

Extraction of aspects from Online Reviews Using a Convolution Neural Network

Vidya Kamma¹
Sridevi Gutta²
Teja Santosh Dandibhotla³

Abstract

The quality of the product is measured based on the opinions gathered from product reviews expressed on a product. Opinion mining deals with extracting the features or aspects from the reviews expressed by the users. Specifically, this model uses a deep convolutional neural network with three channels of input: a semantic word embedding channel that encodes the semantic content of the word, a part of speech tagging channel for sequential labelling and domain embedding channel for domain specific embeddings which is pooled and processed with a Softmax function. This model uses three input channels for aspect extraction. Experiments are conducted on amazon review dataset. This model achieved better results.

Keywords: Neural network, softmax function and extraction.

AMS Classification: 68T07⁴

¹ Research Scholar, Department of Computer Science and Engineering, KoneruLakshmaiah Education Foundation, Vaddeswaram, A.P., India. & Assistant Professor, Department of Computer Science, Neil Gogte Institute of Technology. Mail Id: kammavidya@gmail.com

² Professor, KoneruLakshmaiah Education Foundation, Vaddeswaram, A.P., India. Mail Id:sridevi.gutta2012@gmail.com

³ Associate Professor, Sreenidhi Institute of Science and Technology, Yamnampet, Ghatkesar, Telangana, India. Mail Id:tejasantoshd@gmail.com

⁴Received on June 16th, 2022. Accepted on Aug 10th, 2022. Published on Nov 30th, 2022. doi: 10.23755/rm.v44i0.909. ISSN: 1592-7415. eISSN: 2282-8214. ©The Authors. This paper is published under the CC-BY licence agreement.

1. Introduction

Customers are expressing views and opinions of products purchased. Customers can leverage the information contained in such reviews to identify the best products. These reviews help organizations to identify customer needs. The rise of reviews is gaining greater attention in research community. When it comes to traditional sentiment analysis, talk about opinion on entire sentence. Customer reviews often express opinion on different aspects of product rather than opinion on whole product Schouten, K., & Frasincar, F. (2015). This raised the problem of aspect extraction for sentiment analysis. Many challenge to be addressed, identifying implicit aspects, multiple aspects in single statement.

The major task of AE (Aspect Extraction) to Extracting aspects of customer reviews Cheung, D. 2009, et al. Jakob, N., & Gurevych, I. (2010), extracted aspects of a spatial dataset using conditional random fields, the limitation is it need to work of large datasets and is linear in nature. Hu, M., & Liu, B. (2004), used linguistic models for extracting aspects from reviews. The limitation is rules to be drafted manually. Efficiency based on grammatical accuracy of sentences.

Neural network models are widely used for aspect extraction. Existing CNN based models are lacking in extracting context-level features. Ye, H. et al., (2017) and Dong, F. et al., (2017), used CNN integrated with LSTM for aspect extraction achieved better performance with increase the complexity of the model parameters. In this paper, to overcome above mentioned issues we contributed the following.

To encode contextual information, three different input channels were used: i) Word embeddings + lexicon ii) Domain embeddings iii) POS Tag embeddings. Previous models used either pre-trained embeddings or review embeddings Poria, S. et al., (2016). To increase performance of model we included lexicons in word embedding layer. The first two channels are used to capture semantic and syntactic information. Third channel is used for sequential labelling of aspects.

2. Related work

Among the many tasks involved in aspect level sentiment analysis, a major one is aspect extraction (Liu B, 2012). Current approaches to extracting features are based on deep learning. Macháček, J. (2016) implemented traditional supervised machine learning method which uses biagram (BOW) bag of words. Li and Lam, (2017) performed aspect extraction by considering annotated data and achieved better results. Luo et al., (2018) integrated word embedding with CRF&BiLSTM for better extracting the aspects. Toh, Z., &Su, J. (2016) combined CNN with binary classifier and implemented anhybrid approach. Approach was top rated in SemEval 2016 Task 5 competition. Jihan, N. et al., (2017) used pre-processing pipeline for normalizing the data. They implemented multi-domain feature extraction and predicted aspect category using SVM. Khalil, T. et al (2016) initialized pre-trained word vectors for CNN and bag of words as features to ensemble classifier. To improve classification, they used secondary classifier. Stéphan Tulkens et al., (2020), implemented (CA_t) Contrastive

Attention mechanism an unsupervised approach based on RBF kernel for extraction aspects. Ruidan He et al., (2017), extracted coherent aspects using neural approach. They used attention mechanism in the training phase to minimize irrelevant words. Poria, S. et al., (2016), used Deep CNN with 7 layers plus Linguistics Patterns for extracting aspects in product reviews. Wang et al., (2016) proposed the method which combined both CRF and dependency tree for better aspect extraction. Wang et al., (2014), based on seeding words implemented the extraction of product aspects was done using two semi-supervised models. Collobert et al., (2011) initialized CNN with word embeddings to solve semantic role labelling and named entity recognition problem. YinYang et al., (2017), proposed (CAT-LDA) a two-layer topic model to extract hierarchical aspects i.e parent and child category. Lin Wang (2015), used Restricted Boltzmann Machines to extract aspects. Xu et al., (2017) and Li & Lam (2018), used deep learning model for extracting aspect and opinion items.

3. Model

The proposed model has three input layers. Three embedding layers word-embedding layer enriched word vectors using lexicons, domain embedding layer, POS embedding layer, CNN layers with ReLu-nonlinear activation function, fully connected layer, pooling layer for multiclass labels a softmax classifier is used for labelling $y=\{B,I,O,E,S\}$,with beginning, inside, outside, end and single of the aspect term. Figure 1. Shows the proposed CNN model

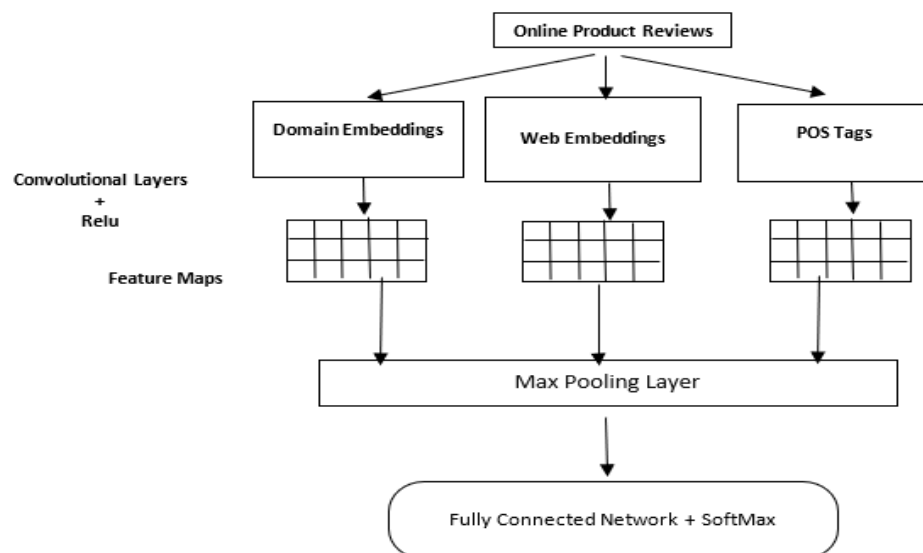


Figure 1. Proposed Convolution Neural Network model

Pre-processing

Mainly pre-processing is performed to clean and convert the reviews for further processing. We used python with NLP libraries are used for removal of numbers, converting acronyms, white spaces, tabs, stop words, words with length of 1 and special characters. Repeated letter words i.e “goooooood” or replaced with original form. All the review sentences are converted to lower case.

Word Embeddings

The three inputs google corpus, amazon reviews, POS tagging are mapped to vectors. Padding is performed to maintain the sentences of equal length. Let the input $w.x_{i+h}$ $x = (x_1, \dots, x_n)$. The three embedding layers representations are W^{gl} for google + lexicon embeddings, W^d for domain embeddings and W^p for POS Tag embeddings are fed to CNN.

Convolution

Convolution layer is mainly used to extract features. Filters of different sizes (3,4,5) are used on each matrix. The stride used in each layer is 1 to tag each word. The Relu is the activation function is used with the convolution operation. The features produced will be of the form.

$$G_i = \max(0, (\sum_{i=1}^n w.x_{i+h} + b)) \quad (1)$$

Where $w.x_{i+h}$ is weight and vector of h-gram in a sentence.

Max Pooling Layer

Pooling operation is applied on the feature maps. The purpose of this layer is to extract maximum efficient features so max pooling is applied. The feature map obtained is the concatenation of three feature maps. Where n, m, o represents the filter sizes for google +lexicon features, domain features and POS features.

$$f = f_{gl}^1 \oplus \dots \oplus f_{gl}^n \oplus f_d^1 \oplus \dots \oplus f_d^m \oplus f_p^1 \oplus \dots \oplus f_p^o \quad (2)$$

To choose hyperparameters we applied cross-validation strategy. Overfitting is the major problem with deep neural networks if the parameters are high. Dropout regularization is added. 50% is considered as the dropout rate to avoid over fitting. And

features are added to fully connected layer $G_i = \max(0, (\sum_{i=1}^n w.x_{i+h} + b))$.

$$K = \alpha(w * f + b) \quad (3)$$

Where w is the weight matrix, b is bias and α is Relu activation function.

Softmax Classifier

This is the final layer used for classification of aspects into classes with highest probability.

$$\hat{y} = \arg \max_j P(y = j / x, w, a) = \arg \max_j (e^{xw_j + a_j} / \sum_k 1^k x e^{xw_k + a_k})$$

w_j : weights of class j and a_j : bias of class j . The above-mentioned process is repeated for the complete training samples.

4. Experiments

These datasets are a collection of Amazon reviews of four categories of products. Product categories included in the analysis of reviews include laptops, smartphones, cameras and wristwatches. From the E-commerce applications, 100 products are considered for each product category. Dataset details are presented in Table 1.

Table.1: Dataset details

Document attributes	Values
Documents reviewed	4,13,841
Each review should include one sentence	1
Max. sentences per review	43
Customer review average	5.98
Review count for the product on average	49.47

Amazon reviews are used to train the domain embeddings. Electronic product reviews are collected from amazon. In this work, word embeddings are obtained from the Google News Corpus Word2Vec embeddings. We remove all stop words and non-English words from the data during pre-processing. POS tags are applied to the obtained words. The table 2 below shows the experimental setup for the modified CNN model, which covers a variety of settings.

Table.2: Experimental setup

Parameters	Value
Filter size	(3,4,5)
No. of feature maps	100
Pooling Function	MAX
Regularization (Dropout Rate)	50%
Activation Function	ReLu
Number of epochs	10

By carefully adjusting the settings of the modified CNN, good accuracy was achieved. According to Pedro Domingos, "Bigger data are better than smart algorithms" which is also the case for our modified CNN model which is clear from accuracy metrics when considering the size of the reviews collection.

Figure-2 and Figure-3 shows that the model has different performances across the four different domains.

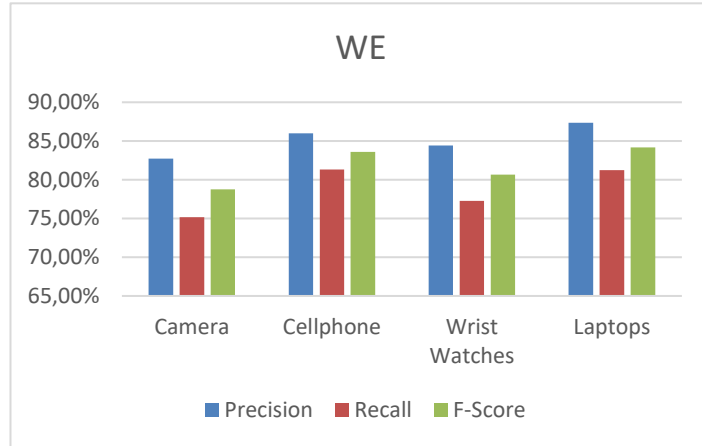


Figure 2: Comparison of the performance with in four domains

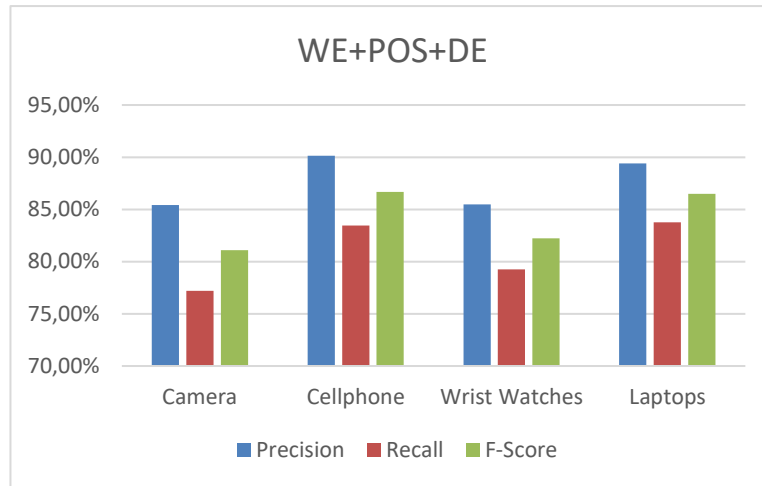


Figure 3: Comparison of the performance with in four domains

Table.3: Impact of the POS feature on the dataset.

Domain	Classifiers	Precision	Recall	F-Score
Camera	WE	80.74%	74.15%	74.76%
Camera	WE+POS+DE	84.42%	75.21%	80.10%
Cell phone	WE	84.01%	80.32%	81.59%
Cell phone	WE+POS+DE	89.05%	83.47%	84.68%
Wrist Watches	WE	82.41%	73.26%	79.67%
Wrist Watches	WE+POS+DE	83.48%	77.25%	81.24%
Laptops	WE	86.35%	80.23%	82.17%
Laptops	WE+POS+DE	86.40%	80.77%	83.49%

POS features were used in conjunction with word embedding features to achieve the

highest accuracy on this dataset. POS play an important role in extraction of aspects as well as word embedding, as demonstrated here. Impact of the POS feature on the dataset is presented in Table.3.

5. Conclusion

In this paper, three channel-based CNN for aspect extraction is introduced. Experiments are conducted on five different domains smartphones, camera, laptop, wrist watches, books. Python is used as a computational environment. We determined the efficiency of using lexicon in word embedding layer. These carefully configured settings have helped the modified CNN architecture achieve good accuracy.

References

- [1] Cheung, D. W. L., Song, I. L., Chu, W., Hu, X. and Lin, J., “Proceedings of the 18th ACM Conference on Information and Knowledge Management”, Association of Computing Machinery, 2009.
- [2] Schouten, K. and Frasincar, F., “Survey on Aspect-Level Sentiment Analysis”, *IEEE Transactions on Knowledge and Data Engineering*, 28(3), pp.813-830, 2015.
- [3] Jakob, N. and Gurevych, I., “Extracting Opinion Targets in A Single-And Cross-Domain Setting with Conditional Random Fields”, *In Proceedings of the conference on empirical methods in natural language processing*. Association for Computational Linguistics, pp.1035-1045, 2010.
- [4] Hu, M. and Liu, B., “Mining and Summarizing Customer Reviews”, *In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp.168-177, 2004.
- [5] Ye, H., Yan, Z., Luo, Z. and Chao, W., “Dependency-Tree Based Convolutional Neural Networks for Aspect Term Extraction”, *In Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Springer, Cham, pp.350-362, 2017.
- [6] Dong, F., Zhang, Y. and Yang, J., “Attention-Based Recurrent Convolutional Neural Network for Automatic Essay Scoring”, *In Proceedings of the 21st Conference on Computational Natural Language Learning (Co NLL 2017)*, pp.153-162, 2017.
- [7] Poria, S., Cambria, E. and Gelbukh, A., “Aspect Extraction for Opinion Mining with A Deep Convolutional Neural Network”, *Knowledge-based Systems*, 108, pp.42-49, 2016.
- [8] Khalil, T. and El-Beltagy, S. R., “Niletmrg At Semeval-2016 Task 5: Deep Convolutional Neural Networks For Aspect Category And Sentiment Extraction”, *In Proceedings of the 10th international workshop on semantic evaluation (SEMEVAL-2016)*, pp.271-276, 2016.

- [9] Liu, B., “Sentiment Analysis and Opinion Mining”, First Edition. San Rafael: Morgan & Claypool Publishers, 2012.
- [10] Macháček, J., “Butknot at Semeval-2016 Task 5: Supervised Machine Learning with Term Substitution Approach in Aspect Category Detection”, *In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pp.301-305, 2016.
- [11] Toh, Z. and Su, J., “Nlangp at Semeval-2016 Task 5: Improving Aspect Based Sentiment Analysis Using Neural Network Features”, *In Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)*, pp.282-288, 2016.
- [12] Jihan, N., Senarath, Y., Tennekoon, D., Wickramarathne, M. and Ranathunga, S., “Multi-Domain Aspect Extraction Using Support Vector Machines”, *In Proceedings of the 29th Conference on Computational Linguistics and Speech Processing (ROCLING 2017)*, pp.308-322, 2017.
- [13] Wang, W., Pan, S. J., Dahlmeier, D. and Xiao, X., “Recursive Neural Conditional Random Fields for Aspect-Based Sentiment Analysis”, 2016. *arXiv preprint arXiv:1603.06679*.
- [14] Li, X., and Lam, W., “Deep Multi-Task Learning for Aspect Term Extraction with Memory Interaction”, *In Proceedings of the conference on empirical methods in natural language processing*, pp.2886-2892, 2017.
- [15] Ye, H., Yan, Z., Luo, Z. and Chao, W., “Dependency-Tree Based Convolutional Neural Networks for Aspect Term Extraction”, *In Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Springer, Cham, pp.350-362, 2017.
- [16] Tulkens, S. and Van Crane burgh, A., “Embarrassingly Simple Unsupervised Aspect Extraction”, 2020. *arXiv preprint arXiv:2004.13580*.
- [17] He, R., Lee, W. S., Ng, H. T. and Dahlmeier, D., “An Unsupervised Neural Attention Model for Aspect Extraction”, *In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 388-397, 2017.
- [18] Wang, T., Cai, Y., Leung, H. F., Lau, R. Y., Li, Q. and Min, H., “Product Aspect Extraction Supervised with Online Domain Knowledge”, *Knowledge-Based Systems*, 71, pp.86-100, 2014.
- [19] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K. and Kuksa, P., “Natural Language Processing (Almost) From Scratch”, *Journal of machine learning research*, 12, pp.2493-2537, 2011.
- [20] Xu, L., Lin, J., Wang, L., Yin, C. and Wang, J., “Deep Convolutional Neural Network based Approach for Aspect-Based Sentiment Analysis”, *Advanced Science and Technology Letters*, 143, pp.199-204, 2017.
- [21] Li, X. and Lam, W., “Deep Multi-Task Learning for Aspect Term Extraction with Memory Interaction”, *In Proceedings of the conference on empirical methods in natural*

language processing, pp.2886-2892, 2017.

[22] Wang, L., Liu, K., Cao, Z., Zhao, J. and De Melo, G., “Sentiment-Aspect Extraction Based on Restricted Boltzmann Machines”, *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. 616-625, 2015.

[23] Yang, Y., Chen, C., Qiu, M. and Bao, F., “Aspect Extraction from Product Reviews Using Category Hierarchy Information”, *In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pp.675-680, 2017.

[24] Domingos, P., “A Few Useful Things to Know About Machine Learning”, *Communications of the ACM*, 55(10), pp.78-87, 2012.

[25] Qiu, G., Liu, B., Bu, J. and Chen, C., “Opinion Word Expansion and Target Extraction Through Double Propagation”, *Computational linguistics*, 37(1), pp.9-27, 2011.