Aspect – Based Sentiment Analysis

A. Punitha[1], R. Keerthi Prabu[2], P. Devanathan[3], A. Sanjai[4], P. Bala[5]

[1]Associate Professor, Department of Computer Science & Engineering, Annamalai University, India
[2,3,4,5]Student, Faculty of Engineering & Technology, Annamalai University, India

Abstract
Aspect-Based Sentiment Analysis (ABSA) is an advanced NLP application that aims to identify aspect terms present in the given review and predict the sentiment associated with those aspect terms. ABSA is better than sentence-based sentiment classification because it considers the aspect terms present in the reviews to determine the sentiment rather than considering the individual sentence. Entrepreneurs could make use of ABSA to understand the customers' opinions about different aspects of their products or services. The task of Aspect-based sentiment analysis can be divided into two subtasks: Aspect Term Extraction (ATE) and Aspect Term Sentiment Classification (ATSC). In this paper, an SVM model is proposed for the task of ATE, and an Attention based LSTM model is proposed for the task of ATSC. The proposed models will be trained and tested on SemEval-2014 dataset.

Keywords: Aspect-Based Sentiment Analysis, NLP, Aspect Term Extraction, Aspect Term Sentiment Classification, SVM and Attention based LSTM.

1.INTRODUCTION

The recent burst in the number of e-commerce websites and social media platforms enabled many people to express their opinions about the products and services they are consuming which eventually increased the quality and quantity of the reviews available for businesses to conduct market research to analyze customers' opinions about their product and services. One of the techniques companies use to analyze users' opinions is called "Aspect Based Sentiment Analysis".

Aspect-based sentiment analysis is an advanced level sentiment analysis task that aims to identify the aspect terms mentioned in a given product review (e.g., Processor, RAM, etc in case of a review about a Laptop or a Smartphone) and predict the sentiment (e.g., Neutral, Positive and Negative) associated with each of those aspect terms.

For example, in the given product review, "battery is really good but its size is too small", there are two aspects "battery" and "size", and sentiments associated with those aspects are "Positive" and "Negative" respectively.

The reviews for performing aspect-based sentiment analysis can be obtained from customer reviews in various online platforms and through social media posts. There are many applications for aspect-based sentiment analysis which includes monitoring customer satisfaction with a specific product or service, Enabling business owners to analyze which aspects of their product or service was liked or disliked by the consumers, Conducting market research and many more.

The task of aspect-based sentiment analysis can be broken into two sub-tasks, (i) Aspect Term Extraction (ATE) and (ii) Aspect Term Sentiment Classification (ATSC). The first task, ATE, involves extraction of all the aspects mentioned in the customer review and the second task involves predicting sentiment polarities of the aspects extracted from the customer reviews.

For the aspect extraction task, A Support Vector Machine model is used to extract the aspects from users' review because they can effectively handle high-dimensional feature spaces, identify complex relationships between features and output labels, and perform well even with a limited amount of training data. And for the ATSC, an Attention based LSTM model is proposed. Attention based LSTM model is preferred over traditional methods because Traditional sentiment analysis methods often rely on feature engineering and use handcrafted features to identify sentiment expressions in the text. However, such approaches can be limited in their ability to capture the nuances of the language and the complex interactions between different aspects and sentiment expressions. Attention based LSTM model, on the other hand, use a neural network architecture that is specifically designed to capture such complex interactions. This model uses multiple layers of attention mechanisms to focus on the most relevant aspects and sentiment expressions in the text. The first layer of attention focuses on the relevant aspects in the text, while the second layer of attention focuses on the sentiment expressions related to those aspects.

2. Related Works

The related works on 'aspect-based sentiment classification' were reviewed below. As mentioned before the task of ABSA can be subdivided into ATE and ATSC.

2.1 Aspect Terms Extraction:

ATE is a classification task involving identifying the aspect terms present in a given review. A number of approaches are proposed by the papers that were reviewed so far. Soujanya Poria, et al., 2016 proposed a convolutional neural network model for ATE task because CNNs can automatically learn relevant features from raw text data, without the need for hand-engineered features. This is achieved by treating the reviews as a sequence of words or tokens, and using convolutional filters to extract n-grams of varying lengths. These n-grams capture local patterns and dependencies in the text, and can be combined and processed by subsequent layers to learn higher-level representations of the review. The output of the CNN can then be used to predict the aspect labels for each token in the review, based on the learned representations. Gunjan Ansari, et al., 2020 proposed KNN machine learning model. Chouikhi, et al., 2023 proposed a Conditional Random Fields (CRF) model because they can model the dependencies between adjacent output labels, which is particularly useful for aspect extraction as the presence of one aspect can affect the likelihood of other aspects appearing in the same sentence. Xin Li, et al., 2018 proposed Recursive Neural Conditional Random Fields (RNCRF) similar to CRF in which the Recurrent network acts as a feature extractor. Łukasz Augustyniak, et al., 2021 proposed an LSTM model for ATE task because of its ability to capture long-term dependencies between words. This is particularly useful for aspect extraction as aspects can often span multiple sentences or paragraphs in a review.

2.2 Sentiment Classification:

Aspect Term Sentiment Classification is an NLP classification task which aims to identify the sentiments associated with a given aspect Term in a given review. Different approaches and models for the ATSC task are proposed by the papers that were reviewed so far. Gayatree Ganu, et al., 2009 proposed SVM model to determine the sentiments associated with the given aspect term in a review. Dehong Ma, et al., 2017 proposed Interactive Attention Networks that could model target and context information interactively. Wei Xue, et al., proposed Gated CNN model, the gated Tanh-ReLU units can selectively output the sentiment features according to the given aspect or entity, the computations of the model could be easily parallelized during training, because convolutional layers do not have time dependency as in LSTM layers, and gating units also work independently. Qingnan Jiang, et al., 2019 proposed CapsNet and CapsNet-BERT models for this task. However, the majority of the papers such as Yequan Wang, et al., Binxuan Huang, et al., and Youwei Song, et al.,proposed attention-based LSTM models to deal with ATSC tasks because of its ability to focus on specific parts of the review based on the given aspect term.

3. Dataset

The proposed model was experimented on the dataset of SemEval-2014 [13] which consists of customer reviews about restaurants and laptops. The proposed models were trained on the laptops data. The laptops data are split into training and testing data which consists of 3045 and 800 annotated reviews respectively. The distribution of the laptop data with respect to the sentiment labels are given in Table 1.

Table 1: Distribution of the laptop data with respect to the sentiment labels

Dataset	Positive	Negative	Conflict	Neutral	Total
LPT-TR	987	866	45	460	2358
LPT-TE	341	128	16	169	654

3.1 Dataset Pre-processing:

Remove dataset entries with no attributes: The reviews with no aspects does not help in training neither aspect extraction model nor the sentiment classification model, Those reviews are needed to be removed in order to reduce the amount of training data and to reduce the training time.
Corrected spelling mistakes: The spelling mistakes present in the reviews of the dataset will result in different vectorization than its intended vectorization of the correct word which will eventually result in the poor training of the models. Hence, it is required to correct all the incorrect spellings.
Add sentiment polarity to the test dataset: Since manually annotated sentiment polarity are needed in the test dataset to evaluate the performance of the sentiment classification model, it is necessary to add sentiment polarity to the testing data.
Pre-process the text data: It includes converting the text into lowercase and removing stopwords and punctuations.
POS tagging: This step involves identifying the parts-of-speech tag of each word in the text data. Then, the words that are not either noun or verb are filtered out.
Convert text data into numerical data: The pre-processed text data is finally converts into an array of numerical data that can be fed as input to the models.

4 Methodology:

4.1 Aspect Term Extraction:

The first step towards dealing with ATE tasks is to pre-process the aspect term labels in training and testing data. The output for the ATE task should be in the format of a numerical array of size 1*n (for example, A

= [a1, a2, a3, … , an]) where n is the number of unique aspect terms present in the training and testing data. The output array A consists of only two values 0 and 1. If a review consists of any aspect term, the corresponding index in the output array A is filled with 1 and the rest of the indices are filled with 0s. For example, A = [0,1,0,0,1, … , 0]. The next step involves training the proposed SVM model along with other machine learning model for the case of comparison.

4.1.1 SVM:

Support Vector Machine (SVM) is a type of supervised machine learning algorithm that is used for classification and regression analysis. The main objective of SVM is to find the optimal hyperplane that can separate the different classes in the dataset. The hyperplane is defined as the line or plane that maximizes the margin between the closest points of the different classes. The points that lie on the margin are called Support Vectors. SVM works by mapping the input data into a high-dimensional feature space using a kernel function. The kernel function is used to transform the data from the input space to a higher-dimensional space where it is easier to separate the classes with a linear hyperplane. SVM can handle both linear and non-linear classification problems by choosing the appropriate kernel function.
The output of the SVM is calculated using the formula given below:

$$Y = 1 \text{ if } \vec{X}.\vec{w} \geq 0 \text{ ; -1 if } \vec{X}.\vec{w} \leq 0 \qquad (1)$$

The formula mentioned above is used for binary classifiers. In order to use this method for multi-class classification one should use a one vs rest classifier which compares the values of between one class and rest of the classes and selects the class that has the highest value.

4.2 Sentiment Classification:

4.2.1 Aspect based embeddings:

After aspect identification, Given a sentence s = [w1, w2, ..., wi , ..., wj , ..., wn] with length n and a target (extracted aspect term) t = [wi , wi+1, ..., wi+m−1] with length m, each word is mapped into a low-dimensional real-value vector, called word embedding. These word embeddings act as numerical features that capture the sentiment polarity of each aspect.

4.2.2 LSTM Model:

The LSTM model is designed to process the aspect-based embeddings in sequence and learn the relationships between the aspects and their associated sentiment polarities. The LSTM model consists of three gates: input gate, output gate, and forget gate. These gates regulate the information flow in the LSTM model.
The input gate is used to quantify the importance of the new information carried by the input. Here is the equation of the input gate,

$$I_t = s(x_t * U_i + H_{t-1} * W_i) \qquad (2)$$

The equation of output gate is given below,

$$O_t = s(x_t * U_o + H_{t-1} * W_o) \qquad (3)$$

Its value will also lie between 0 and 1 because of this sigmoid function. If tanh function is used instead of sigmoid function, the above function can be rewritten as

$$H_t = O_t * \tanh(C_t) \qquad (4)$$

The function of the forget gate is to decide whether to keep the information from the previous time step or forget it. Here is the equation for the forget gate.

$f_t = s(x_t * U_f + H_{t-1} * W_f)$ (5)

Later, a sigmoid function is applied to it. That will make ft a number between 0 and 1. This ft is later multiplied with the cell state of the previous timestamp, as shown below.

$C_{t-1} * f_t = 0 \ldots \text{ if } f_t = 0$ (forget everything) (6)

$C_{t-1} * f_t = 0 \ldots \text{ if } f_t = 1$ (forget nothing) (7)

4.2.3 Attention mechanism:

The attention mechanism [14] is used to weigh the importance of each aspect-based embedding. This allows the model to focus on the most relevant aspects when making predictions. The attention mechanism calculates a weight for each aspect-based embedding based on its relevance to the sentiment polarity prediction. These weights are then used to compute a weighted average of the aspect-based embeddings, which is then fed into the output layer to predict the sentiment polarity.

The weights αij are computed by a SoftMax function given by the following equation:

$\alpha_{ij} = \exp(e_{ij}) \div \sum_{k=1}^{Tx} \exp(e_{ik})$ (8)

$e_{ij} = a(s_{i-1}, h_j)$ (9)

5 Evaluation:

The proposed models SVM and Attention based LSTM are tested on the testing data of the laptops data along with other similar machine learning models. The evaluation metrics of these models are given in Table 2 and 3.

Table 2: Evaluation metrics of the models trained for ATE task

Model	**Accuracy**
Naïve Bayes	99.12%
SVM	80.12%
SGD	70.75%
Decision Tree	59.50%

Table 3: Evaluation metrics of the models trained for ATESC task

Model	**Accuracy**

Naïve Bayes	43.10%
SVM	56.04%
K Nearest Neighbors	49.06%
Decision Tree	49.74%
Random Forest	60.30%
Voting Classifier (Hard)	55.36%
Voting Classifier (Soft)	52.81%
AE-LSTM	86.32%

6 Conclusion:

In conclusion, aspect-based sentiment analysis is a powerful tool for understanding the opinions and attitudes expressed in large volumes of text data. By analyzing the sentiment of individual aspects or features of a product or service, researchers and businesses can gain a more nuanced understanding of customer feedback and use that information to improve their offerings. In this project, various approaches to aspect-based sentiment analysis are explored, including machine learning and deep learning methods. As the field continues to evolve, it is expected to see further developments in the techniques and tools used for aspect-based sentiment analysis, making it an even more valuable tool for businesses and researchers alike.

References:

1. Soujanya Poria, Erik Cambria, Alexander Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network", Knowledge-Based Systems, Volume 108, 2016, Pages 42-49
2. Gunjan Ansari, Chandni Saxena, Tanvir Ahmad, M.N. Doja, "Aspect Term Extraction using Graph-based Semi-Supervised Learning", Procedia Computer Science, Volume 167, 2020, Pages 2080-2090
3. Chouikhi, H., Alsuhaibani, M., Jarray, F., "BERT-Based Joint Model for Aspect Term Extraction and Aspect Polarity Detection in Arabic Text". Electronics 2023, *12*, 515.
4. Xin Li, Lidong Bing, Piji Li, Wai Lam, Zhimou Yang, "Aspect Term Extraction with History Attention and Selective Transformation", arXive 2018.
5. Łukasz Augustyniak, Tomasz Kajdanowicz, Przemysław Kazienko, "Comprehensive analysis of aspect term extraction methods using various text embeddings", Computer Speech & Language, Volume 69, 2021, 101217.
6. Ganu, G., Elhadad, N., & Marian, A., "Beyond the Stars: Improving Rating Predictions using Review Text Content". *International Workshop on the Web and Databases* 2009.
7. Dehong Ma, Sujian Li, Xiaodong Zhang, and Houfeng Wang, "Interactive attention networks for aspect-level sentiment classification", arXiv 2017.
8. Wei Xue and Tao Li, "Aspect based sentiment analysis with gated convolutional networks", arXiv 2018.

9. Qingnan Jiang, Lei Chen, Ruifeng Xu, Xiang Ao, and Min Yang. 2019. “A Challenge Dataset and Effective Models for Aspect-Based Sentiment Analysis”. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pages 6280–6285, Hong Kong, China. Association for Computational Linguistics.
10. Wang, Yequan & Huang, Minlie & Zhu, Xiaoyan & Zhao, Li., “Attention-based LSTM for Aspect-level Sentiment Classification”, 2016, 606-615.
11. Huang, Binxuan and Ou, Yanglan and Carley, Kathleen M, "Aspect Level Sentiment Classification with Attention-over-Attention Neural Networks", 2018, Springer International Publishing, 197-206.
12. Youwei Song, Jiahai Wang, Tao Jiang, Zhiyue Liu, and Yanghui Rao., “Attentional encoder network for targeted sentiment classification”, 2019, arXiv.
13. Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. SemEval-2014 Task 4: Aspect Based Sentiment Analysis. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 27–35, Dublin, Ireland. Association for Computational Linguistics.
14. Bahdanau, Dzmitry, Kyunghyun Cho and Yoshua Bengio. “Neural Machine Translation by Jointly Learning to Align and Translate.”, 2014.