



A Model for Classifying Social Commerce Texts Using Deep Learning

A. Naserasadi¹ , M. Estilaei^{2,*}

¹ Assistant Professor, Department of Computer Science, Zaran Higher Education Complex

² Lecturer, Department of Engineering, Payame Noor University, Tehran, Iran

ARTICLE INFO	ABSTRACT
<p>Article History: Received 1 June 2020 Received in revised form 8 July 2020 Accepted 28 September 2020 Available online 28 September 2020</p>	<p>With the rapid expansion of online commerce, a vast amount of data related to consumer interactions, product reviews, and purchasing behaviors is continuously generated and shared on social media platforms. Effectively analyzing and processing these data streams can provide valuable insights for enhancing social commerce, optimizing customer experience management, improving targeted advertising, and boosting consumer engagement. One of the key challenges in this domain is the automated classification of social commerce texts, which plays a critical role in understanding user sentiments, preferences, and purchasing intentions. In this paper, we propose a deep learning-based model for classifying social commerce texts, leveraging pre-trained language models to enhance feature extraction and improve classification accuracy. The proposed approach first employs a pre-trained language model to generate high-dimensional text feature vectors, which are then utilized in a classification framework to accurately categorize social commerce-related texts. The effectiveness of the model is evaluated using benchmark datasets, demonstrating its capability to accurately classify social commerce texts. The experimental results indicate that the proposed method achieves an average precision score of 0.725 and an average recall score of 0.708, highlighting its potential for real-world applications in automated social commerce analytics.</p>
<p>Keywords: Text Classification, Social Commerce, Deep Learning, Natural Language Processing, Machine Learning</p>	

1. INTRODUCTION

With the rapid advancement of the internet and smartphones, these phenomena have gradually become essential platforms for shopping, entertainment, and travel. In recent years, despite the overall economic conditions worldwide, the volume of transactions in e-commerce has continued to rise, making social commerce a new driving force for the growth of online consumption [1]. Social commerce can enhance transactions by incorporating social network elements such as following, sharing, discussing, and interacting in e-commerce services. Compared to the traditional product-centric e-commerce model, social commerce is human-centric, spreading product information through user reviews and word-of-mouth advertising, thus stimulating and increasing consumer demand. The most common method of this process is utilizing users' textual reviews about products and services [2]. This results in a large volume of texts related to social commerce, and processing these naturally generated texts can be beneficial

* Corresponding Author: uast.info@yahoo.com

Lecturer, Department of Engineering, Payame Noor University, Tehran, Iran



for the further development of social commerce. One type of processing on natural language with multiple applications is text classification. Text classification refers to the automatic organization of a set of texts into several predefined categories based on their content. In social commerce, this process can have various applications such as marketing, customer experience management, and advertising. However, it is important to note that the classification of social commerce texts differs significantly from other text classifications due to their nature and conditions. These differences include the short length of social commerce texts (up to 200 words), the informal and conversational nature of social commerce texts, the use of common internet words and emojis, the wide diversity in social commerce texts due to the variety of products and services, and the dependency of these texts on the nature of the product or service and the use of specific phrases and words in each field [3].

Given the aforementioned points and the importance of the topic, this paper presents a model for classifying social commerce texts using deep learning and the ALBERT language model. In this model, after pre-training the language model with a suitable textual corpus and then fine-tuning its parameters with related datasets, the process of classifying social commerce texts is carried out. The rest of the paper is structured as follows: Section 2 reviews the background and significance of the research, Section 3 presents the proposed model, Section 4 describes the conditions and results of implementing the model, and Section 5 provides a summary, conclusions, and suggestions for future research.

2. BACKGROUND AND SIGNIFICANCE OF RESEARCH

Numerous studies have been conducted on text classification. These studies have mainly focused on three aspects: text feature extraction and representation, selection and improvement of classification algorithms, and evaluation of classification methods [4]. Notable models include those proposed in [5], which used feedforward neural networks for this task. Additionally, recurrent neural networks were employed for text classification in [6] and [7]. Convolutional neural networks were utilized in [8] and [9], while capsule networks were explored in [10] and [11]. Attention mechanisms were used in [12], augmented memory networks in [13], graph neural networks in [14], Siamese neural networks in [15], and hybrid models in [16]. Pre-trained models were used in [17], adversarial learning in [18], and reinforcement learning in [19].

In terms of selecting and improving text classification algorithms, studies like [20] used the ELMO model based on long short-term memory. Additionally, [21] employed encoder-based models like Trans, [22] used irregularly weighted k-nearest neighbor algorithms, and [23] utilized occurrence matrices to leverage local information.

The above research has extensively explored text feature representation and extraction, as well as proposing and improving suitable classification algorithms, which have significantly impacted improving results. However, for texts in specific domains like social commerce with unique structures and conditions, as previously mentioned, text features may not be well-extracted, and entities and words may not fully reflect the precise meaning of the text. Therefore, while text feature representation and extraction are crucial steps for improving social commerce text classification performance, they require specific attention and modifications addressed in this paper.

3. PROPOSED MODEL

Based on the research mentioned in the previous section, to extract effective features for a more accurate examination and classification of social commerce texts, this paper proposes an algorithm based on bidirectional encoding using the ALBERT language model. This algorithm utilizes the pre-trained language model to represent feature vectors at the sentence level in social commerce texts, and then the obtained feature vector is used for classification in the model's classification unit [24].

The proposed classification model in this paper consists of four main sections: pre-processing of social commerce text datasets, pre-training the language model, fine-tuning the language model, and classification. Pre-processing social commerce texts includes pre-training using a suitable textual corpus, pre-processing the datasets, and vector representation of datasets in a specific format. After pre-processing, the data are input into the pre-trained language model, and finally, the classification process is performed.

3.1. Pre-processing and Datasets

This section includes the following:

- Data Collection: For the pre-training process of the ALBERT language model, fine-tuning the model, and training and evaluating the model, a suitable and relevant textual corpus is needed, which must be prepared according to the model's requirements (for more details on the corpus used in this paper, refer to Section 4).
- Data Cleaning and Sorting: This process mainly involves removing special symbols and characters from the text content, eliminating extra spaces and unifying the text, removing stop words, and deleting conversational words in the text that do not impact its meaning (such as "Hi" and "Hello").
- Text Segmentation: This process involves dividing a text into its meaningful units such as sentences and words.
- Stemming: This process involves determining the root of words to better perform natural language processing tasks.

3.2. Language Model

The ALBERT language model used in this paper is a deep learning-based language model derived from the BERT language model but with significantly fewer parameters. The BERT language model is based on a bidirectional encoder implementation, and its encoder unit comprises 6 encoders, while its decoding layer is identical. Each encoder in this model includes two sections: Self-Attention and Feed Forward, allowing each node to focus on the meaning of the current word and help determine the context of the text. Its decoding layer also includes three sections: Self-Attention, En-De Attention, and Feed Forward, enabling it to play a precise role in determining the current content [24 and 25].

The ALBERT language model employs two optimization processes to reduce the size of the BERT model: using embedding layers and sharing parameters in the network's hidden layers. Combining these two approaches results in a base model with only 12 million parameters compared to the 108 million parameters of the BERT model, while in several natural language processing tasks, the ALBERT model's accuracy averages 80.1%, compared to the BERT model's 82.3% [24 and 25].

In the model's usage process, the word embedding procedure is first performed on the input data, and the result is then input into the encoder layer. After the relevant processing in the encoding layer through the Self-Attention mechanism and neural networks, the output is input into the next encoding layer. The use of the Self-Attention mechanism addresses the slowness of recurrent neural network structures and their inability to parallelize. This effectively resolves the issue of long-term dependencies in natural language processing. The main idea of Self-Attention is to calculate each sentence in one step, adjusting the weight ratio of each word through this process so that the system can obtain the higher-dimensional meaning of each word in the sentence, thus obtaining the meaning of the entire sentence through this process. In the Self-Attention mechanism, three new vectors Q-Query, K-Key, and V-Value are used, which result from the dot product of the embedded word vector with a random matrix of initial dimensions 64×512. When a sentence is input, the status of each word relative to other words is calculated using the following formula:

$$Attention(Q, K, V) = \text{soft max}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

In such a way that d_k represents the dimensions of the query and key vectors of each word, and Softmax is the normalized exponential function. The resulting matrix has rows that indicate the attention vector of the respective word in the input sentence. This result contains information about the relationship between the target word and other words in the sentence, represented as a new vector [24].

3.3. Text Classification Algorithm

The general framework of the algorithm used in this paper for classifying social commerce texts using the ALBERT language model is presented in Figure 1.

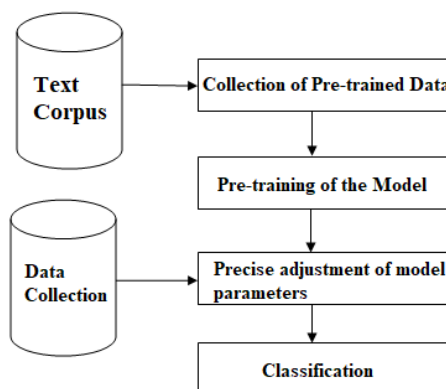


Fig.1. General Framework of the Proposed Model Algorithm

The steps of this algorithm are as follows:

- Pre-training of the Language Model: Initially, a suitable and relevant text corpus is provided. This corpus does not necessarily include social commerce texts and can comprise any type of text. Each data piece in this corpus is stored in a text document with labels, separators, and content. Due to the large size of such a corpus, it is typically divided into smaller parts to enable processing. The feature vector generation program is then invoked to create training data, which is ultimately used to pre-train the language model.

- Fine-tuning of the Language Model Parameters: This process requires datasets directly derived from social commerce texts. These data are also stored in a text document with labels, separators, and content. For final training and fine-tuning of the model parameters, 80% of these data are used for training, while the remaining 20% are reserved for final evaluation of the model.

- Classification Process: For classifying input texts, this paper defines five categories ranging from 1 (very poor or very weak) to 5 (very good or very strong). For classification, the word vectors related to the text content are fed into the classification algorithm. The algorithm assigns a category label to the text based on the trained language model. The accuracy of the classification process is then evaluated by comparing the assigned label with the original label present in the input dataset.

4. MODEL EXECUTION AND RESULTS

4.1. Datasets Used

As mentioned earlier, the model requires two types of datasets: a text corpus for pre-training the language model and a dataset for training and evaluating the final model. In this paper, the Kaggle and Yelp text classification datasets are used as the corpus for pre-training the language model. Additionally, the Amazon, IMDB, and Goodreads datasets are separately used for training and final evaluation of the model. Each of these datasets is independently applied to the model, and the results are analyzed and evaluated.

4.2. Execution Environment Details

The model execution and testing were performed on a system with a 3.7 GHz Intel Xeon Gold 6218 processor, 32 GB RAM, and an NVidia Quadro RTX4000 GPU with 8 GB of DDR6 RAM. The operating system used was Linux Ubuntu 16.04 LTS. Python version 3.9 and TensorFlow version 2.5 were used for programming. Additionally, the NLTK version 3 library was used for preprocessing and other tasks.

4.3. Evaluation Metrics

In evaluating the classification results, the target category is considered correct, while other categories are considered incorrect. The classifier's prediction result on the dataset is defined as correct or incorrect, leading to the formation of a confusion matrix, shown in Table 1.

Table 1. Confusion Matrix of Classification Results

Predicted Class \ Real Class	True	False
	Positive	True Positive (TP)
Negative	False Positive (FP)	True Negative (TN)

The commonly used classification metrics are accuracy, precision, recall, and F1 score, which are also used in this paper. These metrics are calculated as follows:

- Accuracy Rate: The ratio of all correctly predicted values to the total number of cases is calculated as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

- Precision Rate: The ratio of correctly predicted positive cases to all predicted positive cases is calculated as:

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

- Recall Rate: The ratio of correctly predicted positive cases to all actual positive cases is calculated as:

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

- F1 Score: To evaluate the overall aspects of different algorithms, the F1 score concept is proposed based on precision and recall. The F1 score combines precision and recall and is calculated as:

$$F_1 - Score = \frac{2Precision \times Recall}{Precision + Recall} \tag{5}$$

By examining the above formulas, it can be understood that when Precision = 1 and Recall = 1, the F1 score reaches its maximum value of 1. In real-world conditions, achieving a situation where both precision and recall equal 1 is challenging. Thus, when using the F1 score to evaluate the classifier's performance, the closer its value is to 1, the better the classifier's performance.

4.4. Results from Model Execution

The results of executing the model on the introduced training and testing datasets in section 4-1 are presented in Table 2.

Table 2. Results from Model Execution on Reference Datasets

Index \ Data Collection	Accuracy	Precision	Recall	F1 Score
Amazon	0.740	0.749	0.722	0.735
IMDB	0.697	0.701	0.689	0.694
Goodreads	0.722	0.726	0.713	0.719

5. ANALYSIS OF RESULTS

By reviewing the model results, it can be observed that the proposed model has adequate accuracy in classifying social commerce texts. This demonstrates the ALBERT language model's effective impact on extracting meanings.

Moreover, considering that the highest data volume is found in the Amazon, Goodreads, and IMDB datasets, the F1 score across these datasets indicates that the greater the data volume in the dataset, the higher the model's accuracy. This finding suggests that the social commerce text classification algorithm based on the ALBERT language model significantly impacts the text classification process in this field.

6. CONCLUSION

This paper presented a model for classifying social commerce texts using the ALBERT language model. For social commerce text data, based on related datasets, the ALBERT language model's bidirectional transformer structure was utilized to derive sentence meanings and subsequently classify them. The results from executing the model on different datasets demonstrate that the proposed model has suitable quality in classifying the desired texts. However, if specific datasets for this model are prepared based on actual data, the quality can be better assessed. Additionally, evaluating the model's quality when increasing the number of target categories can be particularly important. Alongside these two points, which can serve as appropriate subjects for further research, it should be noted that text classification is only one aspect of extracting value from social commerce big data. Future work can utilize sentiment analysis and entity recognition to enhance model quality and apply this model to automatically generate social commerce advertisements.

Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

Declaration of Interest

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

REFERENCES

- [1] Sheikh, Z., Liu, Y., Islam, T., Hameed, Z., & Khan, I. U. (2019). Impact of social commerce constructs and social support on social commerce intentions. *Information Technology & People*. <https://doi.org/10.1108/ITP-04-2018-0195>
- [2] Bugshan, H., & Attar, R. W. (2020). Social commerce information sharing and their impact on consumers. *Technological Forecasting and Social Change*, 153, 119875. <https://doi.org/10.1016/j.techfore.2019.119875>
- [3] Nakayama, M., & Wan, Y. (2019). The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews. *Information & Management*, 56(2), 271-279. <https://doi.org/10.1016/j.im.2018.09.004>
- [4] Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune BERT for text classification?. In *China National Conference on Chinese Computational Linguistics* (pp. 194-206). Springer, Cham. https://doi.org/10.1007/978-3-030-32381-3_16
- [5] Iyyer, M., Manjunatha, V., Boyd-Graber, J., & Daumé III, H. (2015). Deep unordered composition rivals syntactic methods for text classification. In *Proceedings of the 53rd Annual Meeting of the Association for*

Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) (pp. 1681-1691). <https://doi.org/10.3115/v1/P15-1162>

- [6] Cheng, J., Dong, L., & Lapata, M. (2016). Long short-term memory-networks for machine reading. arXiv preprint arXiv:1601.06733. <https://doi.org/10.18653/v1/D16-1053>
- [7] Liu, P., Qiu, X., Chen, X., Wu, S., & Huang, X. J. (2015). Multi-timescale long short-term memory neural network for modelling sentences and documents. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (pp. 2326-2335). <https://doi.org/10.18653/v1/D15-1280>
- [8] Le, H. T., Cerisara, C., & Denis, A. (2018). Do convolutional networks need to be deep for text classification?. In Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence.
- [9] Hughes, M., Li, I., Kotoulas, S., & Suzumura, T. (2017). Medical text classification using convolutional neural networks. In Informatics for Health: Connected Citizen-Led Wellness and Population Health (pp. 246-250). IOS Press.
- [10] Zhao, W., Ye, J., Yang, M., Lei, Z., Zhang, S., & Zhao, Z. (2018). Investigating capsule networks with dynamic routing for text classification. arXiv preprint arXiv:1804.00538.
- [11] Yang, M., Zhao, W., Chen, L., Qu, Q., Zhao, Z., & Shen, Y. (2019). Investigating the transferring capability of capsule networks for text classification. *Neural Networks*, 118, 247-261. <https://doi.org/10.1016/j.neunet.2019.06.014>
- [12] Yamada, I., & Shindo, H. (2019). Neural attentive bag-of-entities model for text classification. arXiv preprint arXiv:1909.01259. <https://doi.org/10.18653/v1/K19-1052>
- [13] Xiong, C., Merity, S., & Socher, R. (2016). Dynamic memory networks for visual and textual question answering. In International Conference on Machine Learning (pp. 2397-2406). PMLR.
- [14] Peng, H., Li, J., Wang, S., Wang, L., Gong, Q., Yang, R., ... & He, L. (2019). Hierarchical taxonomy-aware and attentional graph capsule RCNNs for large-scale multi-label text classification. *IEEE Transactions on Knowledge and Data Engineering*.
- [15] Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using siamese BERT-networks. arXiv preprint arXiv:1908.10084. <https://doi.org/10.18653/v1/D19-1410>
- [16] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8), 9.
- [17] Bao, H., Dong, L., Wei, F., Wang, W., Yang, N., Liu, X., ... & Li, M. (2020). UniLMv2: Pseudo-masked language models for unified language model pre-training. In International Conference on Machine Learning (pp. 642-652). PMLR.
- [18] Chen, J., Yang, Z., & Yang, D. (2020). MixText: Linguistically-informed interpolation of hidden space for semi-supervised text classification. arXiv preprint arXiv:2004.12239. <https://doi.org/10.18653/v1/2020.acl-main.194>
- [19] Liu, X., Mou, L., Cui, H., Lu, Z., & Song, S. (2020). Finding decision jumps in text classification. *Neurocomputing*, 371, 177-187. <https://doi.org/10.1016/j.neucom.2019.08.082>
- [20] Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. arXiv preprint arXiv:1802.05365. <https://doi.org/10.18653/v1/N18-1202>

- [21] Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training.
- [22] Wu, F., & Zheng, Y. (2016). Adaptive normalized weighted KNN text classification based on PSO. *Scientific Bulletin of the National Mining University*, 1, 109-115.
- [23] Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532-1543). <https://doi.org/10.3115/v1/D14-1162>
- [24] Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2019). ALBERT: A lite BERT for self-supervised learning of language representations. *arXiv preprint arXiv:1909.11942*.
- [25] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.