSTM*) and neural network *BiLSTM* in the **Amazon Industrial and Scientific** dataset.



Figure 7: Classification accuracy between our approach (*PriRec-BiLSTM*) and neural network *BiLSTM* in the **Tweets Sentiment 140** dataset.



Figure 8: Classification accuracy between our approach (*PriRec-BiLSTM*) and neural network *BiLSTM* in the US Airline Tweets dataset.

with primacy/recency effects) in the selected datasets are depicted in *Fig. 4-8*. In each dataset, we ran our approach with different amounts of sentences (X%) to create the primacy and recency sub-lists (X = 10%, X = 20%, and X = 30%). The experimental results show that our proposed

neural network model, in most cases, achieves better performance compared to the basic neural network model (*BiLSTM*). However, we can observe slightly different results depending on the datasets. In few cases, the basic neural network outperforms the neural network with primacy/recency effects.

For the *IDMB* dataset with X = 10%, the basic neural network model *BiLSTM* outperforms *PriRec-BiLSTM* at the sample size of 50000 entries. However, with a higher percentage amounts (X = 20% and X = 30%), the performance of *PriRec-BiLSTM* incrementally increases and therefore triggers a better performance compared to *BiLSTM* (see *Fig. 4*). Similarly, in the *Amazon Smartphones* dataset with X = 10%, there exists an increase tendency in terms of classification accuracy for the neural network *BiLSTM* at the sample size of 4000 entries. However, in two other variants (X=20% and X=30%), *PriRec-BiLSTM* always works better than BiLSTM (see *Fig. 5*). The datasets *Amazon Industrial and Scientific, Tweets Sentiment 140*, and *US Airline Tweets* do not show at any points where *BiLSTM* outperforms *PriRec-BiLSTM* (see *Fig. 6 - 8*). Especially, with X = 30%, the *Industrial and Scientific* dataset even shows a significantly higher performance of *PriRec-BiLSTM* (see *Fig. 500*) entries.

5. Conclusions and Future Work

This paper proposes a neural network of Bidirectional Long Short Term Memory that takes primacy/recency effects into account. To estimate the impacts of these effects, we ran our approach with different datasets and compared the classification accuracy of this approach with this of the neural network *BiLSTM*. The experimental results show that primacy/recency effects positively influence sentiment classification. We have proven that taking into account these effects can help to increase the performance of the *BiLSTM* neural network.

Our approach shows some disadvantages. The first one lines in the method used for sentiment classification, which is dependent on the performance of the logistic regression classifier. Another limitation lies in the neural network structure, which does not consist of many layers, and the method used in the addition layer is pretty simple. Besides, the pre-processing of the text does not involve textual spelling correction, which could add some noise to the datasets. To address the mentioned limitations, we propose some suggestions for future work as follows: (1) adopt other sentiment classification approaches (e.g., *Lexicon-based approach*) to provide better analysis results for primacy/recency effects, (2) train the neural network taking into account the primacy/recency values, and (3) implement a neural network model based on the *attention mechanism* that is tuned or adapted to primacy/recency effects.

References

- L. Kurniasari, A. Setyanto, Sentiment analysis using recurrent neural network, Journal of Physics: Conference Series 1471 (2020) 12–18. doi:10.1088/1742-6596/1471/1/ 012018.
- [2] A. Poushneh, R. Rajabi, Can reviews predict reviewers' numerical ratings? the underlying mechanisms of customers' decisions to rate products using latent dirichlet allocation (lda),

Journal of Consumer Marketing 39 (2022) 230-241. doi:https://doi.org/10.1108/ JCM-09-2020-4114.

- [3] D. M. E.-D. M. Hussein, A survey on sentiment analysis challenges, 2018. URL: http: //dx.doi.org/10.1016/j.jksues.2016.04.002. doi:10.1016/j.jksues.2016.04.002.
- [4] M. Mandl, A. Felfernig, E. Teppan, M. Schubert, Consumer decision making in knowledgebased recommendation, Journal of Intelligent Information Systems 37 (2011) 1–22.
- [5] E. C. Teppan, M. Zanker, Decision biases in recommender systems, Journal of Internet Commerce 14 (2015) 255–275.
- [6] A. Felfernig, G. Friedrich, B. Gula, M. Hitz, T. Kruggel, G. Leitner, R. Melcher, D. Riepan, S. Strauss, E. Teppan, O. Vitouch, Persuasive recommendation: Serial position effects in knowledge-based recommender systems, in: Proceedings of the 2nd International Conference on Persuasive Technology, PERSUASIVE'07, Springer-Verlag, Berlin, Heidelberg, 2007, pp. 283–294.
- [7] A. Felfernig, Biases in decision making, in: Proceedings of the International Workshop on Decision Making and Recommender Systems 2014, CEUR, Bolzano, Italy, 2014, pp. 32–37.
- [8] J. Murphy, C. F. Hofacker, R. Mizerski, Primacy and recency effects on clicking behavior., J. Comput. Mediat. Commun. 11 (2006) 522–535. URL: http://dblp.uni-trier.de/db/journals/ jcmc/jcmc11.html#MurphyHM06.
- [9] T. N. T. Tran, M. Atas, A. Felfernig, R. Samer, M. Stettinger, Investigating serial position effects in sequential group decision making, in: Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization, UMAP'18, ACM, New York, NY, USA, 2018, pp. 239–243.
- [10] B. Agarwal, N. Mittal, Machine Learning Approach for Sentiment Analysis, 2014, pp. 193–208. doi:10.4018/978-1-4666-6086-1.ch011.
- [11] M. Biltawi, W. Etaiwi, S. Tedmori, A. Hudaib, A. Awajan, Sentiment classification techniques for arabic language: A survey, in: 2016 7th International Conference on Information and Communication Systems (ICICS), 2016, pp. 339–346. doi:10.1109/IACS.2016.7476075.
- [12] I. Gupta, N. Joshi, Enhanced twitter sentiment analysis using hybrid approach and by accounting local contextual semantic, Journal of Intelligent Systems 29 (2020) 1611–1625. URL: https://doi.org/10.1515/jisys-2019-0106. doi:doi:10.1515/jisys-2019-0106.
- [13] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, M. Stede, Lexicon-based methods for sentiment analysis, Computational Linguistics 37 (2011) 267–307. doi:10.1162/COLI_a_00049.
- [14] B. Vimalkumar, M. Bhumika, Analysis of various sentiment classification techniques, International Journal of Computer Applications 140 (2016) 22–27. doi:10.5120/ ijca2016909259.
- [15] F. Nipher, On the distribution of errors in numbers written from memory, Transactions of the Academy of Science of St. Louis 3 (2022) ccx-ccxi.
- [16] H. Ebbinghaus, Memory: A contribution to experimental psychology, Annals of neurosciences 20 (2013) 155–156. doi:10.5214/ans.0972.7531.200408.
- [17] E. Kirkpatrick, An experimental study of memory, Psychological Review 1 (1970) 602–609. doi:10.1037/h0068244.
- [18] K. Lashley, The problem of serial order in behavior, Cerebral mechanisms in behavior, the Hixon Symposium (1951).

- [19] K. Y. T. R. Dixon, D. L. H. (Eds.)., Verbal behaviour and behaviour theory, Prince-Hall, 1968, pp. 122–148.
- [20] R. Burke, A. Felfernig, M. H. Göker, Recommender systems: An overview, AI Magazine 32 (2011) 13–18.
- [21] S. Highhouse, A. Gallo, Order effects in personnel decision making, Human Performance 10 (1997) 31–46.
- [22] T. N. T. Tran, C. I. Baumann, A. Felfernig, V.-M. Le, The immunity of users' item selection from serial position effects in multi-attribute item recommendation scenarios, in: P. Brusilovsky, M. de Gemmis, A. Felfernig, E. Lex, P. Lops, G. Semeraro, M. C. Willemsen (Eds.), Proceedings of the 8th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems (IntRS 2021), volume 2948 of *CEUR Workshop Proceedings*, 2021, pp. 101–111. URL: https://recsys.acm.org/recsys21/intrs/.
- [23] T. N. T. Tran, C. Baumann, A. Felfernig, V.-M. Le, The immunity of users' item selection from serial position effects in multi-attribute item recommendation scenarios, in: P. Brusilovsky, M. de Gemmis, A. Felfernig, E. Lex, P. Lops, G. Semeraro, M. Willemsen (Eds.), Proceedings of the 8th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems (IntRS 2021), Online Event, September 25 and September 29, 2021., volume 2948 of *CEUR Workshop Proceedings*, 2021, pp. 101–111. URL: https://recsys.acm.org/recsys21/intrs/, 8th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems : co-located with 15th ACM Conference on Recommender Systems (RecSys 2021), IntRS 2021 ; Conference date: 25-09-2021 Through 29-09-2021.
- [24] T. N. T. Tran, A. Felfernig, N. Tintarev, Humanized recommender systems: State-ofthe-art and research issues, ACM Trans. Interact. Intell. Syst. 11 (2021). URL: https: //doi.org/10.1145/3446906. doi:10.1145/3446906.
- [25] A. Go, R. Bhayani, L. Huang, Twitter sentiment classification using distant supervision, Processing (2009) 1–6. URL: http://www.stanford.edu/~alecmgo/papers/ TwitterDistantSupervision09.pdf.
- [26] E. Kouloumpis, T. Wilson, J. Moore, Twitter sentiment analysis: The good the bad and the omg!, in: ICWSM, 2011, pp. 538–541.
- [27] S. Mohammad, S. Kiritchenko, X. Zhu, NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets, in: Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013), Association for Computational Linguistics, Atlanta, Georgia, USA, 2013, pp. 321–327. URL: https://aclanthology.org/S13-2053.
- [28] A. Pak, P. Paroubek, Twitter as a corpus for sentiment analysis and opinion mining, in: N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, D. Tapias (Eds.), LREC, European Language Resources Association, 2010, pp. 1320–1326. URL: http://dblp.uni-trier.de/db/conf/lrec/lrec2010.html#PakP10.
- [29] D. Stojanovski, G. Strezoski, G. Madjarov, I. Dimitrovski, I. Chorbev, Deep neural network architecture for sentiment analysis and emotion identification of twitter messages, Multimedia Tools and Applications 77 (2018) 32213–32242.
- [30] M. Ahmad, S. Aftab, S. Muhammad, S. Awan, Machine learning techniques for sentiment analysis: A review, International Journal of Multidisciplinary Sciences and Engineering 8 (2017) 2045–7057.

- [31] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Ng, C. Potts, Recursive deep models for semantic compositionality over a sentiment treebank, in: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Seattle, Washington, USA, 2013, pp. 1631–1642. URL: https: //aclanthology.org/D13-1170.
- [32] C. dos Santos, M. Gatti, Deep convolutional neural networks for sentiment analysis of short texts, in: Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, Dublin City University and Association for Computational Linguistics, Dublin, Ireland, 2014, pp. 69–78. URL: https://aclanthology. org/C14-1008.
- [33] A. Severyn, A. Moschitti, Twitter sentiment analysis with deep convolutional neural networks, in: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '15, Association for Computing Machinery, New York, NY, USA, 2015, p. 959–962. URL: https://doi.org/10.1145/2766462.2767830. doi:10.1145/2766462.2767830.
- [34] P. Le, W. Zuidema, Compositional distributional semantics with long short term memory, in: Proceedings of the Fourth Joint Conference on Lexical and Computational Semantics, Association for Computational Linguistics, Denver, Colorado, 2015, pp. 10–19. URL: https: //aclanthology.org/S15-1002. doi:10.18653/v1/S15-1002.
- [35] Y. Wang, M. Huang, L. Zhao, X. Zhu, Attention-based lstm for aspect-level sentiment classification, in: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, 2016, p. 606–615. URL: https://www.microsoft.com/en-us/research/ publication/attention-based-lstm-aspect-level-sentiment-classification/.
- [36] Y. Kim, Convolutional neural networks for sentence classification, Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (2014). doi:10. 3115/v1/D14-1181.
- [37] Q. Qian, M. Huang, J. Lei, X. Zhu, Linguistically regularized LSTM for sentiment classification, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Vancouver, Canada, 2017, pp. 1679–1689. URL: https://aclanthology.org/P17-1154. doi:10.18653/v1/P17-1154.
- [38] J. P. Chiu, E. Nichols, Named entity recognition with bidirectional LSTM-CNNs, Transactions of the Association for Computational Linguistics 4 (2016) 357–370. URL: https: //doi.org/10.1162/tacl_a_00104. doi:10.1162/tacl_a_00104.
- [39] L. Zhang, S. Wang, B. Liu, Deep learning for sentiment analysis: A survey, WIREs Data Mining and Knowledge Discovery 8 (2018). URL: https://doi.org/10.1002/widm.1253. doi:10.1002/widm.1253.
- [40] S. Gupta, T. Kanchinadam, D. Conathan, G. Fung, Task-optimized word embeddings for text classification representations, Frontiers in Applied Mathematics and Statistics 5 (2020). URL: https://doi.org/10.3389/fams.2019.00067. doi:10.3389/fams.2019.00067.
- [41] R. He, J. McAuley, Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering, WWW '16: Proceedings of the 25th International Conference on World Wide Web (2016). doi:10.1145/2872427.2883037.
- [42] J. McAuley, C. Targett, Q. Shi, A. van den Hengel, Image-based recommendations on

styles and substitutes, in: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '15, Association for Computing Machinery, New York, NY, USA, 2015, p. 43–52. URL: https://doi.org/10.1145/2766462.2767755. doi:10.1145/2766462.2767755.