Bima Journal of Science and Technology, Vol. 7 (1) Mar, 2023 ISSN: 2536-6041

DOI: 10.56892/bima.v7i01.394

ENHANCED DEEP NEURAL NETWORK FOR ASPECT-OPINION TERMS CO-EXTRACTION

*AMINU DAU

Department of Computer Science, Hassan Usman Katsina Polytechnic, Katsina State Nigeria.

Corresponding Author: aminu.dau@hukpoly.edu.ng

ABSTRACT

Aspect-Opinion Co-extractions are two important sub-tasks of aspect-based sentiments analysis (ABSA) that involve simultaneous identification of product's aspects and the associated opinion words from user textual reviews. Traditional approaches to the aspect-opinion co-extraction typically depend on the handcrafted and rule-based methods which are known to be labor-intensive and often less accurate. Recently, Convolutional Neural Networks (CNNs) have been widely applied for the aspect-opinion co-extraction task. However, the existing approaches rely solely on the word embedding models such as Glove or word2vec. As such they cannot guarantee more fine-grained semantic information due to the proble of the *distributional hypothesis* . Thus, in this paper, wet propose a lexicalized CNN (LCNN) technique that can help to better capture the fine-grained semantic information for better coextraction process. The proposed method consists of lexicon embeddings in addition to the word embeddings as inputs to the network. For the word embedding input, we use general embedding (GE) which is pre-trained based on a large corpus of Google news and domain-specific embedding (DSE) which is trained based on the Amazon and Yelp reviews. For the lexicon embeddings, we use lexicon resources based on the SenticNet model. The word embedding and lexicon embedding are concatenated and fed into the convolutional network to generate local features which are then max-pooled to generate input to the softmax function for the final aspect-opinion co-extraction task. The proposed model was evaluated using various benchmark datasets and the experimental results have shown that our proposed model performed better than to the baseline approaches.

Keywords: Aspect extraction, Convolutional Neural Network, Word embedding, Lexicon Integration; Artificial Neural Network

INTRODUCTION

Aspect-opinion terms Co-extraction can be carried out using either supervised or unsupervised methods (Z. Chen, Huang, Liu, Shi, & Jin, 2021). For several years, the state of the art supervised methods essentially rely on Conditional Random Fields (CRFs) (Lafferty, Mccallum, Pereira, & Pereira, 2001) or Recurrent Neural Networks (RNNs) (Irsoy & Cardie, 2014; Y. Liu, Wang, & Wang, 2018). Both of these methods have their intrinsic drawbacks. For instance, the CRF-based methods generally require a large number of features to function due to their linearity in nature, the RNNs are very slow to train due to their sequentially dependent cells. The unsupervised methods such as linguistic patterns or syntactic rules (Guang & Bing, 2009; Popescu & Etzioni, 2005) need to be hand-crafted and their performances largely depend on the grammatical structure of the sentence. Recently, some approaches have been proposed to exploits CNN architecture to enhance the accuracy of the models (Poria, Cambria, & Gelbukh, 2016b; Xu, Liu, Shu, & Yu, 2018b). However, most of the existing

DOI: 10.56892/bima.v7i01.394

CNN based methods rely solely on the word embedding techniques such as Glove (Pennington, Socher, & Manning, 2014) or Word2vec (Mikolov, Yih, & Zweig, 2013) as the main semantic features. Even though word embeddings have been indicated to be effective in better learning both semantic and syntactic features of texts. However, due to the distributional hypothesis, word embedding alone cannot guarantee to learn more semantic information of some aspect words. For instance, *"good" and "bad"* are particularly mapped together as neighbours in a latent space while analyzing these words is very critical in real-world applications. Thus, this paper proposes an approach for the aspect-opinion co-extraction using a lexicalized CNN (LCNN) architecture.

Although some approaches have been previously proposed to integrate lexicon resources into the CNN model for the natural language processing tasks (NLP) to some extent (Kunaver & Požrl, 2017; Lan, Zhang, Lu, & Wu, 2016; Shin, Lee, & Choi, 2016; Zhang, Zou, & Gan, 2017). However, the existing lexicon integrated CNN methods particularly focus on the sentence level analysis while little attention has been given to the aspect-opinion words level analysis which is totally different from the sentence classification (analysis). Moreover, the existing methods generally focus on the general pre-trained embeddings (GE) and particularly ignore the domain-specific embeddings (DSE). However, due to the complexity of the NLP tasks such as aspect/opinion terms extraction, we argue that utilizing DSE which is essentially fine-grained embeddings are very crucial for capturing the domain-specific semantic information of the text thereby improving the model performances The proposed model comprised of the multiple inputs to the convolutional layer, namely, word embedding and lexicon embedding. For the word embeddings specifically, we utilize both the general embedding (GE) and the domain-specific embeddings (DE) while for the lexicon embeddings particularly we utilized Sentic lexicon resource (Cambria, Olsher, & Rajagopal, 2014). Following the input layer, is the convolutional layer, pooling layer, and finally the output layer with fully connected SoftMax for the final extraction task. The model was evaluated using different benchmark datasets and the experimental results showed better performances of our proposed model compared to the baseline approaches. The major contributions of the proposed approach include the following:

- We introduced a lexicalized CNN technique for aspect-opinion coextraction based on the user textual reviews leveraging word embeddings and lexicon embeddings.
- We design an approach to investigate the impact of utilizing two different input layers for the CNN network.
- We carry out a series of experiments on the benchmark data sets and the results demonstrated that our proposed method outperforms the baseline methods with significant improvements.

The paper proceeds as follows: Section2, reports the related work, section 3 and section 4 present an overview of the proposed model and experimental study respectively. section 5 and section 6 reports the results and discussion, and the conclusion of the article respectively

RELATED WORK

Aspect and opinion terms extraction have recently become a fast-growing research area in both academia and industries. Several works have been carried out for the aspect extraction, with the frequency-based methods being the earliest approaches (Hu & Liu, 2004; Popescu & Etzioni, 2005) (S. Chen, Wang, Liu, & Wang, 2021). In these methods,

DOI: 10.56892/bima.v7i01.394

some specified constraints are applied for identifying the most frequent nouns or noun phrases in reviews as the aspects candidates. These methods consider only the most popular aspect/opinion words while the low-frequency aspect/opinion words are usually neglected. To address the issue with the frequency-based methods, Poria et al., (Poria, Cambria, Ku, Gui, & Gelbukh, 2014) introduced a rule-based method that exploits linguistic patterns for better performance. Recently, some approaches have been introduced to employ topic modeling based methods (García-Pablos, Cuadros, & Rigau, 2018; Mei, Ling, Wondra, Su, & Zhai, 2007) , with the most widely used model being LDA technique (Blei, Ng, & Jordan, 2003). Topic models basically use the latent topics between variable words and documents, whereby each document comprises a random mix of topics. These methods are limited in that they cannot effectively capture the fine-grained aspects.

With the recent achievement of the artificial neural network in NLP (Da'u & Salim, 2019; Kim, 2014) (Huang et al., 2021), several methods have been introduced to leverage deep learning techniques for the aspect/opinion extraction task. Most of these methods rely on the CNN or RNN models. For example, Poria et al., (Poria, Cambria, & Gelbukh, 2016a) apply a multilayer convolutional model for aspect extraction by tagging each word as an aspect or non-aspect label. To further improve the model performance, the authors additionally applied linguistic features which are then integrated with the pre-trained vectors. Pham (Pham & Le, 2018) proposed a CNN based technique by exploiting multiple embeddings for aspect detection. The model specifically integrates Word2vec, Glove and one-hot-vector to generate a unified feature generation for a better extraction process. The model is capable to learn the shared representation using different CNN units which are then jointly trained with the same objective function. Authors in ref (Xu, Liu, Shu, & Yu, 2018a) introduced a simple CNN based technique named DE-CNN that leverage double embeddings for the aspect extraction. The model uses pre-trained Glove and a domain-dependent embedding that are trained on the Amazon and Yelp reviews using FasText method.

Following their huge success in sequential modeling and word dependency support, the RNN models have been utilized for aspect/opinion extraction tasks. Authors in ref (Jebbara & Cimiano, 2016b) exploited a bidirectional GRU model for the aspect-opinion coextraction. Specifically, the model uses the GRU model to extract product aspects from the user textual review based on the *IOB* sequential labelling. In the second stage, the GRU model is further used to predict the user sentiments associated with the aspects. Authors in ref (Irsoy & Cardie, 2014) uses the Elman-type RNN model to better extract opinion expression by exploiting Google pre-trained word embeddings. The authors demonstrated the power of the model over the shallow RNN and the CRF method. Chen et al., (2017) introduced a bidirectional LSTM based approach for aspect extraction. The authors integrated the LSTM model with the CRF to better classify the number of targets in the sentence for better accuracy. The model is capable to capture the dependencies of words in sentences thereby enhancing the performances of the model. Tran et al., (Tran, Hoang, & Huynh, 2019) proposed an RNN based method for the aspect extraction by integrating BiGRU and CRF techniques. The model particularly uses the embedding layer consisting of GloVe, and the CRF layer to finally predict the aspect term labels. An end to end method called BiDTreeCRF was proposed in (Luo, Li, Liu, Wang, & Unger, 2018) for the aspect extraction. The model is capable to effectively

DOI: 10.56892/bima.v7i01.394

extract syntactic dependencies using top-down and bottom-up propagation in dependency trees. Wang et al., (Wang, Pan, Dahlmeier, & Xiao, 2016b) proposed a recursive neural network model for aspect and opinion co-extraction. The model leverages a dependency-based method that simultaneously uses the CRF and RNN techniques in addition to the handcrafted features. Wang et al. (Wang, Pan, Dahlmeier, & Xiao, 2016a) used the GRU method for aspect extraction. Specifically, the authors design a couple of multi-layer attention networks (CMLA) based on the GRU to simultaneously extract the product aspect and the associated opinion. Thus, learning can be achieved through encoding and decoding the dual propagation of aspects and the associated opinion as well as the constraint to the grammatical relations. Although the above methods have demonstrated good performance compared to their prior approaches, yet, most of these methods solely rely on the word embeddings as the main semantic feature and fail to additionally exploit the lexicon features for improving the model performance. Even though, some approaches have been proposed to integrate lexicons into the CNN methods (Lan et al., 2016; Shin et al., 2016; Zhang et al., 2017) for the NLP task. However, different from our proposed model, these methods particularly focus on sentence-level analysis. In our proposed method, we particularly focus on the aspect-opinion words level analysis.

MATERIALS AND METHODS

In this section, the detailed description of the proposed methods is presented including the problem definition, an overview of the proposed model, and the different components of the model.

Problem Definition

Assuming we have a training set of text review in a particular domain, given by D = $\{d_1, d_2 \ldots . d_n\}$, where n is the number$s$ of sentences in the reviews. For any $d_i \in$ D, there may exist a set of aspect$s$ terms $A = \{a_1, a_2 \ldots . a_m\}$, where each $a_j \in A$ can be a single or a sequence of words referring to some aspect of an item. The ultimate goal here is to train a classifier to identify a set of aspect terms A_j and the associated opinions from each sentence $d_i \in$ D in the user textual review for a particular domain.

Similar to other sequential tagging problems (Jebbara & Cimiano, 2016a; Poria et al., 2016a), this problem can be regarded as a tagging task based on the BIO encoding method where each review sentence d_i is assumed to be comprised of a sequence of words $d_i = \{w_{i1}, w_{i2} \ldots . w_{ir}\}$, and each word $w_{ip} \in d_i$ is annotated as one the following categories: BA, IA, BO, IO, O corresponding to the "*Beginning of the Aspect*", "*Inside of the of Aspect*", "*Beginning of the Opinion*", " *Inside of the of Opinion*" and *"Others*" respectively.

Proposed Model

The proposed leverages word embeddings and lexicon resources for better aspect-opinion co-extraction. The proposed model is an extension of the CNN structure for text modelling introduced in (Kim, 2014). It is made up of multiple input channels, convolution layer, pooling layer, and a fully connected layer with SoftMax function. Figure 1 demonstrates an illustration of the proposed LCNN approach. The detail description of the proposed method is given as follows:

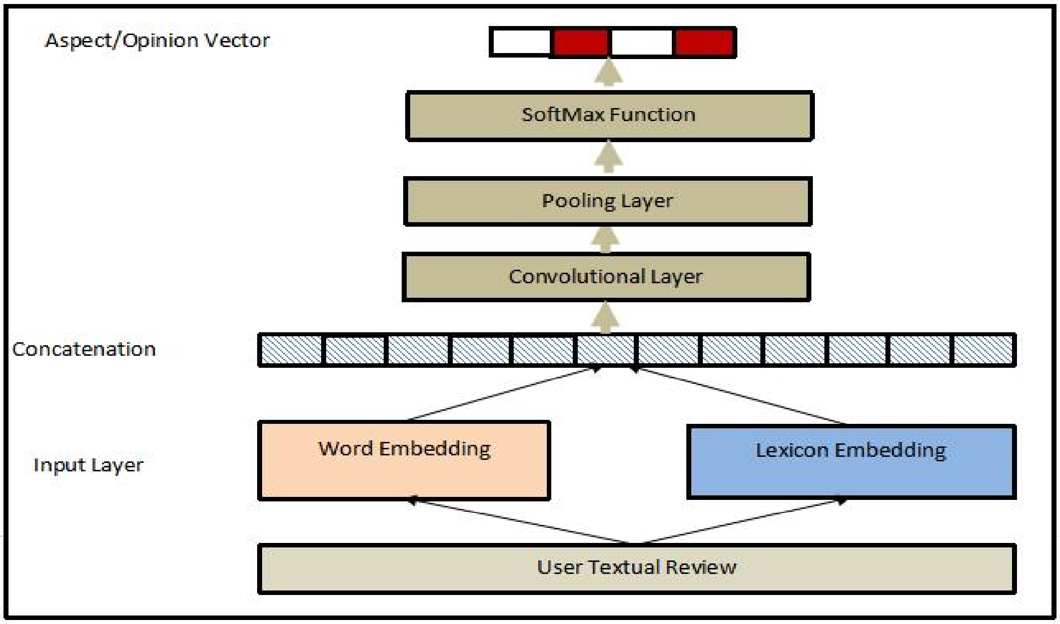

Figure 1: the framework of the proposed model

Input: The input of the proposed LCNN comprises of two sets of vectors: word embedding and lexicon embedding. The word embedding is aimed to better enhance learning the semantic features of the text. To achieve that, specifically, two different word embedding techniques are utilized: General word embedding (GE) (Mikolov et al., 2013) and Domain-specific embedding (DE). More formally, for capturing the semantic information of a sentence i-th word is mapped to a k −dimensional embedding by a Lookup Table as:

$$x_i = LT(W)^i.$$

Where x_i is the $i-th$ row vector of word embedding matrix, W is the semantic feature of a sentence of length n, this is represented as concatenating all its word embeddings orderly. Formally this can be represented as:

$$X_i = [x_1 \ldots x_n], with \ X_i \in R^{n \times k}. \quad (1)$$

For the lexicon embedding, a five-dimensional vector is produced based on the $Sentic$ 3 (Cambria et al., 2014). $SenticNet$ 3 is a graph-based resource for semantic information which offers real-valued scores for some specified Sentics. It involves concept level resource for affective and semantic information. For each of the concept that are part of the knowledge graph, $SenticNet$ 3 provides real-valued scores for five different Sentics (attention, pleasantness, sensitivity, polarity, aptitude). Due to the space limitation, more detail about the $Sentic$ 3 can be found in (Cambria et al., 2014). Thus, for each concept that is represented in $SenticNet$ 3, a d-dimensional feature vector is generated as:

$$Q_j = [q_1 \ldots\ldots, q_m], with \ Q_j \in R^d \quad (2)$$

Where m is the number of words in the sentence and d is the dimension of the vector. After that, both the word embeddings and the lexicon embedding channels are concatenated to generate local features which are then fed into the convolutional layer. More formally the input document matrix can be represented as:

$$S = [X_i \oplus Q_j] \in R^{n(K+d)} \quad (3)$$

Where X_i and Q_j represent the word embedding and lexicon embedding vectors respectively. $\oplus$ is the concatenation operator, n is the number of words, $d \, and \, k$ are the dimensions of the word embedding and lexicon embedding respectively.

Convolution Layer: with the concatenated word embedding and lexicon embedding inputs, the convolution operation is applied to generate the global features from the local features. More formally a convolutional operation produces a feature c_i from a window of words $x_{i:\,i+h-1}$ as:

$$C_i = ReLU(W.\, x_{i:\,i+h-1} + b) \quad (4)$$

Where $ReLU$ is an activation function, $b \in R$ is a bias term. This filter is used to each possible windows of words in the sentence to produce a feature map:

$$c = [\, c_1, c_2 \ldots . c_{n-h+1}] \in R^{n-h+`1} \quad (5)$$

Pooling Layer : after the convolutional operation, the maximum values from different filters are taken as the most informative features. Formally the generated features can be given as $\hat{c} = [\hat{c}_1, \hat{c}_2 \ldots . \hat{c}_n]$ (6)

Output Layer: the generated pooled features are finally used as the input to the fully connected Soft max for the final extraction tasks. The output vector is computed as follows:

$$a^{'} = Sofmax(U^0 . C + b^o) \quad (7)$$

Where, $U^0 \in R^{k \times P}$ is a weight matrix, and $b^0 \in R^k$ is a biased vector and are the parameters to learn.

EXPERIMENTS

To better validate the proposed method, three different datasets are utilized. The first two datasets and the third datasets are from SemEval2014 (Pontiki & Pavlopoulos, 2014) and semeval2015 (Maria, Dimitrios, Haris, & Suresh, 2015) consisting of reviews from the restaurant and laptop domains respectively.

It should be noted that the original datasets only comprise labels for the aspect terms. Thus to better evaluate the performance of our model for both aspect and opinion terms extraction, we utilize labels on the opinion terms provided in (Wang et al., 2016a) and (Wang et al., 2016b) accordingly. The statistics of the datasets are given in Table 1.

Table 1: summary of the datasets

Datasets	SemEval 14-R	SemEval 14-L	SemEval 15-R
Train	3,041	3,045	1,315
Test	800	800	685
Total	**3,841**	**3,845**	**2,000**

To get clean reviews, all the datasets are lowercased and then split into separate sentences. All stop words, special and alphanumeric characters are filtered. For the word embeddings, specifically, Google word2vec (Mikolov et al., 2013) and the Domain-specific embeddings (DSE), are used. The DSE is trained based on the CBOW using Gensim on the Yelp and Amazon reviews for the $restaurant$ and the $laptop$ domains respectively. For the SecticNet, since our proposed method focuses on the word level analysis and not meant for the concept level analysis, all multi-word concepts in the $Sentic$ 3 are removed and single word concepts that are part of vocabulary are reserved. 5-folds cross-validation on the training datasets is used for choosing the hyper-parameters. To better set up the model hyperparameters, particularly, three filter sizes of $(3, 4, 5), with$ 100 feature maps were chosen.

DOI: 10.56892/bima.v7i01.394

To handle the overfitting, dropout regularization with $L2$ constraints of 3 with the dropout rate of 0.5 was used. The model is trained with a stochastic gradient descendent (SGD) with a learning rate of 0.0001 and the mini-batches of 64. All the values were decided according to the grid search algorithm on the validation sets.

Comparative Methods

To better investigate the effectiveness of our proposed method we use different states of the art models for comparison. These include 1) WDEmbd (Yin et al., 2016): a dependency-based method for the aspect extraction that uses word embedding. 2) IHS_RD(Chernyshevich, 2014), 3) DLIREC (Toh & Wang, 2014): the two top systems for SemEval-14 Laptop and SemEval-14 Restaurant domain respectively. 5) CNN+P (Poria et al., 2016a): a deep learning-based method that uses the CNN method for sequence labeling 6) LSTM (P. Liu, Joty, & Meng, 2015): an RNN based method that uses different versions of the RNN. 7) RNCF (Wang et al., 2016b): a dependency-based method that simultaneously uses CRF and RNN methods. 8) CML (Wang et al., 2016a): a multilayer coupled-attentive network that for aspect/opinion terms coextraction. 9) SpanMLT [39] (Zhao, Huang, Zhang, Lu, & Xue, 2020) is a shared span-based learning technique that uses span boundaries for extractin relations between aspect and opinion pairs.

RESULTS AND DISCUSSION

Table 2 demonstrates the performance of the proposed method compared with the state-of-the-art baselines. The results are reported for both aspect term (AT) and the opinion terms (OT) extraction in terms of the F1 score accuracy across all the datasets. It can be seen that the LCNN approach significantly outperforms all the baselines across all the datasets and the improvements are statistically significant based on the t-test at $p<0.05$.

Table 2: Comparison performances with the baseline methods in terms of F1 score (%)

Model	SemEval-14 L		SemEval-14 R		SemEval –15 R	
	AT	OT	AT	OT	AT	OT
WDEmb	75.16	-	84.97	-	69.73	-
DLIREC	73.78	-	84.01	-	-	-
HIS_RD	74.55	-	79.62	-	-	-
CNN+P	82.26	-	87.17	-	-	-
LSTM	72.73	74.98	81.15	80.22	64.30	66.43
CMLA	77.80	80.17	85.29	83.18	70.73	73.68
RNCF	78.42	79.44	84.93	84.11	67.74	67.62
SpanMLT	77.87	80.51	85.24	85.79	71.07	75.02
LCNN	**84.05**	**84.37**	**88.43**	**85.34**	**75.92**	**75.05**

Compared to the WDEmb method which is a deep learning-based method and particularly exploits word embeddings in addition to the word dependencies, our proposed model achieves improvements of 11.83 %, 4.07%, and 8.88% on the SemeEval-014 L, SemeEval-014 R and SemeEval-015 R datasets respectively, in terms of the F1 score accuracy. Similarly, compared to DLIREC and HIS_RD approaches which are the winning models in the SemEval-14, in the laptop and restaurant domain respectively, our model performs better with improvements of 13.92%, 5.26 % and 12.74%, 11.07% on the restaurant and laptop domain respectively. It can also be shown that our model outperforms the RNCF model which is the CRF base method, with gains of 7.18%, 4.12, 12.08, and 6.21%, 1.46%, 11.0% on the SemeEval-014 L, SemeEval-014 R and SemeEval-015 R

DOI: 10.56892/bima.v7i01.394

datasets for the aspect and opinion terms extraction respectively. Similarly, compared to CML model which typically uses multi-layered attention networks, our model outperformed it with gains of 8.03%, 5.24%, 3.68% and, 2.60 %, 7.34%, 1.86% on the SemeEval-14 L, SemeEval-14 R and SemeEval-15 R datasets for the aspect and opinion terms extraction respectively. More over, the CNN+P which is also a CNN based deep learning approach utilizing heuristic and rule-based features with addition to the linguistic features, still our proposed approach outperforms the model with improvements of 2.18% and 1.45% on the SemEval-14 and SemEval-14 R respectively. More importantly, the recently published aspect-opinion term coextraction extraction which is based on shared span learning method, still our proposed model outperfromed it with a significant improvement.

Model Ablation

To examine the performance of different components of the proposed approach, we design four different settings of the model as follows:

- LCNN+GE: which uses the General Embedding (GE) as the only input and ignore the lexicon embeddings. This is used to examine the influence of the lexicon features on the model performances.
- LCNN+DSE: which uses the domain-specific embedding (DSE) as the input to the network and ignore the lexicon embedding. This setting is used to investigate the impact of the DSE in the model performance.
- LCNN+GE+LX: which utilizes both the GE and lexicon embedding. This setting is particularly used to assess the impact of the lexicon resources in addition to the general word embedding.
- LCNN+DSE+LX: which uses lexicon resources in addition to the DSE embedding and used to examine the influence of the lexicon with the domain-specific word embeddings.

Table 1: Performance of the various settings of the LCNN in terms of F1 scores (%)

Model	**SemEval-14 L**		**SemEval-14 R**		**SemEval -15 R**	
	AT	**OT**	**AT**	**OT**	**AT**	**OT**
LCNN+GE	80.67	81.15	84.55	83.87	68.98	69.28
LCNN+DSE	82.01	82.81	85.68	85.04	70.05	70.37
LCNN+GE+Lx	82.62	83.05	87.25	86.21	71.76	72.23
LCNN+DSE+Lx	**84.05**	**84.37**	**88.43**	**85.34**	**75.92**	**75.05**

The experiment results of the various settings of the approach are recorded in Table 3 in terms of the F1 accuracy in each dataset. From Table 3 one can see that different results are reported on different variants across various datasets. For example, the LCNN+DSE+LX variant outperforms all other versions of the model while the LCNN+GE performs relatively worst among all the versions in all the cases. As can be observed from Table 3, all the versions of the proposed method except for LCNN+GE, show relatively competitive performances across different datasets. This simply shows the influence of domain-specific embeddings (DSE). Also, as can be seen from Table 3, all the variants of the proposed model which exploit lexicon resources show relatively competitive performances compared to other versions that rely solely on the word embedding alone, across all the datasets. This particularly demonstrates the advantage of integrating the lexicon feature in addition to either of the general embedding (GE) and domain-specific embedding (DSE) compared to the method that purely relies on the word embeddings.

For more analysis, we also investigate the sensitivity of the word embedding dimension with regards to the model performance in terms of the F1 score accuracy. To this end, we explore the different embedding dimensions from 25 to 375 with an interval of 25. Figure 2 and 3 demonstrates the experimental results on the different variants of the proposed method for the aspect/opinion terms extraction respectively. It can be seen that the best accuracy is typically within the range of 300 and 150. This essentially shows that the model is essentially sensitive to the word embedding dimensions provided it is within the appropriate range.

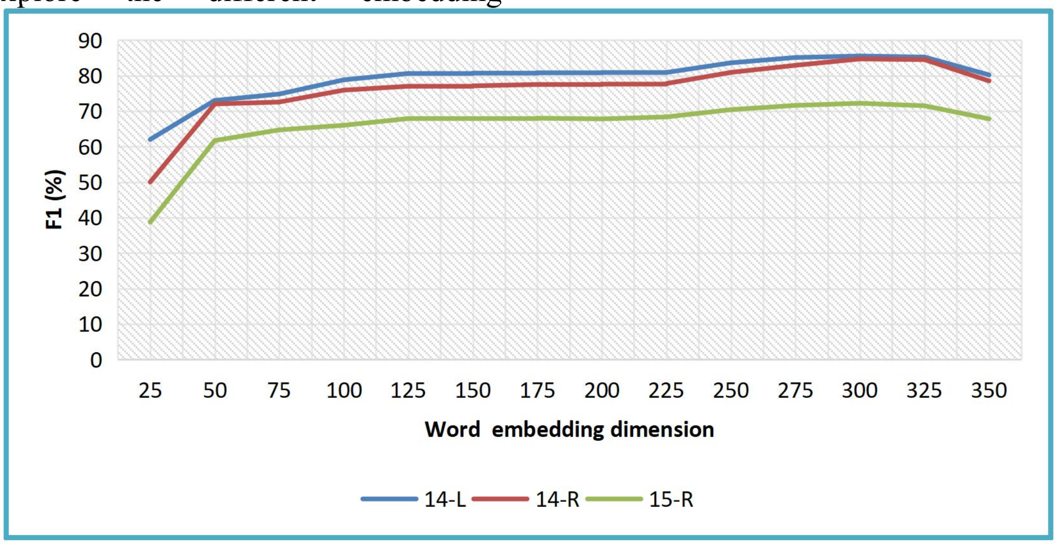

Figure 2: Model sensitivity to the word embedding dimension on the aspect extraction

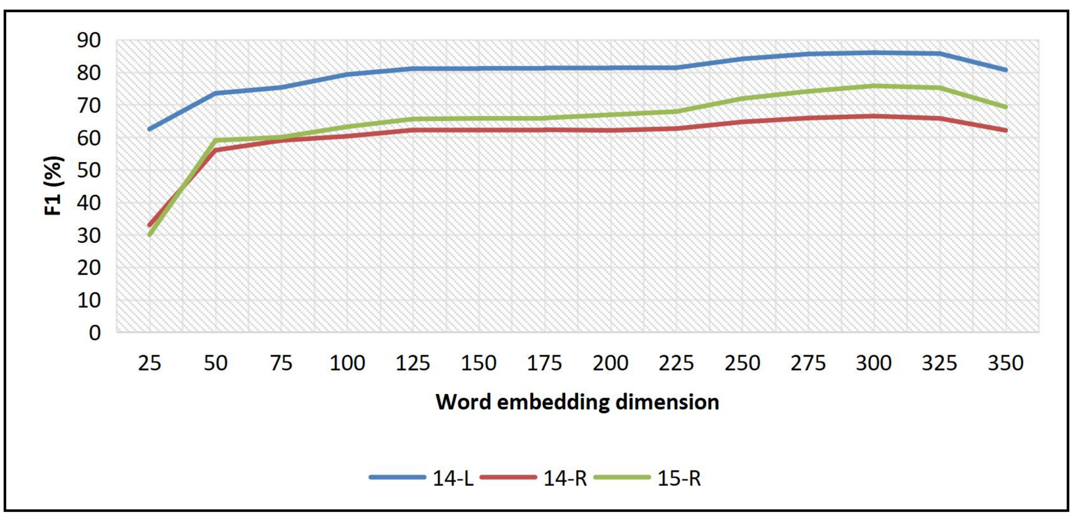

Figure 3: Model sensitivity to the word embedding dimension on the opinion terms extraction

An important advantage of our proposed approach is that it utilizes both the word embedding and lexicon resource as the input to the CNN method. The enable the approach to better capture the semantic and syntactic

DOI: 10.56892/bima.v7i01.394

features thereby performing better than the baseline methods.

CONCLUSION

This research proposed a lexicalized convolutional neural network method for the aspect-opinion coextraction. The proposed approach utilized multiple inputs to the CNN, namely, word embedding and lexicon embeddings. For the word embedding channel, general word embedding and domain-specific embedding are utilized. For the lexicon embedding input channel, lexicon resource from Sentic is utilized. The word embedding l and the lexicon embedding channels are concatenated and fed to the convolutional layer which is then pooled to generate global features as the input to the softmax max function for the final extraction task. The proposed model was evaluated using various datasets and the experimental study has demonstrated the proposed methods outperform the baselines. Our experimental results reaffirm many of the previous results which show that leveraging lexicon information could enhance the accuracy of the text processing tasks including aspect extraction.

REFERENCES

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research 3 (2003) 993-1022*, *3*(4–5), 993–1022.

Cambria, E., Olsher, D., & Rajagopal, D. (2014). SenticNet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis. *Proceedings of the National Conference on Artificial Intelligence*, *2*, 1515–1521.

Chen, S., Wang, Y., Liu, J., & Wang, Y. (2021). Bidirectional machine reading comprehension for aspect sentiment triplet extraction. *Proceedings Of The AAAI Conference On Artificial Intelligence*, *35*(14), 12666–12674.

Chen, T., Xu, R., He, Y., & Wang, X. (2017). Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Systems with Applications*, *72*, 221–230.

Chen, Z., Huang, H., Liu, B., Shi, X., & Jin, H. (2021). Semantic and syntactic enhanced aspect sentiment triplet extraction. *ArXiv Preprint ArXiv:2106.03315*.

Chernyshevich, M. (2014). IHS RandD Belarus: Cross-domain extraction of product features using CRF. *SemEval*, 309–313.

Da'u, A., & Salim, N. (2019). Aspect extraction on user textual reviews using multi-channel convolutional neural network. *PeerJ Computer Science*, *2019*(5), 0–16.

García-Pablos, A., Cuadros, M., & Rigau, G. (2018). W2VLDA: Almost unsupervised system for Aspect Based Sentiment Analysis. *Expert Systems with Applications*, *91*, 127–137.

Guang, Q., & Bing, L. (2009). Opinion Word Expansion and Target Extraction through Double Propagation. *Computational Linguistics*, *37*(1), 1–19. Retrieved from https://scholar.google.com.br/scholar?start=220&q=tudonotítulo:+Extraction+Emotion+OR+Feeling+OR+Sentiment+OR+Opinion+OR+Personality+OR+Subjectivity&hl=pt-BR&as_sdt=0,5&as_ylo=2005&as_yhi=2015#11

Hu, M., & Liu, B. (2004). Mining opinion features in customer reviews. *Proceedings of the National Conference on Artificial Intelligence*, 755–760.

Huang, L., Wang, P., Li, S., Liu, T., Zhang, X., Cheng, Z., … Wang, H. (2021).

DOI: 10.56892/bima.v7i01.394

First target and opinion then polarity: Enhancing target-opinion correlation for aspect sentiment triplet extraction. *ArXiv Preprint ArXiv:2102.08549*.

Irsoy, O., & Cardie, C. (2014). Opinion mining with deep recurrent neural networks. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 720–728.

Jebbara, S., & Cimiano, P. (2016a). Aspect-Based Relational Sentiment Analysis Using a Stacked Neural Network Architecture. *On Artificial Intelligence, 29 August-2*, 1–9.

Jebbara, S., & Cimiano, P. (2016b). Aspect-Based Sentiment Analysis Using a Two-Step Neural Network Architecture. *Semantic Web Evaluation Challenge*.

Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. *Proceedings of the 2014 Emperical Methods in Natural Language Processing (EMNLP)*, 23–31.

Kunaver, M., & Požrl, T. (2017). Diversity in recommender systems – A survey. *Knowledge-Based Systems*, *123*, 154–162.

Lafferty, J., Mccallum, A., Pereira, F. C. N., & Pereira, F. (2001). Conditional Random Fields. *Proceedings of the 18th International Conference on Machine Learning 2001 (ICML 2001)*, 282–289. Retrieved from http://repository.upenn.edu/cis_papers%5Cnhttp://repository.upenn.edu/cis_papers%5Cnhttp://repository.upenn.edu/cis_papers.

Lan, M., Zhang, Z., Lu, Y., & Wu, J. (2016). Three Convolutional Neural Network-based models for learning Sentiment Word Vectors towards sentiment analysis. *Proceedings of the International Joint Conference on Neural Networks*, *2016-Octob*, 3172–3179.

Liu, H., Hu, Z., Mian, A., Tian, H., & Zhu, X. (2014). A new user similarity model to improve the accuracy of collaborative filtering. *Knowledge-Based Systems*, *56*, 156–166.

Liu, P., Joty, S., & Meng, H. (2015). Fine-grained opinion mining with recurrent neural networks and word embeddings. *Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, 1433–1443.

Liu, Y., Wang, J., & Wang, X. (2018). Learning to recognize opinion targets using recurrent neural networks. *Pattern Recognition Letters*, *106*, 41–46.

Luo, H., Li, T., Liu, B., Wang, B., & Unger, H. (2018). *Improving Aspect Term Extraction with Bidirectional Dependency Tree Representation*. Retrieved from http://arxiv.org/abs/1805.07889.

Maria, P., Dimitrios, G., Haris, P., & Suresh, M. (2015). SemEval-2015 Task 12: Aspect Based Sentiment Analysis. *Proceedings Ofthe 9th International Workshop on Semantic Evaluation (SemEval 2015)*, 486–495.

Mei, Q., Ling, X., Wondra, M., Su, H., & Zhai, C. (2007). Topic sentiment mixture: Modeling facets and opinions in weblogs. *16th International World Wide Web Conference, WWW2007*, 171–180.

Mikolov, T., Yih, W., & Zweig, G. (2013). Linguistic regularities in continuous space word representations. *Proceedings of NAACL-HLT*, 746–751. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Linguistic+Regularities+in+Continuous+Space+

DOI: 10.56892/bima.v7i01.394

Word+Representations#0%5Cnhttps://www.aclweb.org/anthology/N/N13/N13-1090.pdf

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543.

Pham, D. H., & Le, A. C. (2018). Exploiting multiple word embeddings and one-hot character vectors for aspect-based sentiment analysis. *International Journal of Approximate Reasoning*, *103*, 1–10.

Pontiki, M., & Pavlopoulos, J. (2014). SemEval-2014 Task 4: Aspect Based Sentiment Analysis. *Proceedings of the 8th International Workshop on Semantic Evaluation*, (SemEval), 27–35.

Popescu, A., & Etzioni, O. (2005). Extracting Product Features and Opinion from Reviews. *Human Language Technology and Empirical Methods in Natural Language Processing, Vancouver, British Columbia*, 339–346.

Poria, S., Cambria, E., & Gelbukh, A. (2016a). Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems*, *108*, 42–49.

Poria, S., Cambria, E., & Gelbukh, A. (2016b). Aspect Extraction for Opinion Miningwith a Deep Convolutional Neural Network. *Knowledge-Based Systems*, *108*, 42–49.

Poria, S., Cambria, E., Ku, L.-W., Gui, C., & Gelbukh, A. (2014). A Rule-Based Approach to Aspect Extraction from Product Reviews. *Second Workshop on Natural Language Processing for Social Media (SocialNLP)*, 28–37.

Shin, B., Lee, T., & Choi, J. D. (2016). *Lexicon Integrated CNN Models with Attention for Sentiment Analysis*. 149–158.

Toh, Z., & Wang, W. (2014). DLIREC: Aspect Term Extraction and Term Polarity Classification System. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, 235–240.

Tran, T. U., Hoang, H. T. T., & Huynh, H. X. (2019). Aspect extraction with bidirectional GRU and CRF. *RIVF 2019 - Proceedings: 2019 IEEE-RIVF International Conference on Computing and Communication Technologies*, 1–5.

Wang, W., Pan, S. J., Dahlmeier, D., & Xiao, X. (2016a). *Coupled Multi-Layer Attentions for Co-Extraction of Aspect and Opinion Terms*.

Wang, W., Pan, S. J., Dahlmeier, D., & Xiao, X. (2016b). Recursive Neural Conditional Random Fields for Aspect-based Sentiment Analysis. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP-16)*, 616–626.

Xu, H., Liu, B., Shu, L., & Yu, P. S. (2018a). Double embeddings and cnn-based sequence labeling for aspect extraction. *ACL 2018 - 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)*, *2*, 592–598.

Xu, H., Liu, B., Shu, L., & Yu, P. S. (2018b). Double Embeddings and CNN-based Sequence Labeling for Aspect Extraction. *Proceedings Ofthe 56th Annual Meeting Ofthe Association for Computational Linguistics*, 592–598.

Yin, Y., Wei, F., Dong, L., Xu, K., Zhang, M., & Zhou, M. (2016). Unsupervised word and dependency path

DOI: 10.56892/bima.v7i01.394

embeddings for aspect term extraction. *IJCAI International Joint Conference on Artificial Intelligence*, *2016-Janua*, 2979–2985.

Zhang, Z., Zou, Y., & Gan, C. (2017). Textual sentiment analysis via three different attention convolutional neural networks and cross-modality consistent regression. *Neurocomputing*, *275*, 1407–1415.

Zhao, H., Huang, L., Zhang, R., Lu, Q., & Xue, H. (2020). Spanmlt: A span-based multi-task learning framework for pair-wise aspect and opinion terms extraction. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 3239–3248.