Check for updates

RESEARCH ARTICLE

Modelling sentiments based on objectivity and subjectivity with self-attention mechanisms [version 1; peer review: 3 approved with reservations]

Hu Ng ¹, Glenn Jun Weng Chia ¹, Timothy Tzen Vun Yap ¹, Vik Tor Goh ²

¹Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Selangor, 63100, Malaysia

²Faculty of Engineering, Multimedia University, Cyberjaya, Selangor, 63100, Malaysia

V1 First published: 04 Oct 2021, 10:1001
<https://doi.org/10.12688/f1000research.73131.1>

Latest published: 17 May 2022, 10:1001
<https://doi.org/10.12688/f1000research.73131.2>

Abstract

Background: The proliferation of digital commerce has allowed merchants to reach out to a wider customer base, prompting a study of customer reviews to gauge service and product quality through sentiment analysis. Sentiment analysis can be enhanced through subjectivity and objectivity classification with attention mechanisms.

Methods: This research includes input corpora of contrasting levels of subjectivity and objectivity from different databases to perform sentiment analysis on user reviews, incorporating attention mechanisms at the aspect level. Three large corpora are chosen as the subjectivity and objectivity datasets, the Shopee user review dataset (ShopeeRD) for subjectivity, together with the Wikipedia English dataset (Wiki-en) and Internet Movie Database (IMDb) for objectivity. Word embeddings are created using Word2Vec with Skip-Gram. Then, a bidirectional LSTM with an attention layer (LSTM-ATT) imposed on word vectors. The performance of the model is evaluated and benchmarked against classification models of Logistics Regression (LR) and Linear SVC (L-SVC). Three models are trained with subjectivity (70% of ShopeeRD) and the objectivity (Wiki-en) embeddings, with ten-fold cross-validation. Next, the three models are evaluated against two datasets (IMDb and 20% of ShopeeRD). The experiments are based on benchmark comparisons, embedding comparison and model comparison with 70-10-20 train-validation-test splits. Data augmentation using AUG-BERT is performed and selected models incorporating AUG-BERT, are compared.

Results: L-SVC scored the highest accuracy with 56.9% for objective embeddings (Wiki-en) while the LSTM-ATT scored 69.0% on subjective embeddings (ShopeeRD). Improved performances were observed with data augmentation using AUG-BERT, where the LSTM-ATT+AUG-BERT model scored the highest accuracy at 60.0% for objective embeddings and 70.0% for subjective embeddings, compared to 57% (objective) and 69% (subjective) for L-SVC+AUG-BERT, and 56%

Open Peer Review

Approval Status ? ✓ ?

	1	2	3
version 2 (revision) 17 May 2022		✓ view	
version 1 04 Oct 2021	? view	? view	? view

1. **Ashima Yadav** , Bennett University, Greater Noida, India
2. **Hien D. Nguyen** , University of Information Technology, VNU-HCM, Ho Chi Minh City, Vietnam
3. **Marco Polignano** , University of Bari 'Aldo Moro', Bari, Italy

Any reports and responses or comments on the article can be found at the end of the article.

(objective) and 68% (subjective) for L-SVC.

Conclusions: Utilizing attention layers with subjectivity and objectivity notions has shown improvement to the accuracy of sentiment analysis models.

Keywords

Sentiment analysis, subjectivity, objectivity, attention mechanism, neural nets.

This article is included in the **Research Synergy** Foundation gateway.

This article is included in the **Artificial Intelligence and Machine Learning** gateway.

Corresponding author: Hu Ng (nghu@mmu.edu.my)

Author roles: **Ng H:** Conceptualization, Methodology, Project Administration, Supervision, Validation, Writing – Review & Editing; **Chia GJW:** Data Curation, Methodology, Resources, Software, Writing – Original Draft Preparation; **Yap TTV:** Conceptualization, Methodology, Supervision, Validation, Writing – Review & Editing; **Goh VT:** Conceptualization, Validation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2021 Ng H *et al.* This is an open access article distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Ng H, Chia GJW, Yap TTV and Goh VT. **Modelling sentiments based on objectivity and subjectivity with self-attention mechanisms [version 1; peer review: 3 approved with reservations]** F1000Research 2021, 10:1001 <https://doi.org/10.12688/f1000research.73131.1>

First published: 04 Oct 2021, 10:1001 <https://doi.org/10.12688/f1000research.73131.1>

Introduction

The proliferation of digital commerce, especially in Malaysia, has allowed many local merchants to reach out to a wider customer base. In order to attract customer's attention, merchants always compete to offer better price and higher quality of services. Besides that, they also seriously consider the customer feedback or reviews in order to gauge service and product quality.¹

By exploring the sentiment tendency of customer reviews, it can provide a good reference for other customer before the purchasing decision is made. Besides, it helps merchants to improve service quality and customer satisfaction.

Sentiment analysis is aimed to determine the sentiment as well as polarity on part of a text. Normally, language terms are under two form of statements, namely fact statement and a non-fact statement, which are known as objective and subjective in categorical terms.² Facts are objective terms like events entities and their properties. On the other hand, a non-fact statement is subjective and usually related to an individual's sentiments, personal beliefs, opinion, perspective, feelings or thoughts.

This paper adopted attention segment³ to a neural network, LSTM, by creating attention-weighted features, namely Long Short Term Memory with Attention (LSTM-ATT)⁴ to create attention-weighted features. It aims to introduce these features at the input level to the neural network, so that the performance of sentiment can be increased.

In business operations, sentiment classification is the automated process of classifying customer reviews in text and labeling them as positive, negative, or neutral. It employs Natural Language Processing (NLP) routine to interpret subjective data, so that it can help merchants to understand how customers feel about their products or services. In this paper, sentiment classification is performed by non-deep neural network classifiers, namely Logistic Regression (LR) and Linear SVC (L-SVC) together with Deep Learning (Artificial Neural Networks) classifier by adopting attention mechanisms (LSTM-ATT).

Word embedding

Word embeddings are a scheme to convert human language to a word representation that is understandable by computers. The word representation is in the form of a real-valued vector that encodes the meaning of the word, so that the words that are closer in the vector space are expected to be similar in meaning.

Collobert⁵ declared that a distinction word vector and proper training can increase the performance of NLP works especially the sentiment analysis. Word embedding can be classified into two types; contextual and non-contextual embeddings. Non-contextual embedding does not consider the effects of arrangement of words in a particular sentence, while contextual embedding does the opposite.

For non-contextual embedding, Mikolov *et al.* initiated Word2Vec.⁶ The word2vec algorithm uses a neural network model to learn word associations from a large corpus of text. Once trained, such a model can detect synonymous words or suggest additional words for a partial sentence. Bengio *et al.*⁷ and Collobert *et al.*⁸ enhanced it by implementing Neural Net Language Model (NNLM). Bojanowski *et al.*⁹ made enhancement on Word2Vec by applying n-grams and cables to obtain higher performance in Word Similarity assignments that involved various types of languages and was able to show big enhancements on morphology rich languages, in particular, German datasets such as GUR350 and GUR65¹⁰ and ZG222.¹¹ Bhagat *et al.*¹² applied unigrams to extra individual words from Twitter messages and multiple machine learning techniques to perform sentiment analysis. Ebner *et al.*¹³ employed three simple bag-of-words representations, where a text is represented as the bag (multiset) of its words, namely pooling encoders, pre-trained word embeddings, and unigram generative regularization to regularize incorporating auxiliary discriminative tasks that managed to reduce training time and model size while maintaining high performance. Gayatry¹⁴ employed Count Vectorizer to convert each word into its corresponding vectors.

For context embedding, Peters *et al.*¹⁵ modified LSTM neural nets to create Embedding from Language Models (ELMo) that were able to show better results than the Stanford Tree-bank model (SST-5) from the research work by Socher *et al.*¹⁶ Devlin *et al.*¹⁷ constructed BERT along with Transformers and Attention Mechanism.³ The role of BERT is not limited to embedding functions but also become a language model that is capable to exceed ELMo on General Language Understanding Evaluation assignments (GLUE) from the research outcomes by Wang *et al.*¹⁸ Liu *et al.*¹⁹ enhanced BERT by developing A Robustly Optimized BERT Pre-Training Approach (RoBERTa). RoBERTa omits the Next Sentence Prediction task and applies an unfixed masking configuration rather than static Masked Language Modelling (MLM). Sangeetha²⁰ proposed multi-head attention fusion model of word and context embedding for sentiment analysis of student feedback.

Methods

Ethics approval

Ethical Approval Number: EA1602021 (From Technology Transfer Office (TTO), Multimedia University).

Datasets

Three large corpora datasets were chosen to denote objectivity and subjectivity datasets correspondingly. IMDb²¹ and Wiki-en²² were chosen as the objectivity datasets while ShopeeRD²³ was chosen as the subjectivity dataset.

IMDb consists of 50K of movie reviews with contents based on the true plot and written with a neutral point of view (NPOV). Wiki-en consists of 4677K of records based on Wikipedia that forced the articles to be factual and follow the NPOV policy. ShopeeRD consists of 208K customer reviews taken from the Shopee Code League 2020 Data Science and Data Analytics competition. ShopeeRD's entries are based on customer experiences, which are potentially judgemental and opinionated.

Wiki-en was used as the objectivity corpus for word embedding, while IMDb was used for objectivity sentiment analysis. 70% of the ShopeeRD was used as the subjectivity corpus for word embedding and the remaining 30% for subjectivity sentiment analysis. Figure 1 displays the mapping of datasets.

Data preparations and word embedding

The reviews and records from the datasets underwent a set of data cleaning steps which included emoji cleaning, text cleaning such as repeated character elimination, punctuation (e.g., ?, ! or,) elimination, stop word (e.g., becomes, against, or at) elimination, lemmatization, case lowering and normalization (normalizing non-English writing into English writing).

Word embedding is carried out to transform the reviews into floating-point numbers that are stored in a high dimension array, which forms a dictionary that the computer is able to obtain word vectors from. The word embedding must be large enough to represent millions of words and for each word is denoted as a high dimension vector. In this paper, one word is represented as a 300-dimension vector.

Word2Vec by Mikolov et al.²⁴ is a word embedding method that consist of two structural design, namely Skip-gram and Continuous Bag-of-Words (CBOW). In the CBOW model, the distributed representations of context (or surrounding words) are combined to predict the word in the middle, while in the Skip-gram model, the distributed representation of the

Figure 1. The mapping of datasets.

input word is used to predict the context. It has been proven that the Skip-gram structure has been shown better results in comparison to Continuous Bag-of-Words.²⁵⁻²⁷ Hence, this paper utilizes Word2Vec Skip-Gram structure to perform word embedding.

ShopeeRD and Wiki-en were trained into embeddings of 300d (300-dimension), with a factor of five negative examples, window dimension of five tokens, and elimination of small sentences. The two embeddings (subjectivity and objectivity) were trained for ten repetitions.

Models

To prevent over-fitting or one model favouring towards a particular embedding, two models (LR and L-SVC) were applied for this paper. In general, a sentence vector is produced from the formation of word vectors.²⁸ Nevertheless, this paper assumes that certain letterings might not apply any weight or produce any consequence, therefore an attention layer, which is adopted from Vaswani *et al.*³ was produced as a substitute. Self-attention is capable of allocating 'attention' to an important vector (keyword). This permits the structural design in a way to highlight attention-ed vectors.²⁹

For that reason, a model integrating attention segments was recommended, and the structural design is presented in Figure 2. The word vectors worked through the attention layer, creating attention-weighted features. By adapting LSTM neural nets, both the original embedding and the attention-weighted embedding are concatenated to create sentiment features.

This paper adopted attention-weighted features model is called Long Short Term Memory with Attention (LSTM-ATT)⁴ with intention to improve the sentiment performance. These features at the input level to the neural network and go through a few dense layers to flatten the output. Finally, Rectified Linear Unit (RELU), a non-linear activation function is applied to produce the sentiment results. The model, LSTM-ATT, is then evaluated against LR and L-SVC. The workflow of the sentiment analysis on IMDb and ShopeeRD with multiple models is illustrated in Figure 3.

Results and discussion

Design of experiments

The experiments were performed in Python, utilizing the scikit-learn library for machine learning as well as the BERT model architecture. Three models (LR, L-SVC and LSTM-ATT) were trained with the objectivity (Wiki-en) and subjectivity (70% of ShopeeRD) embeddings. Ten-fold cross-validation was applied during the training. After that, the models were tested against the objectivity (IMDb) test set and the subjectivity (20% of ShopeeRD) test set to eliminate bias.

Figure 2. Structural design of LSTM-ATT.

Figure 3. Sentiment analysis.

The experiments were based on benchmark comparison, embedding comparison and model comparison with 70-10-20 train-validation-test splits. The validation was carried out to perform parameter tuning, so that the best results among the models could be obtained.

Quality of embeddings

Figures 4 and 5 demonstrate the t-distributed Stochastic Neighbor Embedding (t-SNE) plots for Wiki-en and ShopeeRD embeddings on the top 15 nearest words to the word 'happy'. The t-SNE for both datasets revealed that word similarities are discovered in the embeddings, for instance, 'glad', 'pleased', 'excited' are grouped together with 'happy'.

The words 'very' and 'good' having closeness to 'happy' were only found in t-SNE for Wiki-en only. Meanwhile the words 'satisfied' and 'wonderful' having closeness to 'happy' were found in t-SNE for ShopeeRD. Furthermore, outliers like 'everyone' and 'everybody' were found to appear in the t-SNE for Wiki-en. This shows that the two embeddings are different in nature.

Sentiment analysis

The three models namely LR, L-SVC and LSTM-ATT were evaluated in terms of their performance in sentiment analysis. The accuracy of the three models is presented in Table 1. L-SVC obtained the highest accuracy (56.9%) for objectivity embedding, whereas LSTM-ATT obtained the highest accuracy (69.0%) for subjectivity embedding. L-SVC performed better than LR properly due to L-SVC attempting to exploit the margin between the closest support vectors whereas LR exploits the posterior class probability.³⁰

From Table 1, there is possible limitation factor cause by the capacity of the training data, therefore the size of the dataset is increased through the data augmentation technique.³¹ As LR has a simpler architecture, data augmentation is not considered, and the focus is made on L-SVC and LSTM-ATT. Table 2 presents the outcome of data augmentation.

Figure 4. t-SNE plots of 'happy' on the Wiki-en.

Figure 5. t-SNE plots of 'happy' on the ShopeeRD.

Table 1. Accuracy of three classifiers.

Data	Model		
	LR	L-SVC	LSTM-ATT
Objective embedding	0.5338	0.5685	0.5604
Subjective embedding	0.6418	0.6892	0.6902

Table 2. Accuracy of two classifiers with augmentation technique.

Data	Model	
	L-SVC+AUG-BERT	LSTM-ATT+ AUG-BERT
Objective embedding	0.5746	0.5991
Subjective embedding	0.6907	0.7004

Table 3. Comparison of our models with others research work on ShopeeRD.

Data	Model		
	L-SVC+AUG-BERT [our method]	LSTM-ATT+AUG-BERT [our method]	Multinomial Naïve Bayes ¹⁴
ShopeeRD			
Training (110K), Testing (36K)	-	-	0.58
Training (145K), Testing (41K)	0.69	0.70	-

From Table 2, it is found that the accuracy of models with the augmented data are found to be better than the models, although not by much. The LSTM-ATT+AUG-BERT was able to beat L-SVC+AUG-BERT on both objective and subjective embeddings.

To the best of our knowledge, there is only one sentiment analysis result from Gayatry's work¹⁴ that accepted by Shopee Code League 2020 Data Science.³² Table 3. shows the comparison of our models with Gayatry's work on ShopeeRD.

To the best of our knowledge, there is no research work on objectivity sentiment analysis on IMDB without any involvement of pre-train data from taken IMDB, as all of them used 50% of total dataset for training and another 50% for testing. In this paper, we trained the models by Wiki-en and test on IMDB.

Conclusions

This paper has presented word embeddings for both objectivity and subjectivity contexts by applying Word2Vec. Analyzing the embedding using the t-distributed stochastic neighbour embedding plot shows that there are some similarities between the two embeddings, but a majority of them are dissimilar. Three models namely, LR, L-SVC and LSTM-ATT were employed to evaluate the performance of the adopted embedding technique. The attention model adopted was able to perform sentiment analysis well with the requirement of more data was fed into the model utilizing AUG-BERT data augmentation. Models with differing architectures will be explored in future work.

Data availability

Underlying data

- Compiled Movie reviews from the Internet Movie Database (IMDb) : <https://datasets.imdbws.com/>,²¹ cited on 6 August 2021.

The data are available for personal and non-commercial use, as stipulated by the owner (IMDb).

- A complete copy of all Wikimedia wikis, in the form of wikitext source and metadata embedded in XML: <https://dumps.wikimedia.org/backup-index.html>,²² cited on 6 August 2021.

The data are available under the terms of the Creative Commons Attribution-Share-Alike 3.0 License.

- Product reviews from the Shopee e-commerce platform, created for the Shopee Code League 2020 Data Science and Data Analytics competitions: <https://www.kaggle.com/davydev/shopee-code-league-20>,²³ cited on 6 August 2021.

The data are available for personal and non-commercial use, as stipulated by the owner (Shopee).

References

1. Vanaja S, Belwal M: **Aspect-level sentiment analysis on e-commerce data.** *2018 Int Conf Inventive Res Computing Applications (ICIRCA)*. IEEE; 2018, July; (pp. 1275–1279).
2. Sahu I, Majumdar D: **Detecting factual and non-factual content in news articles.** *Proc fourth ACM IKDD conferences on data sciences*. 2017, March; (pp. 1–12).
[Publisher Full Text](#)
3. Vaswani A, Shazeer N, Parmar N, et al.: **Attention is all you need.** *In Advances in neural information processing systems*. 2017; (pp. 5998–6008). arXiv preprint arXiv:1706.03762.
4. Lee WS, Ng H, Yap TTV, et al.: **Attention Models for Sentiment Analysis Using Objectivity and Subjectivity Word Vectors.** In: Alfred R, Iida H, Haviluddin H, et al. (eds) *Computational Science and Technology. Lecture Notes in Electrical Engineering*. Singapore: Springer; 2021; vol 724.
[Publisher Full Text](#)
5. Collobert R, Weston J, Bottou L, et al.: **Natural language processing (almost) from scratch.** *J Machine Learn Res*. 2011; **12**(ARTICLE), 2493–2537.
6. Mikolov T, Sutskever I, Chen K, et al.: **Distributed representations of words and phrases and their compositionality.** *arXiv preprint arXiv:1310.4546*. 2013.
7. Bengio Y, Ducharme R, Vincent P, et al.: **A neural probabilistic language model.** *J Machine Learning Res*. 2003; **3**: 1137–1155.
8. Collobert R, Weston J: **A unified architecture for natural language processing: Deep neural networks with multitask learning.** *Proc 25th Int Con Machine learning*. 2008, July; (pp. 160–167).
[Publisher Full Text](#)
9. Bojanowski P, Grave E, Joulin A, et al.: **Enriching word vectors with subword information.** *Transactions of the Association for Computational Linguistics*. 2017; **5**: 135–146.
10. Gurevych I: **Using the structure of a conceptual network in computing semantic relatedness.** *Int Conf Natural Language Processing*. Berlin, Heidelberg: Springer; 2005, October; (pp. 767–778).
11. Zesch T, Gurevych I: **Automatically creating datasets for measures of semantic relatedness.** *Proc Workshop Linguistic Distances*. 2006, July; (pp. 16–24).
12. Bhagat A, Sharma A, Chettri S: **Machine Learning Based Sentiment Analysis for Text Message.** *Int J Computing Technol*. 2020.
13. Ebner S, Wang F, Van Durme B: **Bag-of-Words Transfer: Non-Contextual Techniques for Multi-Task Learning.** *Proc 2nd Workshop Deep Learning Approaches for Low-Resource NLP (DeepLo 2019)*. 2019, November; (pp. 40–46).
14. cited on 6 August 2021.
[Reference Source](#)
15. Peters ME, Neumann M, Iyyer M, et al.: **Deep contextualized word representations.** *arXiv preprint arXiv:1802.05365*. 2018.
16. Socher R, Perelygin A, Wu J, et al.: **Recursive deep models for semantic compositionality over a sentiment treebank.** *Proc 2013 Conf Empirical methods in natural language processing*. 2013, October; (pp. 1631–1642).
17. Devlin J, Chang MW, Lee K, et al.: **Bert: Pre-training of deep bidirectional transformers for language understanding.** *arXiv preprint arXiv:1810.04805*. 2018.
18. Wang A, Singh A, Michael J, et al.: **GLUE: A multi-task benchmark and analysis platform for natural language understanding.** *arXiv preprint arXiv:1804.07461*. 2018.
[Publisher Full Text](#)
19. Liu Y, Ott M, Goyal N, et al.: **Roberta: A robustly optimized bert pretraining approach.** *arXiv preprint arXiv:1907.11692*. 2019.
20. Sangeetha K, Prabha D: **Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for LSTM.** *J Ambient Intelligence Humanized Computing*. 2020; 1–10.
21. **IMDB movie review data.** *IMDB.com*. cited on 6 August 2021.
[Reference Source](#)
22. Wikimedia.org: **Wikimedia Downloads.** *Wikimedia.org*. n.d. cited on 6 August 2021.
[Reference Source](#)
23. **Shopee Code League 2020 Data Science.** *kaggle.com*. cited on 6 August 2021.
[Reference Source](#)
24. Mikolov T, Sutskever I, Chen K, et al.: **Distributed representations of words and phrases and their compositionality.** *arXiv preprint arXiv:1310.4546*. 2013.
25. Jang B, Kim I, Kim JW: **Word2vec convolutional neural networks for classification of news articles and tweets.** *PloS one*. 2019; **14**(8): e0220976.
[Publisher Full Text](#)
26. Caselles-Dupré H, Lesaint F, Royo-Letelier J: **Word2vec applied to recommendation: Hyperparameters matter.** *In Proc 12th ACM Conf Recommender Systems*. 2018, September; (pp. 352–356).
27. Li B, Drozd A, Guo Y, et al.: **Scaling word2vec on big corpus.** *Data Science and Engineering*. 2019; **4**(2): 157–175.
[Publisher Full Text](#)
28. Liu H: **Sentiment analysis of citations using word2vec.** *arXiv preprint arXiv:1704.00177*. 2017.
29. Chorowski J, Bahdanau D, Serdyuk D, et al.: **Attention-based models for speech recognition.** *In Advances in neural information processing systems*. 2015; (pp. 577–585). arXiv preprint arXiv: 1506.07503.
30. Sa'id AA, Rustam Z, Wibowo VVP, et al.: **Linear Support Vector Machine and Logistic Regression for Cerebral Infarction Classification.** *2020 Int Conf Decision Aid Sciences Application (DASA)*. 2020, pp. 827–831.
[Publisher Full Text](#)
31. Shi L, Liu D, Liu G, et al.: **AUG-BERT: An Efficient Data Augmentation Algorithm for Text Classification.** *Int Conf Communications, Signal Processing Systems*. Singapore: Springer; 2019, July; (pp. 2191–2198).
[Publisher Full Text](#)
32. cited on 6 August 2021.
[Reference Source](#)

Open Peer Review

Current Peer Review Status: ? ? ?

Version 1

Reviewer Report 08 April 2022

<https://doi.org/10.5256/f1000research.76757.r126996>

© 2022 Polignano M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Marco Polignano

Department of Computer Science, University of Bari 'Aldo Moro', Bari, Italy

The authors present an approach to performing sentiment analysis by comparing different embeddings and classification strategies.

Several weaknesses need to be addressed by the authors to enable publication:

1. Insufficient detail is provided on the motivations for the article. Why is what is presented relevant to scientific research?
2. Numerous approaches in the literature use an attention level LSTM model to do sentiment analysis. What does this work present that is innovative?
3. Technical and implementation details are not provided. Source code is not shared. This makes the proposed contribution non-replicable.
4. It is unclear why non-contextual embeddings were used and not the newer ones based on BERT and ELMO. It is unclear how the data is used in the two phases of embedding space creation and sentiment analysis. Further details on the experimental protocol should be reported.
5. The authors show accuracy as the only metric. This metric is not applicable if the datasets are unbalanced. Unfortunately, such analysis is not performed by the authors. I suggest including the scores of additional metrics such as precision, recall, and f1 measure.
6. A statistical validation of the results has not been conducted, so the differences obtained are not supported by real-world evidence and may be due to chance.

Therefore, it is suggested to further investigate these issues to make the contribution more robust and relevant to the community.

References

1. Polignano M, Basile P, de Gemmis M, Semeraro G: A comparison of word-embeddings in

emotion detection from text using bilstm, cnn and self-attention. *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*. 2019. 63-68

2. Polignano M, Basile P, De Gemmis M, Semeraro G, Basile A: Italian BERT language understanding model for NLP challenging tasks based on tweets. *6th Italian Conference on Computational Linguistics, CLiC-it 2019*. 2019; **2481**: 1-6

3. Polignano M, de Gemmis M, Narducci F, Semeraro G: Do You Feel Blue? Detection of Negative Feeling from Social Media. **10640**: 321-333 [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

No

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Natural Language Processing, Machine Learning, Sentiment Analysis, Recommender Systems

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 23 March 2022

<https://doi.org/10.5256/f1000research.76757.r126992>

© 2022 Nguyen H. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Hien D. Nguyen

University of Information Technology, VNU-HCM, Ho Chi Minh City, Vietnam

This paper introduced an architecture to adopt attention segment to a neural network, called LSTM-ATT, to create attention-weighted features. Through that, sentiment classification is performed by non-deep neural network classifiers, this method is called Logistic Regression (LR) and Linear SVC (L-SVC) together with Deep Learning classifier by LSTM-ATT. The results show that L-SVC scored the highest accuracy with 56.9% for objective embeddings (Wiki-en) and LSTM-ATT scored 69.0% on subjective embeddings. The authors also experimented by integrating the proposed method with data augmentation using AUG-BERT. The LSTM-ATT+AUGBERT model scored the highest accuracy at 60.0% for objective embeddings and 70.0% for subjective embeddings, and 57% (objective) and 69% (subjective) for L-SVC+AUG-BERT.

This paper is well organized. However, the authors should revise as follows:

1. The authors should explain more the importance about modeling sentiment based on objectivity and subjectivity.
2. This study has to present more the architectures in Figure 2 and Figure 3.
3. The results need to be explained more their meaning. The authors should compare the proposed method with other similar methods, such as:

Nguyen, H., Huynh, T., Hoang, S., et al. (2020). Language-oriented Sentiment Analysis based on the Grammar Structure and Improved Self-attention Network. ENASE 2020, pp. 339-346¹

Zainuddin, N., Selamat, A., Ibrahim, R. (2018). Hybrid sentiment classification on twitter aspect-based sentiment analysis. Applied Intelligence 48(5): 1218 – 1232²

4. The references [14] and [32] should be added their titles.

In my opinion, this paper need to be revised before the final decision.

References

1. Nguyen H, Huynh T, Hoang S, Pham V, et al.: Language-oriented Sentiment Analysis based on the Grammar Structure and Improved Self-attention Network. *Proceedings of the 15th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE 2020)*. 2020. 339-346 [Publisher Full Text](#)
2. Zainuddin N, Selamat A, Ibrahim R: Hybrid sentiment classification on twitter aspect-based sentiment analysis. *Applied Intelligence*. 2017. [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Intelligent system, knowledge engineering, data science, automated reasoning, machine learning

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 13 May 2022

Hu NG

Dear Prof Nguyen,

We are greatly appreciative of the insightful comments and helpful suggestions that you have provided.

The following are our response on the issues that you have highlighted:

Your Comment

1. The authors should explain more the importance about modeling sentiment based on objectivity and subjectivity.

Our Response

We are of the opinion that once subjectivity is identified the sentiment accuracy can be further increased as objective statements should not contribute to this.

Your Comment

2. This study has to present more the architectures in Figure 2 and Figure 3.

Our Response

The references have also been corrected.

Your Comment

3. The results need to be explained more their meaning. The authors should compare the proposed method with other similar methods, such as:

Nguyen, H., Huynh, T., Hoang, S., et al. (2020). Language-oriented Sentiment Analysis based on the Grammar Structure and Improved Self-attention Network. ENASE 2020, pp. 339-3461

Zainuddin, N., Selamat, A., Ibrahim, R. (2018). Hybrid sentiment classification on twitter aspect-based sentiment analysis. Applied Intelligence 48(5): 1218 – 12322

Our Response

As we are using different datasets, we could not compare the results as is. However, we will perform this in future work. In addition, we take these papers into consideration and have cited them in the paper.

Nguyen, H. D., Huynh, T., Hoang, S. N., Pham, V. T., & Zelinka, I. (2020, May). Language-oriented Sentiment Analysis based on the Grammar Structure and Improved Self-attention Network. In *ENASE* (pp. 339-346).

4. The references [14] and [32] should be added their titles.

Our Response

The references have also been corrected.

Competing Interests: No competing interests were disclosed.

Reviewer Report 28 February 2022

<https://doi.org/10.5256/f1000research.76757.r123733>

© 2022 Yadav A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ashima Yadav

Bennett University, Greater Noida, Uttar Pradesh, India

This paper has presented word embeddings for both objectivity and subjectivity contexts by applying Word2Vec. Analyzing the embedding using the t-distributed stochastic neighbour embedding plot shows that there are some similarities between the two embeddings, but a majority of them are dissimilar. Three models namely, LR, L-SVC and LSTM-ATT were employed to evaluate the performance of the adopted embedding technique. The attention model adopted was able to perform sentiment analysis well with the requirement of more data was fed into the model utilizing AUG-BERT data augmentation. The authors have very well addressed the problem. However, I still have following suggestions for them:

1. The authors should give a more detailed description of the deep learning based methods used in sentiment analysis. (Sentiment analysis using deep learning architectures: a review¹) (A Weighted Text Representation framework for Sentiment Analysis of Medical Drug Reviews²)
2. Also, they should tell how their proposed model is different from the other baseline models (apart from giving higher accuracy) as the area of sentiment classification have many popular models.

3. Explain the role of attention mechanism in sentiment analysis. Can word-level and sentence level attention give better accuracy as explained in (A Language-independent Network to Analyze the Impact of COVID-19 on the World via Sentiment Analysis³)?
4. In the Introduction section, a bullet-wise summary of the significant contribution of this work is required, which could highlight the motivation for this study.
5. There should be a Related work section in the manuscript which should explain how their proposed model differs from baseline methods or the advantages of their model over the baseline ones.

References

1. Yadav A, Vishwakarma D: Sentiment analysis using deep learning architectures: a review. *Artificial Intelligence Review*. 2020; **53** (6): 4335-4385 [Publisher Full Text](#)
2. Yadav A, Vishwakarma DK: A Weighted Text Representation framework for Sentiment Analysis of Medical Drug Reviews. *2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM)*. 2020. [Publisher Full Text](#)
3. Yadav A, Vishwakarma D: A Language-independent Network to Analyze the Impact of COVID-19 on the World via Sentiment Analysis. *ACM Transactions on Internet Technology*. 2022; **22** (1): 1-30 [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Sentiment analysis, deep learning, machine learning, attention mechanism

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 13 May 2022

Hu NG

Dear Prof Yadav,

We are greatly appreciative of the insightful comments and helpful suggestions that you have provided.

The following are our response on the issues that you have highlighted:

Your Comment

1. The authors should give a more detailed description of the deep learning based methods used in sentiment analysis.

Our Response

Additional papers have been cited to provide more context on the deep learning based methods in the Literature Review.

Your Comment

2. Also, they should tell how their proposed model is different from the other baseline models (apart from giving higher accuracy) as the area of sentiment classification have many popular models.

Our Response

This model places emphasis on attention mechanisms which in this context perform better than other baseline models. Remarks about this is added to the Methods.

Your Comment

3. Explain the role of attention mechanism in sentiment analysis. Can word-level and sentence level attention give better accuracy as explained in (A Language-independent Network to Analyze the Impact of COVID-19 on the World via Sentiment Analysis3)?

Our Response

The attention mechanism allows the model to utilize the most relevant parts of the input sequence in a flexible manner, by a weighted combination of all of the encoded input vectors, with the most relevant vectors being attributed the highest weights. We have not looked into sentence level attention as the focus of this paper is on word-level, and sentence-level will be explored in future work.

Your Comment

4. In the Introduction section, a bullet-wise summary of the significant contribution of this work is required, which could highlight the motivation for this study.

Our Response

A new paragraph has been added to highlight the contribution of this research work. We believe the paragraph format is more appropriate as follow the format of the journal

Your Comment

5. There should be a Related work section in the manuscript which should explain how their proposed model differs from baseline methods or the advantages of their model over the baseline ones.

Our Response

The use of attention mechanisms in this context proves to be beneficial, and this is the advantage over baseline methods.

Our Response

New references added

Yadav, A., & Vishwakarma, D. K. (2020). Sentiment analysis using deep learning architectures: a review. *Artificial Intelligence Review*, 53(6), 4335-4385.

Yadav, A., & Vishwakarma, D. K. (2020, September). A weighted text representation framework for sentiment analysis of medical drug reviews. In *2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM)* (pp. 326-332). IEEE.

Yadav, A., & Vishwakarma, D. K. (2021). A Language-independent Network to Analyze the Impact of COVID-19 on the World via Sentiment Analysis. *ACM Transactions on Internet Technology (TOIT)*, 22(1), 1-30.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research