International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-4, November 2019

Efficient Aspect Finding in Sentiment Analysis using Optimal Binary Search Tree

K. Chitra, A. Tamilarasi, T. Kavitha, S. Hemalatha

Abstract— Sentiment Analysis is an interesting research area. It useful for analyze the user comments with the help of natural language processing. While consider a huge data and improper data, the sentiment classification with better accuracy is a very big challenge in sentiment analysis. The social media records relevant to Sentiment analysis is a difficult one to identify with a people's situation, thoughts, and view in the direction of a certain event, which has a number of applications such as election guess and product assessment. This paper focuses on the basic idea of optimal binary search Tree Technique to find the optimal solution (aspect) from a sentence among the several aspects; specifically, we perform the classification process appropriate to various features, which is used to identify the best one aspect among the several aspect terms in a sentence. In the proposed system, We have introduced optimal binary search Tree aspect level sentiment analysis technique. Through this we are getting best and efficient aspect level solution compare to General Tree. The main novelty of the paper is analyzing how aspect sentiment extracted from each word in a sentence through Dictionary based approach with help of the syntactic Patterns. It provides a well understanding of the appropriate role of every words to obtain optimal Aspect in an efficiency way.

Keywords: Sentiment classification, Aspect level Sentiment Analysis, Supervised Technique, Optimal Binary Search Tree Technique.

I. INTRODUCTION

Sentiment classification acts as an important task in daily life, in political aspect, commodity production, and commercial activities. In Sentiment analysis, identify a result for the accurate and timely classification of emotion is a very tough task.

Generally, a human facing different kinds of feelings or emotions such as happy, sad, fear, anger, etc in their daily life. This will depend on human situations in a particular period. In terms of computers, these feelings or emotions are called as sentiments. The sentiments are extracted from customer feedback about a particular entity are called as sentiment detection or emotion detection. Then classify those sentiments belongs to what category such as best, average or worst case is referred as sentiment classification. Sentiment Classification is used to classifying the text in accordance with sentimental information associated with the text. The proposed paper discussed about a sentiment analysis along with aspect category. It is the process of classifying sentiment polarity with respect to a collection of aspects in a sentence. The difficult task of Aspect Based Sentiment Analysis is how to effectively retrieve the particular aspects from the whole sentence. The aspect scenario can be composed of different tasks such as Extraction of aspect from the sentence and classify those aspects sentiment classification (Liu, 2012).

1.1 Aspect Extraction

The process of aspect extraction is to find aspects that have been evaluated. Consider the following sentence case.

"The Lifetime of this phone Battery is very long".

In above sentence, the aspect is Battery Lifetime and the phone represented as an object. There are many methods are there to extract the opinion. In which most important or leading method is a Sequential strategy because these are under the supervised techniques. By using this technique, we perform the training with the help of manually labeled data. That means in a corpus, one needs to manually annotate aspects and no - aspects. Generally, the target of the aspect extraction is to extract the topic from a sentence. The below figure represents the flow how the aspect is extracted from the sentence and achieve optimal aspect.

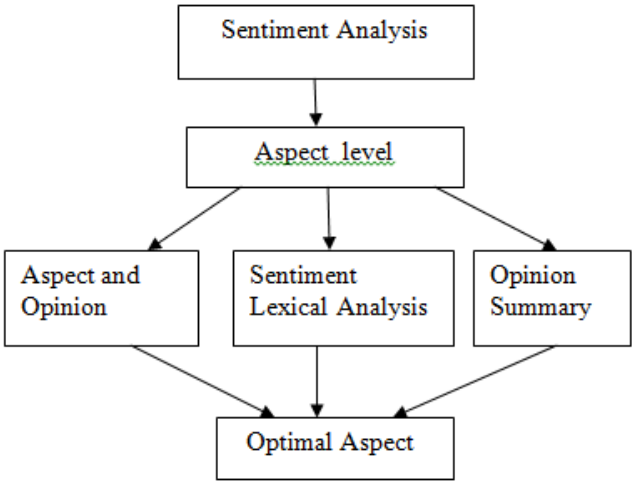


Fig.1 Aspect extraction flow Diagram

Manuscript published on November 30, 2019.
* Correspondence Author
Mrs. K. Chitra, Assistant Professor(s), Department of Computer Applications, Kongu Engineering College, Perundurai, Tamilnadu, India.
Dr. A. Tamilarasi, Professor and Head, Department of Computer Applications, Kongu Engineering College, Perundurai, Tamilnadu, India.
Mrs. T. Kavitha, Assistant Professor(s), Department of Computer Applications, Kongu Engineering College, Perundurai, Tamilnadu, India.
Mrs. S. Hemalatha, Assistant Professor(s), Department of Computer Applications, Kongu Engineering College, Perundurai, Tamilnadu, India.

1.2 Aspect Sentiment Analysis

The task of Aspect sentiment analysis is to find out whether the opinions on different aspects belong to several cases like best, worst, or average. To determine these tasks we need classification approaches. This approach can have two categories like supervised learning and lexicon-based approach. In a Supervised learning case, we need a two set of documents one for training set and another for a test set to perform the classification process. In a Lexicon-based technique, classification process is carried out by making the difference between the features of a given text against sentiment lexicons whose values are priorly determined. Sentiment lexicon having list of words and expressions used to express the user opinions. It does not need training data priorly in order to classify the data. The main problem for learning methods is to finding the scope of every sentiment expression, whether the sentence may holding the aspect or not. While doing the classification process, a set of aspect dependent features are generated through dependency parser for predicting accuracy. In a lexicon-based approach, sentiment phrases as the focused resource. This method (Ding et al., 2008) has following steps to allot a polarity to an aspect: Notify the sentiment words and phrases, apply sentiment shifters, aggregate opinions using an aggregation function (e.g. Hu and Liu (2004)).

Example: The food is tasty, cost is normal but the environment is not good.

In the above sentence, there are three aspects such as food, cost and environment. Food related positive aspect, Cost related to neutral and environment related to negative aspect respectively. An Aspect term is usually referred an entity.

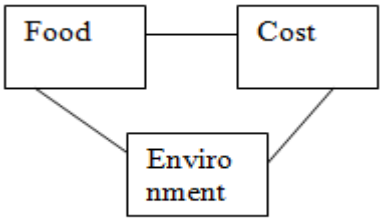


Fig. 2. Sentiment Dependencies Between Multiple Aspects in a one sentence.

1.3 Examples: Analyzing Positive And Negative Aspect Level Sentiment Analysis

Example 1: This saree is expensive one but not that much attractive. This sentence belongs to negative opinion because of high cost and not that much look. Here, Cost and attractive is belongs to aspect category.

Example 2: Mary not gives that much attention to the movie because of noisy environment. This sentence belongs to neutral opinion. Where, the environement is an aspect.

Example 3: I know she is a having good characteristics. This sentence belongs to positive opinion because of good characteristics. Here, character is an aspect.

Example 4: I dislike all malpractice activities

This sentence belongs to negative case. Here, malpractice is an aspect.

II. RELATED WORK

The article Hai HaDo [2], et al [2019] elucidated sentiment analysis specifically for aspect case and to provide better result of granularity, It concentrates on two things like retrieving extraction and perform classification of entity reviews related to objective-dependent tweets. In this article, they have proposed a comparative study of Deep learning approaches for aspect-oriented sentiment analysis to place various approaches in perspective. They were preferred SemEval 2014-2016 Dataset and analyze the DNN model Performance in aspect category.

The article Mikhaylov [3], et al [2019] elucidated aspect-based sentiment analysis using machine learning methods. This paper have discussed about an example method, dictionary-based approach, syntactic along with semantic analysis of natural language text to produce a plurality of syntactic-semantic structures representing the part of the natural language text and do the evaluation using one or more text characteristics produced by the syntactic-semantic analysis, use a classifier function to determine polarities associated with one or more aspect terms and generating a report.

The article Jean Maillard [5], et al [2019] elucidate a model related to CKY chart parser, and analyze its downstream performance on a natural language inference task and a reverse dictionary process. Again, they explained how its performance can be extended with attention mechanism, which completely exploits the parse chart, by attending over all possible subspans of the sentence. They found that their approach is competitive against same models of comparable size and outperforms Tree-LSTMs that use trees produced by a parser. Finally, they have an another architecture based on a shift-reduce parser. They perform tree analysis induced by both their models, to examine whether they are consistent with each other and across re- runs, and whether they resemble the trees produced by a standard parser.

The article SoujanyaPoria [4][18] et al [2019] elucidated a new pattern to concept oriented sentiment analysis that combines several process like linguistics, common-sense computing, and machine learning for producing the better accuracy related to predicting appropriate polarity. The feelings are carried out by a concept basis and identift the dependency relation of the input sentence. Specifically, those achieve a well understanding of the contextual role of each and every concept within the sentence. consequently, obtain a polarity detection engine that well performs statistical methods.

The article Pinlong Zhao [21], et al [2019] elucidated a sentiment classification model related to graph convolution networks that can effectively notify the sentiment dependencies between many aspects in a single sentence. This model introduced bidirectional attention mechanism along with position encoding to reproduce aspect specific representations between each aspect and its context words, then using GCN technique to extract the sentiment dependencies between different aspects in a sentence. Then they have to evaluate the system on the SemEval 2014 datasets and produced the effective result.

The article **Xiao Ma** [19] et al [2019], elucidated a model which can be divided in two steps: (1) the first one introduces position attention mechanism to model the explicit position context between the aspect and its context words with the aim of dealing with aspects an order and the second model denoted as how to model multi-aspects within one opinionated sentence all at once by applying position attention mechanism. The system empirically assess the proposed method on the SemEval 2014 datasets and hopeful experimental output turn out that the proposed approach obtained efficient performance compared to some other attention-based methods.

III. PROPOSED METHODOLOGY

3.1 Supervised Techniques (Dictionary Based Approach)

There are various supervised algorithms are available to extract the opinion from social medial websites. In our proposed methodology, we applied Dictionary based approach to extract the opinion words from the customer feedback as a sentence format. Then the aspects are identified from that sentence with help of syntactic patterns rules.

3.2 Tree Based Sentiment analysis

There are so many researches are doing their research on sentiment analysis including tree concept. Here we focused optimal binary search tree based aspect level sentiment analysis to obtain the optimal solution dynamically.

Figure 2a shows the tree representation of a one sentence called as collection of words.

Each node in the AST is an abstract component in a sentence. A node N with collection of Children nodes such as n_1 , n2 ⋯ , n_m represents the constructing process of the component that denoted as $R \rightarrow c_1 \cdots c_m$.

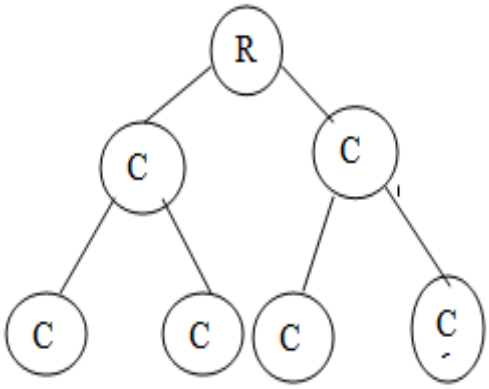


Fig.3 Tree Representation of a one sentence

3.3 Optimal Binary Search Tree

Optimal binary search tree is a one kind of binary search Tree. While we construct a binary search tree, there are so many possibilities are available. Among the several possibilities, we have to find the optimum one for producing effective and efficiency searching solution. For this reason, we have introduced OBSTAS algorithm to find the optimal aspect from the sentence collected from the customer feedback relevant to restaurant data.

Consider the following Aspect related to restaurant data.

Aspect 1: Environment (W)

Aspect 2: Food Quality (X)

Aspect 3: Service (Y)

Aspect 4: Cost (Z)

The above stated are the different aspects. From these aspects, we have constructed binary search Tree. There are several possibilities are obtained and find the optimal one among them.

Fig 4. Possibilities of binary search tree (Stated few here)

Fig. 4a

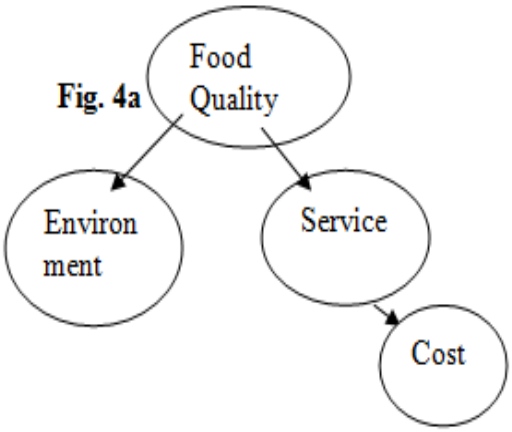


Fig. 4b

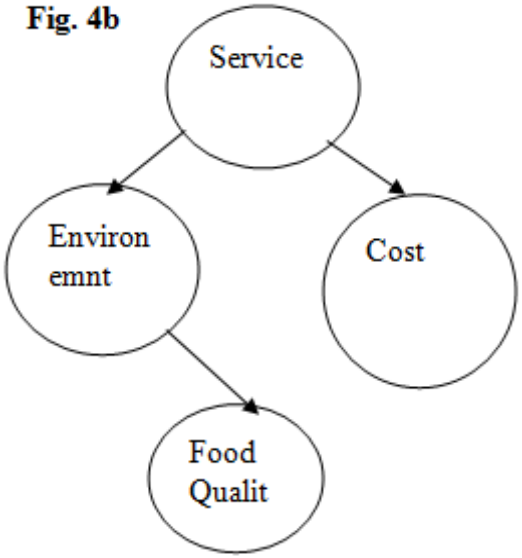
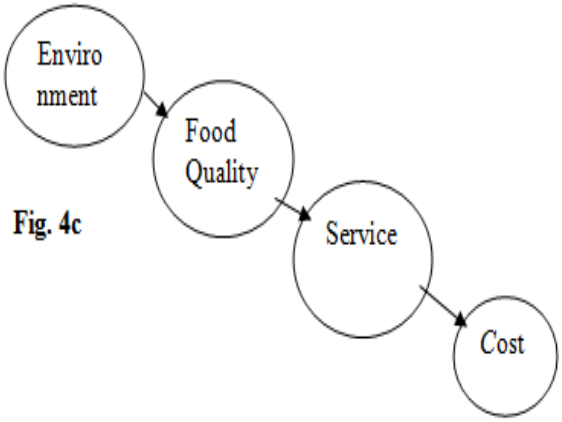


Fig. 4c



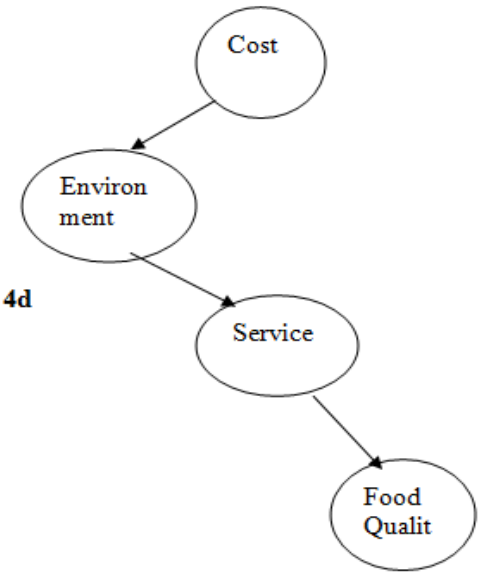


Fig. 4d

3.3.1 Formula for finding optimal solution

1. Cost[Q,R] = lowest{Cost[Q, m-1] + Cost[m+1, j]} + Summation of Q to R probabilities.
2. Cost [Q,Q] = probabilities of Q for 1<=Q<=n.
3. Cost [Q, Q-1] = zero for 1<=Q<=n+1.

BY applying the second formula, we have make entry for all the aspects value relevant to corresponding row and column (key)

By applying third formula, we have an diagonal representation (value is 0).

By applying first formula, we have to find the optimal solution.

3.3.2 Optimal Binary Search Tree - Aspect sentiment Analysis Algorithm (OBSTAS Algorithm)

OBSTAS(Aspect, p ,c)

Input : Aspects to insert in BST, probability value for each keys, number of aspects.

Assign the each aspect to different variable.

Output: lowest cost to make optimal Binary search Tree to predict best aspect.

```
Begin
  Define cost matrix of size c x c
  for Q:= 0 to c-1, do
    cost [Q, Q] := p[Q]
  done
  for length :=  2 to c, do
    for Q:= 0 to (c-length+1), do
      R := Q + length – 1
      cost[Q, R] := ∞
      for m in range Q to R, done
        if m > Q, then
          c1 := cost[Q, m-1]
        else
          c1 := 0
        if m < R, then
          c1 := c1 + cost[m+1, R]
        c1 := c1 + sum of p from Q to R
        if c1 < cost[Q, R], then
          cost[Q, R] := c1
      done
    done
  done
  return cost[0, c-1]
End
```

IV. DATASET & EXPERIMENTAL RESULT

We have taken restaurant data for the experimental purpose. The data's are collected from kaggle websites related to restaurant dataset. From the customer feedback, we have to find the positive probability value for each aspect by taking all customer details. Then apply the optimal binary search tree algorithm for finding best aspect among the several possibilities of other aspects like Environment, Food Quality, Service, and Cost.

Data	Aspect	**Positive**	Neutral	Negative
Restaurant	Environment	**2000**	5000	3000
	Food Quality	**8000**	1000	1000
	Service	**5000**	3000	2000
	Cost	**7000**	2000	1000

Table 1: Root Table Construction

	0	1	2	3	4
1	0	1	2	2	**2 (Root)**
2		0	2	3	3
3			0	3	4
4				0	4
5					0

Result

The Optimal Aspect of given positive Case is X.

OBST Positive Aspect output Representation look like this according to the Table 1.

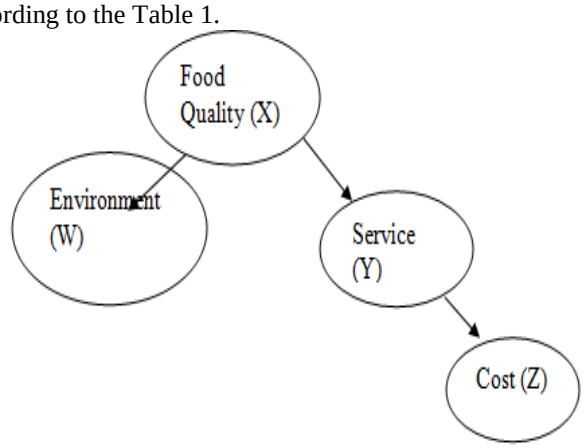


Fig. 5

V. CONCLUSION

In this proposed paper concentrated on various entities relevant to restaurant data by applying tree based sentiment analysis. The main motivation of this paper is to identify the optimal aspects in one sentence among the several aspects. The datasets are collected from the social media websites relevant to restaurant and then the Customer feedback collected. From the feedback, we have to extract the aspect using supervised technique. Then constructed the binary search Tree for the extracted aspect along with the positive probability value for each aspect. Among the several positive aspect, finally we found the (minimum cost) optimal aspect by applying optimal binary search tree algorithm for improving searching efficiency. In future, we improve this work along with deep learning Techniques to obtain the better performance.

REFERENCES

1. RaniaOthman, RamiBelkaroui, RimFaiz "Extracting product features for opinion mining using public conversations in twitter", Procedia Computer Science, Volume 112, Pages 927-935, 2017.
2. Hai Ha Do, P.W.C Prasad, Abeer Alsadoon "Deep Learning for Aspect-Based Sentiment Analysis:A Comparative Review", Published in Expert Systems with Applications, Volume 118, Pages 272-299, 15 March 2019.
3. Maksim Borisovich Mikhaylov, Konstantin Alekseevich Pasechnikov, "Aspect-based sentiment analysis and report generation using machine learning methods, United States Patent, Patent No. : US 10, 198, 432, B2, February 5, 2019.
4. SoujanyaPoria, ErikCambria, GrégoireWinterstein, Guang-BinHuang, "Sentic patterns: Dependency-based rules for concept-level sentiment analysis", Knowledge Based Systems,Volume 69, Pages 45-63, October 2014.
5. Jean Maillard , Stephen Clark and Dani Yogatama , "Jointly learning sentence embeddings and syntax with unsupervised Tree-LSTMs", Natural Language Engineering, Volume 25, Issue 4, Pages. 433-449 July 2019.
6. Jaspreet Singh, gurvinder Singh & Rajinder Singh, "Optimization of sentiment analysis using machine learning classifiers", SpringerOpen, Human-centric Computing and Information Sciences volume 7, Article number: 32, Published: 11 December 2017.
7. KumarRavi, VadlamaniRavi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications", Knowledge Based Systems, Volume 89, Pages 14-46, November 2015.
8. Basant Agarwal, Soujanya Poria, Namita Mittal, Alexander Gelbukh, Amir Hussain, "Concept-Level Sentiment Analysis with Dependency-Based Semantic Parsing: A Novel Approach" Springer, Volume 7, Issue 4, Pages 487–499, August 2015.
9. T. H. Nguyen and K. Shirai, "PhraseRNN: Phrase recursive neural network for aspect-based sentiment analysis,", Volume:Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing., Pages 2509–2514, September 2015.
10. Tun Thura Thet, Jin-Cheon Na, Christopher S.G. Khoo, "Aspect-based sentiment analysis of movie reviews on discussion boards", Journal of Information Science, Volume 36, Issue 6, November 15, 2010.
11. K.G. Nandakumar, Dr.T.Christopher "Opinion Mining: A Survey", International Journal of Computer Applications (0975 –8887), Volume 113, No.2, March 2015.
12. Hu han, Jin liu, and Guoli liu, "Attention-Based Memory Network for Text Sentiment Classification", IEEE Access, Volume 6, 2018.
13. Muhammaad Zubair, Aurangzeb Khan, Shakeel Ahmad, FazalMasudKundi and Asghar, "A Review of Feature Extraction in Sentiment Analysis", Journal of Basic and Applied Scientific Research, ISSN 2090-4304, 4(3)181-186, 2014.
14. WalaaMedhat, AhmedHassan, HodaKorashy, "Sentiment analysis algorithms and applications: a survey", Ain Shams Engineerin Journal, Volume 5, Issue 4, Pages 1093-1113, December 2014 .
15. Yong Zhang, Meng Joo Er , Rui Zhao ; Mahardhika Pratama, "Multiview Convolutional Neural Networks for MultidocumentExtractive Sumarization", IEEE Transactions on Cybernetics, volume. 47, No. 10, October 2017.
16. Yoon Kim, "Convolutional neural networks for sentence classification," in Proc. Conf. Empirical Methods Natural Lang. Process., Doha, Qatar, Pages. 1746_1751, 2014.
17. Chen Zhang, Qiuchi Li, Dawei Song, "Syntax-Aware Aspect Level Sentiment Classification with Proximity-Weighted Convolution Network", Published in SIGIR'19 Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval", Pages 1145-1148, ISBN: 978-1-4503-6172-9, July 21 - 25, 2019.
18. Soujanya Poria, Nir Ofek, Alexander Gelbukh, Amir Hussain, Lior Rokach, Dependency Tree-Based Rules for Concept-Level Aspect-Based Sentiment Analysis", SpringerLink, CCIS, volume 475, Pages 41-47, 04 October 2014.
19. 19.Xiao Ma, JiangfengZeng, LimeiPeng, GiancarloFortino, YinZhang, "Modeling multi-aspects within one opinionated sentence simultaneously for aspect-level sentiment analysis, Elsevier,Future Generation Computer Systems, Volume 93, Pages 304-311XiaoMa, April 2019.
20. 20.LiJianqiang, FuXianghua, M.A.Masud, Huang, Joshua Zhexue, "Learning distributed word representation with multi-contextual mixed embedding," Knowledge Based System, Volume. 106, Pages 220 - 230, 15 August 2016.
21. Pinlong Zhaoa, Linlin Houb, Ou Wua, Modeling Sentiment Dependencies with Graph Convolutional Networks for Aspect-level Sentiment Classification, Computer Science, Computation and Language, Volume, 11 June 2019.