Journal of Engineering Sciences
Vol 10, Issue 11, Nov /2019
ISSN NO:0377-9254

AN APPROACH OF SENTIMENT ANALYSIS BASED ON CATEGORY DETECTION AND DEAL WITH IMBALANCED DATA USING MACHINE LEARNING TECHNIQUES

M. JANARDHAN[1], M. RUFUS KUMAR[2]

[1]Asst. Professor, Dept. of CSE, PBR VITS, Kavali, A.P, India.

[2]M.Tech, Dept. of CSE, PBR VITS, Kavali, A.P, India.

Abstract – In this paper we consider online consumer reviews to assist purchase-decision making has become increasingly popular. To process the user reviews and find the useful information for making decision of purchase most of existing systems are presented. But one can hardly read all reviews to obtain a fair evaluation of a product or service. A subtask to be performed by such a framework would be to find the general aspect categories addressed in review sentences, for which this paper presents two methods. the first method presented is an unsupervised method that applies association rule mining on co-occurrence frequency data obtained from a corpus to find these aspect categories. While not on par with state-of-the-art supervised methods, the proposed unsupervised method performs better than several simple baselines, a similar but supervised method, and a supervised baseline, with an F1-score of 67%. The second method is a supervised variant that outperforms existing methods with an F1-score of 84%. And also we extend our work to deal with the imbalanced dataset using Synthetic Minority Over-sampling Technique and Modified Synthetic Minority Over-sampling Technique further improving the performance of the system.

Keywords — Sentiment Analysis, Co-occurrence Data, imbalanced Data.

I. INTRODUCTION

Word of mouth (WoM) has always been influential on consumer decision-making. Family and friend are usually asked for advice and recommendations before any important purchase-decisions are made. These recommendations can both have short as well as long term influence on consumer decision-making [1].

With the Web, WoM has greatly expanded. Anyone who wishes to share their experiences can now do so electronically. Social media, like Twitter and Facebook allow for easy ways to exchange statements about products, services, and brands. The term for this expanded form of WoM is electronic WoM (EWoM).

Over the last few years, EWoM has become increasingly popular [2]. One of the most important forms of EWoM communication are product and service reviews [3] posted on the Web by consumers. Retail companies such as Amazon and Bol have numerous reviews of the products they sell, which provide a wealth of information, and sites like Yelp offer detailed consumer reviews of local restaurants, hotels, and other businesses. Research has shown these reviews are considered more valuable for consumers than market-generated information and editorial recommendations [4]–[6], and are increasingly used in purchase decision-making [7].

The information that can be obtained from product and service reviews is not only beneficial to consumers, but also to companies. Knowing what has been posted on the Web can help companies improve their products or services [8].

However, to effectively handle the large amount of information available in these reviews, a framework for the automated summarization of reviews is desirable [9]. An important task for such a framework would be to recognize the topics (i.e., characteristics of the product or service) people write about. These topics can be fine-grained, in the case of aspect-level sentiment analysis, or more generic in the case of aspect categories.

As one can see, aspect categories are usually implied, that is, the names of the categories are not explicitly mentioned in the sentence. The same holds for fine-grained aspects: while most of them are referred to explicitly in a sentence, some are only implied by a sentence. For example, in the sentence below, the implied fine-grained aspect is “staff,” whereas the implied aspect category is “service.”

When the aspect categories are known beforehand, and enough training data is available, a supervised machine learning approach to aspect category detection is feasible, yielding a high performance [11]. Many approaches to find aspect categories are supervised [11]–[14]. However, sometimes the flexibility inherent to an unsupervised method is desirable.

In this paper, both an unsupervised and a supervised method are proposed that are able to find aspect categories based on co-occurrence frequencies. The unsupervised method uses spreading activation on a graph built from word co-occurrence frequencies in order to detect aspect categories. In addition, no assumption has to be made that the implicit aspects are always referred to explicitly, like it is done in [15]. The proposed unsupervised method uses more than just the literal category label by creating a set of

explicit lexical representations for each category. The only required information is the set of aspect categories that is used in the data set. The supervised method on the other hand uses the co-occurrences between words, as well as grammatical relation triples, and the annotated aspect categories to calculate conditional probabilities from which detection rules are mined.

And also we extend our work to deal with imbalanced data using Synthetic Minority Over-sampling Technique (SMOTE). Further improving performance we modified SMOTE.

II. RESEARCH METHOD

A) Unsupervised Method

The proposed unsupervised method (called the spreading activation method) uses co-occurrence association rule mining in a similar way as [15], by learning relevant rules between notional words, defined as the words in the sentence after removing stop words and low frequency words, and the considered categories. This enables the algorithm to imply a category based on the words in a sentence. To avoid having to use the ground truth annotations for this and to keep this method unsupervised, we introduce for each category a set of seed words, consisting of words or terms that describe that category.

These words or terms are found by taking the lexicalization of the category, and its synonyms from a semantic lexicon like WordNet. For example, the ambience category has the seed set {ambience, ambiance, atmosphere}. With the seed words known, the general idea of implicit aspect detection can be exploited to detect categories as well. The idea is to mine association rules of the form [notional word → category] from a co-occurrence matrix. Each entry in this co-occurrence matrix represents the frequency degree of two notional words co-occurring in the same sentence. Stop words, like the and and, as well as less frequent words are omitted because they add little value for determining the categories in review sentences.

The reason why we choose to mine for rules similar to that of [15]'s, and do not consider all notional words in the sentence at once to determine the implied categories, like [21], is based on the hypothesis that categories are better captured by single words. If we have for example categories like food and service all it takes to categorize sentences is to find single words like chicken, staff, or helpful.

Association rules are mined when a strong relation between a notional word and one of the aspect categories exists, with the strength of the relation being modeled using the co-occurrence frequency between category and notional word.

We distinguish between two different relation types: 1) direct and 2) indirect relations. A direct relation between two words A and B is modeled as the positive conditional probability P(B|A) that word B is present in a sentence given the fact that word A is present. An indirect relation between two words A and B exists when both A and B have a direct relation with a third word C. This indicates that A and B could be substitutes for each other, even though their semantics might not be the same. Without checking for indirect relations, substitutes are usually not found since they do not co-occur often together. A visual example of an indirect relation can be found in Fig. 1.

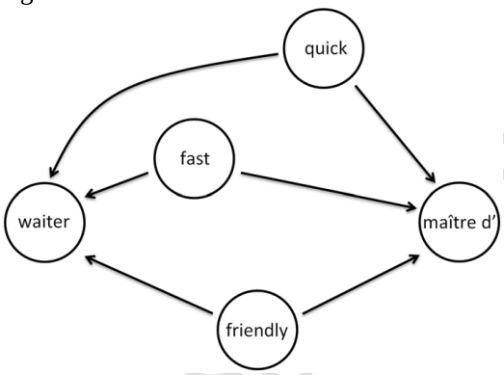

Fig. 1: Example of an indirect relation

To exploit the direct, as well as the indirect relation information between notional words and seed words, the spreading activation algorithm is utilized, which is a method to search for associative networks. Spreading activation has been successfully applied in various fields. For that, a network data structure is needed, consisting of vertices connected by links, as depicted in Fig. 1. The vertices are labeled and the links may receive direction and/or weights to model the relations between vertices. The search process of finding an associative network is initiated by giving each vertex an activation value. These initial values determine the area of the search as the activation values are iteratively spread out to other, linked, vertices.

In our case we want to use spreading activation to find, for each category, a network of words associated with the category's set of seed words. To do this, a network data structure is created, having vertices for all notional words and edges to model the direct relations between these words. In the network data structure all notional words receive an initial activation value of zero except for the category's seed words, which receive positive activation values. In the first iterative step of the spreading activation algorithm, these positive activation values are spread out to other words directly related to the seed words, based on the strength of the direct relation. In this way, words that have strong direct relations with the seed words receive high association values. The following iterative steps will be looking for words with high association values that are then activated and will spread out their activation value to other words directly related to them. In this way, notional words that are indirectly related to one of the seed words are also identified. The end result will be a network of notional words, each with their own activation value, the higher the activation value, the

more related the notional word will be to the category.

The data network structure used for the spreading activation algorithm will consist of vertices that represent the notional words, and links between two vertices representing a strictly positive co-occurrence frequency. Each link represents the direct relation between two notional words and receives weight equal to the conditional probability that word A co-occurs with word B, given that B appears in a sentence. This also means that the links receive direction as the conditional probability is not symmetric, making the data network structure a co-occurrence digraph.

Once each category has its own associative network, rules can be mined of the form [notional word → category] from vertices in these networks, based on the activation value of the vertex. Since the same word can be present in multiple associative networks, one word might trigger multiple aspect categories. Based on the words in the sentence, a set of rules is triggered and their associated aspect categories are assigned to the sentence. Fig. 2 illustrates how the unsupervised method works on a simple example corpus, with a decay factor of 0.9 and firing threshold of 0.4. The example shows how an associative network for the category food is found and rules are extracted.

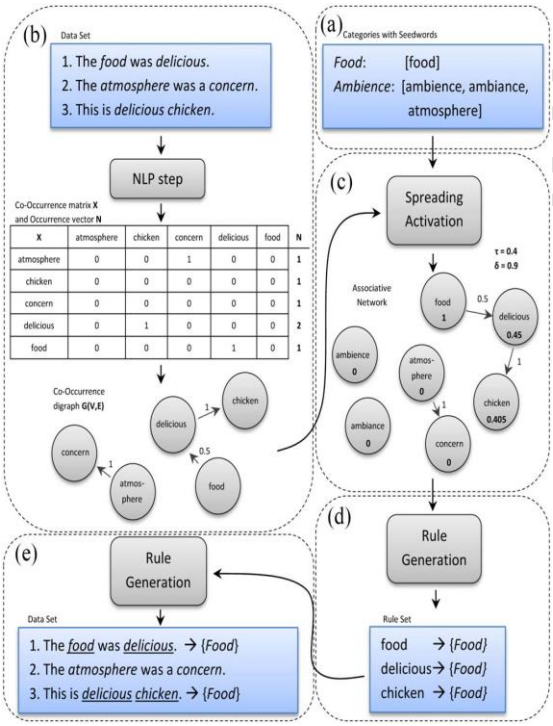

Fig. 2: Example flowchart of the unsupervised method.

i) Algorithm

The method can best be described according to the Algorithm 1.

Algorithm 1: Spreading Activation Algorithm

input : category c
input : vertices V
input : seed vertices S_c
input : weight matrix W
input : decay factor δ
input : firing threshold τ_c
output: activation values $A_{c,i}$ for category c

```
foreach s ∈ S_c do
    A_{c,s} ← 1
end
foreach i ∈ V \ S_c do
    A_{c,i} ← 0
end
F ← S_c
M ← S_c
while M ≠ ∅ do
    foreach i ∈ M do
        foreach j ∈ V do
            A_{c,j} ← min{A_{c,j} + A_{c,i} · W_{i,j} · δ, 1}
        end
    end
    M ← ∅
    foreach i ∈ V \ F do
        if A_{c,i} > τ_c then
            add i to F
            add i to M
        end
    end
end
```

B) Supervised Method

Similar to the first method, the supervised method (called the probabilistic activation method) employs co-occurrence association rule mining to detect categories. We borrow the idea from to count co-occurrence frequencies between lemmas and the annotated categories of a sentence. However, low frequency words are not taken into account in order to prevent overfitting. This is achieved using a parameter αL, similar to the unsupervised method. Furthermore, stop words are also removed.

In addition to counting the co-occurrences of lemmas and aspect categories, the co-occurrences between grammatical dependencies and aspect categories are also counted. Similar to lemmas, low frequency dependencies are not taken into account to prevent overfitting, using the parameter αD. Dependencies, describing the grammatical relations between words in a sentence, are more specific than lemmas, as each dependency has three components: 1) governor word; 2) dependent word; and 3) relation type. The added information provided by dependencies, may provide more accurate predictions, when it comes to category detection. Knowing whether a lemma is used in a subject relation or as a modifier can make the difference between predicting and not predicting a category.

Once the conditional probabilities are computed and the thresholds are known, unseen sentences from the test set are processed. For each unseen sentence we check whether any of the lemmas or dependency forms in that sentence have a conditional probability greater than its corresponding threshold, in which case the corresponding category is assigned to that

sentence. Fig. 3 illustrates how the supervised method works on a very simple test and training set.

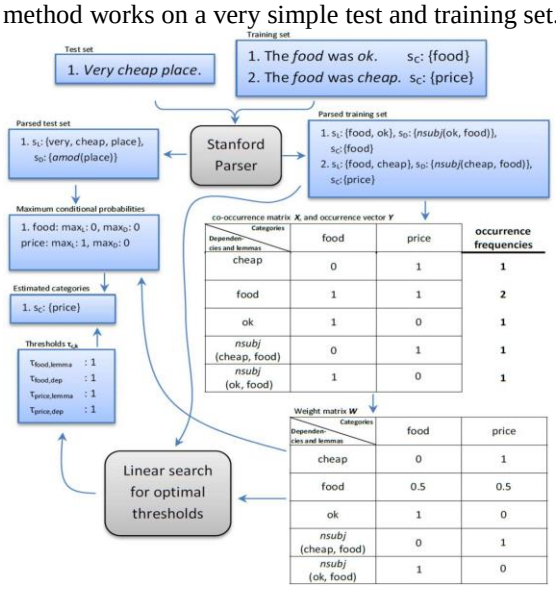

Fig. 3: Example flowchart of the supervised method.

i) Algorithm

The method can best be described according to the following Algorithm 2.

Algorithm 2: Identify Category Set C and Compute Weight Matrix W

input : training set
input : occurrence threshold θ
output: category set C, Weight matrix W

$C, X, Y \leftarrow \emptyset$
foreach *sentence* $s \in$ *Training set* **do**
  // s_k are the lemmas/dependecies of s
  foreach $s_k \in \{s_L, s_{D_1}, s_{D_2}, s_{D_3}\}$ **do**
    foreach *dependency forms/lemmas* $j \in s_k$ **do**
      // count dependency form/lemma occurrence j in Y
      if $j \notin Y$ **then**
        add j to Y
      end
      $Y_j \leftarrow Y_j + 1$
      // s_C are the categories of s
      foreach *category* $c \in s_C$ **do**
        // Add unique categories in category set C
        if $c \notin C$ **then**
          add c to C
        end
        // count co-occurrence (c, j) in X
        if $(c, j) \notin X$ **then**
          add (c, j) to X
        end
        $X_{c,j} \leftarrow X_{c,j} + 1$
      end
    end
  end
end
// Compute conditional probabilities
foreach $(c, j) \in X$ **do**
  if $Y_j > \theta$ **then**
    $W_{c,j} \leftarrow X_{c,j}/Y_j$
  end
end

The final step is to predict the aspect categories for each unseen sentence s ∈ test set. From all lemmas and dependency forms sL, sD1 , sD2 , and sD3 in sentence s we find the maximum conditional probability P(c|j), as described in (4), for each category c ∈ C. Then, if any of these maximum conditional probabilities surpasses their threshold τc,k, category c is assigned as an aspect category for sentence s. The pseudo-code for this step is shown in Algorithm 3.

Algorithm 3: Estimating Categories for the Test Set

input : training set
input : test set
input : occurrence threshold θ
output: Estimated categories for each sentence in the test set

$W, C \leftarrow$ Algorithm 2(Training set, θ)
$\tau_{c,L}, \tau_{c,D_1}, \tau_{c,D_2}, \tau_{c,D_3} \leftarrow$ LinearSearch (Training set, W, C)
// Processing of review sentences
foreach *sentence* $s \in$ *test set* **do**
  foreach *category* $c \in C$ **do**
    // Obtain maximum conditional probabilities $P(c|j) = W_{c,j}$ per type, for sentence s
    $\max_{c,L} \leftarrow \max_{l \in s_L} W_{c,l}$
    $\max_{c,D_1} \leftarrow \max_{d_1 \in s_{D_1}} W_{c,d_1}$
    $\max_{c,D_2} \leftarrow \max_{d_2 \in s_{D_2}} W_{c,d_2}$
    $\max_{c,D_3} \leftarrow \max_{d_3 \in s_{D_3}} W_{c,d_3}$
    if $max_{c,L} > \tau_{c,L}$ or $max_{c,D_1} > \tau_{c,D_1}$ or $max_{c,D_2} > \tau_{c,D_2}$ or $max_{c,D_3} > \tau_{c,D_3}$ **then**
      estimate category c for sentence s
    end
  end
end

In this proposed system we consider the aspect category balanced dataset to achieve the sentiment analysis. In our extension work we deal with imbalanced data for that we explore machine learning techniques such as Synthetic Minority Over-sampling Technique.

C) Extension Work

i) Problem definition

This problem is predominant in scenarios where anomaly detection is crucial like electricity pilferage, fraudulent transactions in banks, identification of rare diseases, etc. In this situation, the predictive model developed using conventional machine learning algorithms could be biased and inaccurate.

This happens because Machine Learning Algorithms are usually designed to improve accuracy by reducing the error. Thus, they do not take into account the class distribution / proportion or balance of classes.

One of the main challenges faced by the utility industry today is electricity theft. Electricity theft is

the third largest form of theft worldwide. Utility companies are increasingly turning towards advanced analytics and machine learning algorithms to identify consumption patterns that indicate theft.

However, one of the biggest stumbling blocks is the humongous data and its distribution. Fraudulent transactions are significantly lower than normal healthy transactions i.e. accounting it to around 1-2 % of the total number of observations. The ask is to improve identification of the rare minority class as opposed to achieving higher overall accuracy.

Machine Learning algorithms tend to produce unsatisfactory classifiers when faced with imbalanced datasets. For any imbalanced data set, if the event to be predicted belongs to the minority class and the event rate is less than 5%, it is usually referred to as a rare event.

Example of Imbalanced Data

Let's understand this with the help of an example.

Ex: In an utilities fraud detection data set you have the following data:

Total Observations = 1000

Fraudulent Observations = 20

Non Fraudulent Observations = 980

Event Rate= 2 %

The main question faced during data analysis is – How to get a balanced dataset by getting a decent number of samples for these anomalies given the rare occurrence for some them?

The conventional model evaluation methods do not accurately measure model performance when faced with imbalanced datasets.

Standard classifier algorithms like Decision Tree and Logistic Regression have a bias towards classes which have number of instances. They tend to only predict the majority class data. The features of the minority class are treated as noise and are often ignored. Thus, there is a high probability of misclassification of the minority class as compared to the majority class.

Evaluation of a classification algorithm performance is measured by the Confusion Matrix which contains information about the actual and the predicted class.

Actual	Predicted	
	Positive	Negative
Positive Class	True Positive (TP)	False Negative (FN)
Negative Class	False Positive(FP)	True Negative(TN)

Accuracy of Model = (TP+TN) / (TP+FN+FP+TN)

However, while working in an imbalanced domain accuracy is not an appropriate measure to evaluate model performance. For eg: A classifier which achieves an accuracy of 98 % with an event rate of 2 % is not accurate, if it classifies all instances as the majority class. And eliminates the 2 % minority class observations as noise.

ii) Synthetic Minority Over-sampling Technique (SMOTE)

This technique is followed to avoid overfitting which occurs when exact replicas of minority instances are added to the main dataset. A subset of data is taken from the minority class as an example and then new synthetic similar instances are created. These synthetic instances are then added to the original dataset. The new dataset is used as a sample to train the classification models.

Total Observations = 1000

Fraudulent Observations = 20

Non Fraudulent Observations = 980

Event Rate = 2 %

A sample of 15 instances is taken from the minority class and similar synthetic instances are generated 20 times

Post generation of synthetic instances, the following data set is created

Minority Class (Fraudulent Observations) = 300

Majority Class (Non-Fraudulent Observations) = 980

Event rate= 300/1280 = 23.4 %

The advantages of SMOTE Mitigates the problem of overfitting caused by random oversampling as synthetic examples are generated rather than replication of instances. No loss of useful information.

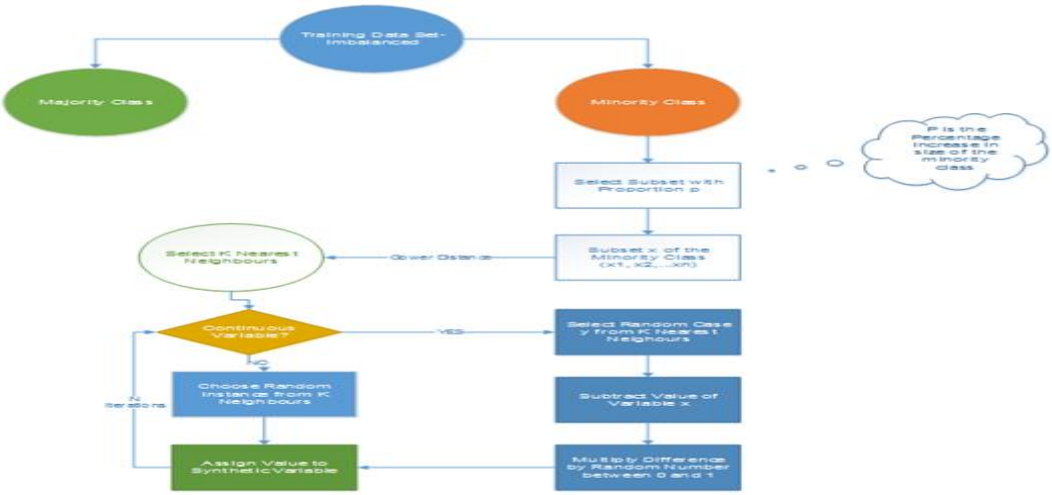

Fig. 4: Synthetic Minority Oversampling Algorithm

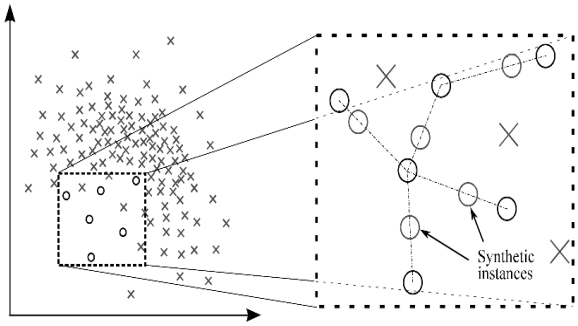

Fig. 5: Generation of Synthetic Instances with the help of SMOTE

iii) Modified SMOTE

It is a modified version of SMOTE. SMOTE does not consider the underlying distribution of the minority class and latent noises in the dataset. To improve the performance of SMOTE a modified method MSMOTE is used.

This algorithm classifies the samples of minority classes into 3 distinct groups – Security/Safe samples, Border samples, and latent nose samples. This is done by calculating the distances among samples of the minority class and samples of the training data.

Security samples are those data points which can improve the performance of a classifier. While on the other hand, noise are the data points which can reduce the performance of the classifier. The ones which are difficult to categorize into any of the two are classified as border samples.

While the basic flow of MSOMTE is the same as that of SMOTE (discussed in the previous section). In MSMOTE the strategy of selecting nearest neighbors is different from SMOTE. The algorithm randomly selects a data point from the k nearest neighbors for the security sample, selects the nearest neighbor from the border samples and does nothing for latent noise.

III. RESULTS ANALYSIS

For the evaluation of the proposed methods, the training and test data from SemEval-2014 [10] are used. It contains 3000 training sentences and 800 test sentences taken from restaurant reviews. Each sentence has one or more annotated aspect categories. Fig. 4 shows that each sentence has at least one category and that approximately 20% of the sentences have multiple categories. With 20% of the sentences having multiple categories, a method would benefit from being able to predict multiple categories. This is one of the reasons why association rule mining is useful in this scenario as multiple rules can apply to a single sentence.

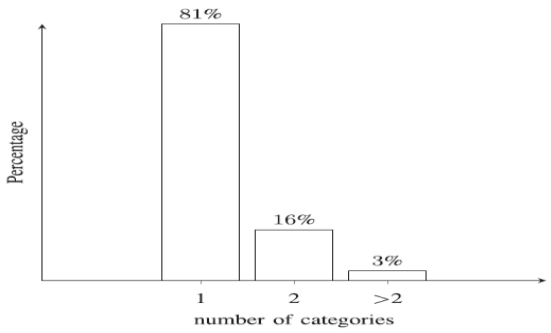

Fig. 4: Distribution of number of aspect categories per sentence.

Fig. 5 presents the relative frequency of each aspect category, showing that the two largest categories, food and anecdotes/miscellaneous, are found in more than 60% of the sentences. This should make these categories easier to predict than the other categories, not only because of the increased chance these categories appear, but also because there is more information about them.

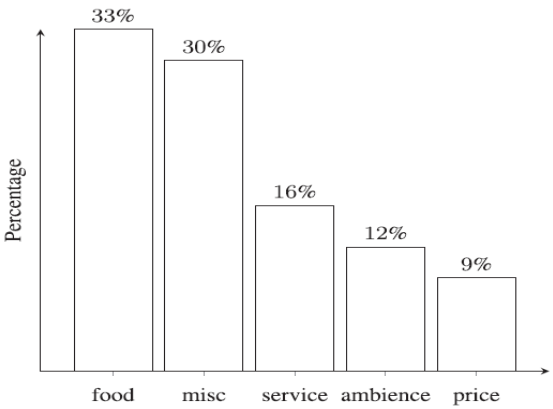

Fig. 5: Relative frequency of the aspect categories.

Last, in Fig. 6, the proportion of implicit and explicit aspect categories is shown. It is clear that using techniques related to implicit aspect detection is appropriate here, given that more than three quarters of the aspect categories is not literally mentioned in the text.

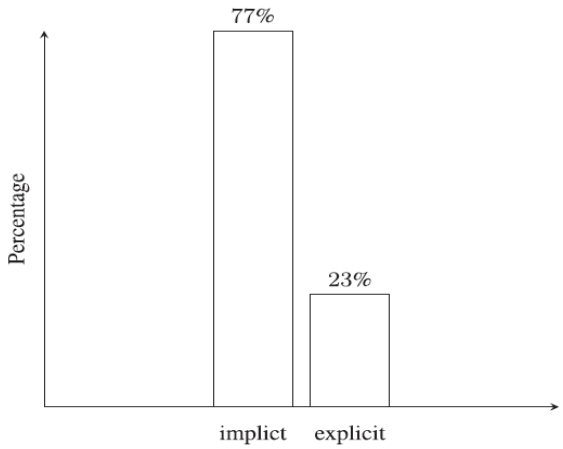

Fig. 6: Ratio between implicit aspect categories and explicitly mentioned ones.

Because both unsupervised and supervised method work best for well-defined aspect categories, the last category in this data set, anecdotes/miscellaneous

poses a challenge. It is unclear what exactly belongs in this category, and its concept is rather abstract. For that reason, we have chosen not to assign this category using any of the actual algorithms, but instead, this category is assigned when no other category is assigned by the algorithm. The characteristics in Fig. 4 also show that the use of anecdotes/miscellaneous as a "fallback" is justified given its large size and the fact that every sentence has at least one category.

A) Unsupervised Method

Table I displays, for each aspect category, the chosen firing threshold together with the resulting precision, recall, and F1-score on the test set. The category anecdotes/miscellaneous is estimated when none of the other four categories are chosen in the sentence.

From Table I, one can conclude that this approach has difficulty predicting the category ambience. This might be due to the nature of that particular category, as it is often not specified in a sentence by just one word, but is usually derived from a sentence by looking at the sentence as a whole.

Table I. Chosen Firing Thresholds and Their Evaluation Scores on the Test Set

Category	TP's	FP's	FN's	τ_c	precision	recall	F_1
food	313	103	105	0.22	75.1%	74.4%	74.8%
service	100	4	72	0.19	96.2%	58.1%	72.5%
ambience	41	10	77	0.09	80.4%	34.8%	48.5%
price	52	16	31	0.09	79.0%	54.2%	64.3%
misc.	163	159	71	-	50.6%	70.9%	59.1%
all	852	157	173	-	70.0%	64.7%	67.0%

B) Supervised Method

For the supervised method we use the training set to learn the parameters and co-occurrence frequencies, after which we evaluate the method on the test set. To see the impact the dependency indicators have, this method is executed separately for the dependency indicators, lemma indicators and a combined version where both lemma and dependency indicators are used, and evaluated on the test set. Tables II–III show the results.

Table II. Evaluation Scores of the Supervised Method with both Dependency and Lemma Indicators on the Test Set

Category	TP's	FP's	FN's	precision	recall	F_1
food	371	51	47	87.9%	88.8%	88.3%
service	159	32	13	83.2%	92.4%	87.6%
ambience	83	28	35	73.8%	70.3%	72.5%
price	74	8	9	90.2%	89.2%	89.7%
anecdotes/misc.	165	38	69	81.3%	70.5%	75.5%
all	852	157	173	84.4%	83.1%	83.8%

Table III. Evaluation Scores of the Supervised Method with Only Dependency Indicators on the Test Set

Category	TP's	FP's	FN's	precision	recall	F_1
food	343	45	75	88.4%	82.1%	85.1%
service	152	27	20	84.9%	88.4%	86.6%
ambience	62	34	56	64.6%	52.5%	57.9%
price	61	5	22	92.4%	73.5%	81.9%
anecdotes/misc.	165	38	69	81.3%	70.5%	75.5%
all	783	149	242	84.0%	76.4%	80.0%

IV. CONCLUSION

In this paper we have presented two methods for detecting aspect categories that is useful for online review summarization. The first, unsupervised, method, uses spreading activation over a graph built from word co-occurrence data, enabling the use of both direct and indirect relations between words. This results in every word having an activation value for each category that represents how likely it is to imply that category. While other approaches need labeled training data to operate, this method works unsupervised. The major drawback of this method is that a few parameters need to be set beforehand, and especially the category firing thresholds (i.e., τc) need to be carefully set to gain a good performance. We have given heuristics on how these parameters can be set.

The second, supervised, method uses a rather straightforward co-occurrence method where the co-occurrence frequency between annotated aspect categories and both lemmas and dependencies is used to calculate conditional probabilities. If the maximum conditional probability is higher than the associated, trained, threshold, the category is assigned to that sentence. Evaluating this approach on the official SemEval-2014 test set [10], shows a high F1-score of 83%.

REFERENCES

1. P. F. Bone, "Word-of-mouth effects on short-term and long-term product judgments," J. Bus. Res., vol. 32, no. 3, pp. 213–223, 1995.

2. R. Feldman, "Techniques and applications for sentiment analysis, " Commun. ACM, vol. 56, no. 4, pp. 82–89, 2013.

3. S. Sen and D. Lerman, "Why are you telling me this? An examination into negative consumer reviews on the Web," J. Interact. Marketing, vol. 21, no. 4, pp. 76–94, 2007.

4. B. Bickart and R. M. Shindler, "Internet forums as influential sources of consumer information," J. Consum. Res., vol. 15, no. 3, pp. 31–40, 2001.

5. D. Smith, S. Menon, and K. Sivakumar, "Online peer and editorial recommendations, trust, and choice in virtual markets," J. Interact. Marketing, vol. 19, no. 3, pp. 15–37, 2005.

6. M. Trusov, R. E. Bucklin, and K. Pauwels, "Effects of word-of-mouth versus traditional marketing: Findings from an Internet social networking site," J. Marketing, vol. 73, no. 5, pp. 90–102, 2009.

7.M. T. Adjei, S. M. Noble, and C. H. Noble, "The influence of C2C communications in online brand communities on customer purchase behavior," J.

Acad. Marketing Sci., vol. 38, no. 5, pp. 634–653, 2010.
8. B. Pang and L. Lee, "Opinion mining and sentiment analysis," Found. Trends Inf. Retrieval, vol. 2, nos. 1–2, pp. 1–135, 2008.
9. C.-L. Liu, W.-H. Hsaio, C.-H. Lee, G.-C. Lu, and E. Jou, "Movie rating and review summarization in mobile environment," IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 42, no. 3, pp. 397–407, May 2012.
10. M. Pontiki et al., "SemEval-2014 Task 4: Aspect based sentiment analysis,"in Proc. 8th Int. Workshop Semantic Eval. (SemEval), Dublin, Ireland, 2014, pp. 27–35.
11. S. Kiritchenko, X. Zhu, C. Cherry, and S. M. Mohammad, "NRCCananda-2014: Detecting aspects and sentiment in customer reviews," in Proc. 8th Int. Workshop Semantic Eval. (SemEval), Dublin, Ireland, 2014, pp. 437–442.
12. T. Brychcin, M. Konkol, and J. Steinberger, "UWB: Machine learning approach to aspect-based sentiment analysis," in Proc. 8th Int. Workshop Semantic Eval. (SemEval), Dublin, Ireland, 2014, pp. 817–822.
13. C. R. C. Brun, D. N. Popa, and C. Roux, "XRCE: Hybrid classification for aspect-based sentiment analysis," in Proc. 8th Int. Workshop Semantic Eval. (SemEval), Dublin, Ireland, 2014, pp. 838–842.

14. G. Castellucci, S. Filice, D. Croce, and R. Basili, "UNITOR: Aspect based sentiment analysis with structured learning," in Proc. 8th Int. Workshop Semantic Eval. (SemEval), Dublin, Ireland, 2014, pp. 761–767.
15. Z. Hai, K. Chang, and J.-J. Kim, "Implicit feature identification via co-occurrence association rule mining ," in Proc. 12th Int. Con. Comput. Linguist. Intell. Text Process. (CICLing), Tokyo, Japan, 2011, pp. 393–404.
16. K. Schouten and F. Frasincar, "Survey on aspect-level sentiment analysis," IEEE Trans. Knowl. Data Eng., vol. 28, no. 3, pp. 813–830, Mar. 2016.
17. Q. Su, K. Xiang, H. Wang, B. Sun, and S. Yu, "Using pointwise mutual information to identify implicit features in customer reviews," in Computer Processing of Oriental Languages. Beyond the Orient: The Research Challenges Ahead (LNCS 4285), Y. Matsumoto, R. Sproat, K.-F. Wong, and M. Zhang, Eds. Berlin, Germany: Springer, 2006, pp. 22–30.
18.Q. Su et al., "Hidden sentiment association in Chinese Web opinion mining," in Proc. 17th Conf. World Wide Web (WWW), Beijing, China, 2008, pp. 959–968.
19.X. Zheng, Z. Lin, X. Wang, K.-J. Lin, and M. Song, "Incorporating appraisal expression patterns into topic modeling for aspect and sentiment word identification," Knowl. Based Syst., vol. 61, no. 1, pp. 29–47, 2014.
20. W. Wang, H. Xu, and W. Wan, "Implicit feature identification via hybrid association rule mining," Expert Syst. Appl. Int. J., vol. 40, no. 9, pp. 3518–3531, 2013.
21.Y. Zhang and W. Zhu, "Extracting implicit features in Online customer reviews for opinion mining," in Proc. 22nd Int. Conf. World Wide Web Companion (WWW Companion), 2013, pp. 103–104.
22. G. Qiu, B. Liu, J. Bu, and C. Chen, "Opinion word expansion and target extraction through double propagation," Comput. Linguist., vol. 37, no. 1, pp. 9–27, 2011.
23. K. Schouten, F. Frasincar, and F. de Jong, "COMMIT-P1WP3: A co-occurrence based approach to aspect-level sentiment analysis," in Proc. 8th Int. Workshop Semantic Eval. (SemEval), Dublin, Ireland, 2014, pp. 203–207.

AUTHORS

M. Janardhan has received MCA degree and M.Tech degree from SV University, Tirupathi and JNTU Anantapur respectively. He dedicated to teaching from last 10 years. He guided 20 P.G and 20 U.G student batches. Currently Working as Assistant Professor in PBR VITS, Kavali, Andhra Pradesh, India.
M. Rufus Kumar has received B.Tech degree in CSE from PBR VITS affiliated to JNTU, Anantapur in 2015 and pursuing M.Tech degree in CSE from PBR VITS, Kavali affiliated to JNTU, Anantapur in 2019.