

Application of ABSA from Business perspective on Restaurant reviews

Vedashree Paithankar¹, Shreya Patil², Vinita Doiphode³, Jithisha Koshy⁴, Ashwini Bhamre⁵,
Mukta Kore⁶

^{1,2,3,4}Student, Department of Information Technology Engineering, P.E.S.'s Modern College of Engineering,
Pune, Maharashtra, India

^{5,6}Asst. Professor, Dept. of Information Technology Engineering, P.E.S.'s Modern College of Engineering,
Pune, Maharashtra, India

Abstract - Due to the evolution and vast use of the Internet, millions of reviews are posted. So, there is a need to analyze them efficiently to identify the emotion of customers behind the reviews posted. In the case of restaurant reviews, it is a tedious and intricate task to manually understand what the customer has to say in each review at that particular moment. In this paper, we have proposed a system that will extract aspects from restaurant reviews and classify the sentiments of reviews as positive or negative. This will benefit the business owner in understanding the customers' opinions towards various features of the restaurant, obtain sharp results, and make respective changes or advancements in real-time leading to business growth. This system considers the owner and business point of view. CNN (Convolutional Neural Networks) is executed on Sem-Eval 2016 dataset to extract aspect and sentiment categories and the results are visualized in the form of word clouds and bar graphs for clear understanding to the owner.

Key Words: Aspects, Sentiments, CNN, Sem-Eval 2016, word cloud, bar graphs

1. INTRODUCTION

Opinions given by the customers are very crucial for any organization helping them to identify the organization's strengths and weaknesses which results in the generation of various approaches that aim to improve the services.

The availability of a wide communication medium presents customers with the ability to write and express their experiences regarding a particular product or service which is being availed. Unfortunately, some users or customers may post unethical and fake reviews to promote or even criticize the brand name. Considering this scenario, the proposed system considers reviews only from the people physically visiting the restaurant. Our system pivots on analyzing the reviews given by individuals visiting the restaurant and generating an output which is favorable for business growth and development.

This technique used to scrutinize the opinions and judgments from textual data in positive and negative is called Sentiment Analysis. The analysis of opinions is

achieved on 3 distinct levels- Document level, Sentence level, and Aspect level.

We have implemented our system on an Aspect level as it gives a granular level insight of an entire review to the owner of the restaurant. Here aspect is a distinct trait or element of a product or service. For example- "The food is tasty" where "food" displays the aspect whereas "tasty" indicates a positive reaction of the reviewer. This positive emotion or attitude expressed is nothing but the sentiment.

Aspect-Based Sentiment Analysis (ABSA) is the method that includes the words linked to the aspects and recognizes the sentiment corresponding with every aspect. Aspect based analysis explores more detailed information behind a text. We are using Aspect-Based Sentiment Analysis to extract Aspect and Sentiment Categories of respective reviews using CNN (Convolutional Neural Network) and visualizing them in the form of word cloud and bar graph

1.1 LITERATURE SURVEY

[1] Reviews given by the customers are very valuable and help in making better decisions. A document is analyzed and the overall meaning is extracted using opinion mining. Aspect level opinion mining classifies customer reviews based on important features. In aspect level opinion mining comprises of 2 major parts. The first part is the opinion word identification while the second part is orientation detection.[2] There are number of different approaches used to perform aspect extraction which are frequency based, relation-based, supervised learning, topic modelling. [3] Part-of-Speech (POS) tagging is a technique which is used to extract the aspect and sentiment words. The words denoting nouns, adjective and verb are identified by Pos-tagger. Additionally, the syntactic grammatical relation between the words of review sentence is established using technique called dependency parsing. The parser used in [2] is Stanford parser. [4] CNN has been widely used to implement text classification related tasks. [4] presents an improved CNN architecture for aspect extraction, which does not require any domain specific features or pre-processing. [4] demonstrates how size and the domain of corpus used to train the word2vec model can affect the accuracy of CNN

models used for aspect extraction. [5] Word cloud is an image which is broadly used in diverse domains as a means to give an overview by distilling text down to those words that appear with highest frequency.

2. DESIGN AND ARCHITECTURE

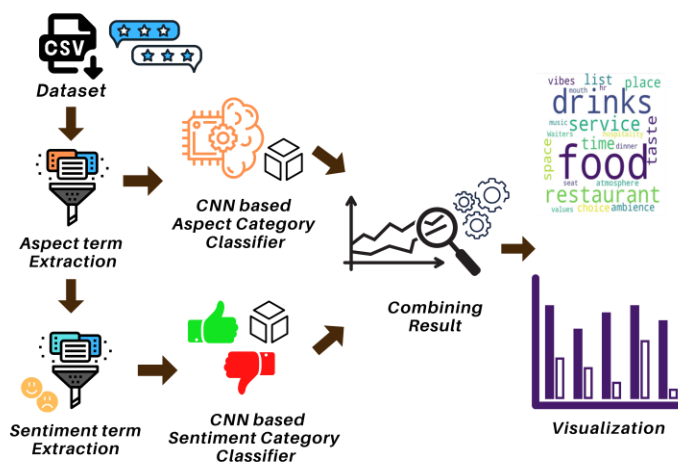


Fig-1: Architecture Diagram

The proposed system consists of Aspect Category Classifier and Sentiment Category Classifier. To build this system, Sem-Eval 2016 dataset is trained which contains restaurant reviews labeled with Aspect Categories and Sentiment Categories.

The Aspect Categories specified are ambiance, service, food, restaurant, and miscellaneous whereas the Sentiment Categories are positive and negative. Initially, the aspect terms are obtained using Spacy's Noun chunk dependency parser. These aspect terms are transformed into vector format using the Bag of Words embedding technique and then these vectors are fed to the Aspect Category Classifier model which is built using CNN and Keras library. The reviews are sorted into the respective 5 aspect categories.

Subsequently, adjectives and adverbs are extracted as sentiment terms using the POS tagger of Spacy. These are then fed to a similar CNN and Keras model i.e. the Sentiment Category Classifier which analyzes the reviews into positive and negative.

In the end, these results are visualized as word cloud and bar graph for clear understanding.

3. TECHNOLOGIES USED

3.1 Machine Learning

Machine learning (ML) is a concept that facilitates a computer program to learn and adapt to new data without human intervention. It consists of building complex algorithms that help the computer program spontaneously

grasp and, consequently, improve based on past knowledge and experience. Sentiment analysis using ML approach helps in extracting opinions about significant information from a large amount of textual data which can be profitable for qualitative business improvement.

3.2 Supervised Learning

Supervised learning is an ML subdomain that maps an input to an output based on sample input and its corresponding output present in the labeled dataset. A supervised learning algorithm inspects the training data and generates a function which is utilized for mapping fresh examples.

Supervised learning is good at classification. Therefore, we are using this technique to sort the new reviews given by customers into respective labeled aspects and sentiment categories from the dataset.

3.3 Aspect Based Sentiment Analysis (ABSA)

Aspect-based sentiment analysis is a level of sentiment analysis in which aspects or features are extracted from textual data and sentiments (positive or negative) are taken out regarding these aspects. It involves breaking down text data into smaller fragments allowing us to obtain more granular and accurate insights from the data.

Moreover, the owner of the restaurant can obtain a strong perspective regarding each aspect and emotion corresponding to that particular aspect. This gives a clear understanding of the owner in deciding exactly which segment of the restaurant demands more attention and development.

3.4 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN), a class of artificial neural networks is extensively applied for image classification and recognition. Recently, it is also used for intercepting difficulties faced in Natural Language Processing (NLP) associated tasks. CNN comprises of neurons that self-optimize by learning. Accordingly, CNN is executed for Aspect and Sentiment Category extraction and classification considering that it gives good results because of its high capacity to automatically identify features from reviews.

4. DATA PROCESSING

4.1 Dependency Parser

Dependency parsing is the process of interpreting the grammatical form based on the dependencies among the words in a sentence where various tags describe the association between two words in that sentence. These tags are the dependency tags. It derives a dependency parse tree

that symbolizes its grammatical arrangement and illustrates the relationship between "head" words and words which modify these heads. For example, in the phrase "cold pizza", the word cold modifies the meaning of the noun pizza. Therefore, a dependency exists from morning to cold in which "pizza" acts as the head and "cold" acts as dependent or child.

This system applies Spacy's Noun chunk dependency parser to obtain the nouns as aspect terms from reviews that are further converted into a vector format and given as input to Aspect Category Classifier.

4.2 Word Embedding

Many Deep Learning algorithms cannot discern plain text in its raw form. This data needs to be transformed into numbers and fed as input for performing any task. And with a large amount of data that can be present in any document, it is compelling to extract and understand information out of it. Word Embedding maps a word using a dictionary to a vector which can be a one-hot encoded vector denoting 1 for presence and 0 for the absence of the word in the particular document.

The system considers Bag of Words (BoW) embedding approach which forecasts the likelihood of a word given in a context. BoW is implemented to convert the extracted aspect terms into vector format which is fed as input to the Aspect Category Classifier. Being probabilistic, BoW commonly performs superior to other methods and does not have high memory requirements.

4.3 POS-Tagging

Part-of-Speech (POS) tagging is an approach of assigning one of the parts of speech like nouns, verbs, adverbs, etc. including their sub-categories, to the respective word in a text (corpus) based on both its meaning and its context — i.e., its association with adjacent and associated terms in a phrase, sentence, or paragraph.

Spacy's POS tagger is used to particularly recognize adjectives and adverbs as sentiment terms from reviews which will be fed as input to the Sentiment Category Classifier.

5. VISUALIZATION

Data visualizations like charts and graphs provide businesses an approach to deliver essential information at a glance. A stunning and comprehensible visualization format is needed to highlight crucial textual data points.

5.1 Word Cloud

A word cloud can compose dull data engrossing and instantly convey relevant information. Word cloud is an image comprising words used in a particular document in

which the size of the individual words symbolizes its frequency or importance. So, the often a specific word appears in the text, the prominently it will appear in the word cloud.

Word clouds are used for analyzing consumer feedback enabling us to understand what customers like or dislike



Fig-2: Aspect terms Word Cloud

about our business. Word clouds are applied for visualizing aspect terms i.e. noun chunks like food, taste, price, ambiance, etc. to highlight the most used terms for the business owner to summarize the service's opinion and feedback efficiently.

5.2 Bar Graph

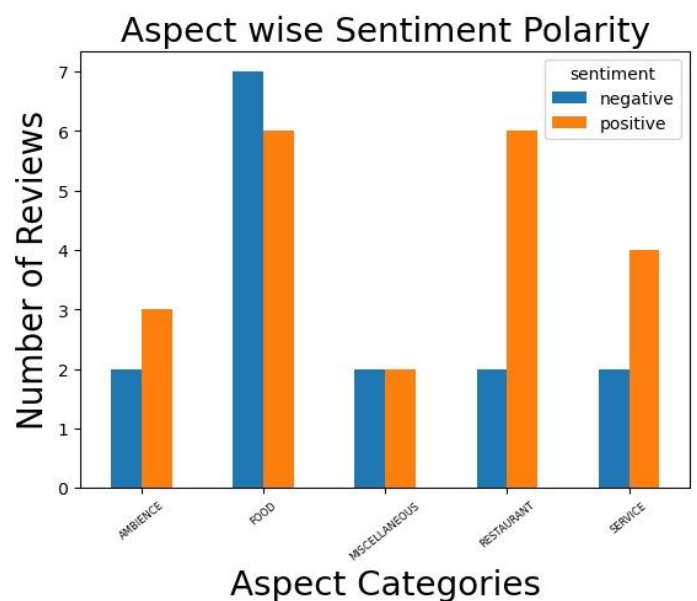


Fig-3: Aspect Categories vs Number of Reviews Bar Graph

A double bar graph is employed to show the number of reviews entered proportional to their respective aspect category describing positive or negative opinions of that feature. This will benefit the restaurant owner to see the number of positive or negative reviews for restaurant, food, ambiance, service, and miscellaneous aspect categories furthermore making him/her understand the good or bad outcomes of the business.

6. CONCLUSION

Therefore, we have proposed an Aspect based Sentiment Analysis model using CNN that consists of Aspect Categorization Model and Sentiment Category Classifier. Techniques like POS Tagging, Dependency parser and Bag of Words are used and to extract sentiment and aspect information from new incoming reviews. The combination of these models generates visualized output in the form of Word cloud and Bar graph.

This ABSA system is salable and benefits the business owner to scrutinize the reviews and distinguish mistakes and repair them. This proposed model helps in the enhancement of business strategies and business growth. In future work, more specific divisions of aspects and sentiments can be considered. Analyzing similes, sarcasm and emoticons can similarly be taken into account as they further express strong opinions regarding restaurant reviews.

7. REFERENCES

- [1] I. K. C. U. Perera and H. A. Caldera, "Aspect based opinion mining on restaurant reviews," *2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCI)*, Beijing, 2017, pp. 542-546, doi: 10.1109/CIAPP.2017.8167276
- [2] More, Pratima & Ghotkar, Archana. (2016). A Study of Different Approaches to Aspect-based Opinion Mining. *International Journal of Computer Applications*. 145. 11-15. 10.5120/ijca2016910712.
- [3] A. S. Shafie, N. M. Sharef, M. A. Azmi Murad and A. Azman, "Aspect Extraction Performance with POS Tag Pattern of Dependency Relation in Aspect-based Sentiment Analysis," *2018 Fourth International Conference on Information Retrieval and Knowledge Management (CAMP)*, Kota Kinabalu, 2018, pp. 1-6, doi: 10.1109/INFRKM.2018.8464692.
- [4] N. Jihan, Y. Senarath and S. Ranathunga, "Aspect Extraction from Customer Reviews Using Convolutional Neural Networks," 2018 18th International Conference

on Advances in ICT for Emerging Regions (ICTer), Colombo, Sri Lanka, 2018, pp. 215-220, doi: 10.1109/ICTER.2018.8615575.

- [5] Florian Heimerl, Steffen Lohmann, Simon Lange, Thomas Ertl, "Word Cloud Explorer: Text Analytics based on Word Clouds", 2014 47th Hawaii International Conference on System Science.
- [6] Suresh Manandhar, Mohammad AL-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, Véronique Hoste, Marianna Apidianaki, Xavier Tannier, Natalia Loukachevitch, Evgeny Kotelnikov, Nuria Bel, Salud María Jiménez-Zafra, Gülşen Eryiğit, "SemEval-2016 Task 5: Aspect Based Sentiment Analysis".
- [7] Yamashita, R., Nishio, M., Do, R.K.G. *et al.*, "Convolutional neural networks: an overview and application in radiology", *Insights Imaging* **9**, 611-629 (2018).
- [8] Wei Wang, Jianxun Gang, "Application of Convolutional Neural Network in Natural Language Processing", 2018 International Conference on Information Systems and Computer Aided Education (ICISCAE)
- [9] Richa Sharma, Shweta Nigam and Rekha Jain, "OPINION MINING OF MOVIE REVIEWS AT DOCUMENT LEVEL", *International Journal on Information Theory (IJIT)*, Vol.3, No.3, July 2014.
- [10] : Fernando Enríquez, Jose A. Troyano, Tomás López-Solaz, An approach to the use of word embeddings in an opinion classification task, *Expert Systems With Applications* (2016), doi: 10.1016/j.eswa.2016.09.005

BIOGRAPHIES



Vedashree Paithankar

Student at Dept. of Information Technology, P.E.S.'s Modern College of Engineering, Pune-05, Maharashtra, India



Shreya Patil

Student at Dept. of Information Technology, P.E.S.'s Modern College of Engineering, Pune-05, Maharashtra, India

**Vinita Doiphode**

Student at Dept. of Information Technology, P.E.S.'s Modern College of Engineering, Pune-05, Maharashtra, India

**Jithisha Koshy**

Student at Dept. of Information Technology, P.E.S.'s Modern College of Engineering, Pune-05, Maharashtra, India

**Ashwini Bhamre**

Asst. Professor at Dept. of Information Technology, P.E.S.'s Modern College of Engineering, Pune-05, Maharashtra, India

**Mukta Kore**

Asst. Professor at Dept. of Information Technology, P.E.S.'s Modern College of Engineering, Pune-05, Maharashtra, India