Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 171 (2020) 22–31

Procedia Computer Science

www.elsevier.com/locate/procedia

Third International Conference on Computing and Network Communications (CoCoNet'19)

An Unsupervised Hierarchical Rule Based Model for Aspect Term Extraction Augmented with Pruning Strategies

Manju Venugopalan[a], Deepa Gupta[a *]

Dept of Computer Science and Engineering, Amrita School of Engineering, Bengaluru, Amrita Vishwa Vidyapeetham,India

Abstract

Sentiment Analysis is an interesting field in text analytics which has emerged from an individual's curiosity to know what others think or feel about an entity or experience and hence take a decision. Aspect level sentiment analysis is a sub-field which operates at a more fine grained level to meet the demand of the end users who are no longer satisfied with the overall sentiment about the entity but are keen to know what features of the entity are matters of concern and what inturn is the sentiment reflected on each of these features/aspects. Aspect term extraction is the first and the most difficult step in aspect level sentiment analysis which attempts to discover the features of the product that are discussed about in the reviews. The proposed approach presents an unsupervised hierarchical rule based approach for aspect term extraction oriented towards a high recall and augmented by pruning strategies for filtering the false positives. The proposed model has reported an appreciable recall of 81.9 and 68.7 on Restaurant and Laptop domains respectively on SemEval 2014 dataset and has also been compared with the state of art models.

Keywords: Aspect term extraction; rule based; word embedding; syntax and semantics;SemEval2014

1. Introduction

The current era has witnessed an exponential growth of online data available in different forms. This in turn owes to the technological advancements that help us be connected round the clock which facilitates sharing our experiences and retrieving what others share in a short span of time. This enormous volume of online data is a valuable corpus for researchers from different fields. Natural Language Processing (NLP) is one such domain which has the goal of making machines understand text the way humans do by learning from large corpora. An interesting domain in NLP which is inspired by the instinct to know what others do or feel is Sentiment Analysis[1,2,3] which automates the task of

10.1016/j.procs.2020.04.303

extracting user opinion or sentiment from a piece of text. The availability of opinion-rich resources in the form of social networking sites, review websites, blogs etc. activates an immense amount of research activity in the field. The applications of Sentiment Analysis are quite varied ranging from a consumer taking input from others experiences on buying a product or visiting a tourist destination to influencing decision making in product design strategies of companies or even predictions for stock markets. The varied applications of the domain and text being the largest volume of data available has contributed to growing popularity of Sentiment Analysis. Most of the initial research works in the field were classification based approaches which tried to categorize the sentiment expressed in a sentence or a review text as Positive, Negative or Neutral[4]. Later the challenges evolved towards exploration of sentiment classification[5,6] on data from social networking sites. Models which experimented semantic models like building polarity lexicons[7,8] which reflect contextual polarity have been other active areas of research in the field.

With the advent of time, the research in sentiment analysis saw a shift towards a sub field called Aspect Level Sentiment Analysis[9] owing to the curiosity of knowing what aspects or features of the product or experience is talked about in the reviews. Aspect level Sentiment Analysis is a field in sentiment analysis which arose on demand. Consumers are no longer satisfied with an aggregate visualization of sentiment in the reviews, rather a more fine-grained representation of opinion in terms of the product features and their respective sentiments is in demand. The first sub-task in Aspect Level Sentiment Analysis is to identify the features or aspects of the product or the experiences that are discussed in the review which is referred to *aspect term extraction*. The next sub-task *sentiment classification* identifies the sentiment towards each aspect term extracted in the previous sub-task and hence labels the sentiment as Positive, Negative or Neutral. For example, in the review "*I enjoyed the new Windows 8 and touchscreen functions",* the aspect terms are *"Windows 8"* and *"touchscreen functions"* and the sentiment towards both the aspect terms are Positive.

Sufficiently large volumes of labeled data have been available for building supervised models for pure sentiment classification applications. But the availability of aspect level labelled data is very limited in terms of size and domains. Appreciable research works have been explored in this area most of them being supervised in nature. The availability of large volumes of training data labelled at the aspect level across different domains is always a challenge and hence demands the need for unsupervised models. The proposed approach puts forward an unsupervised hierarchical rule based approach for the sub-task of aspect term extraction augmented by pruning strategies.

The paper is structured as follows. Section 2 discusses a few prominent works in the area of aspect term extraction. Section 3 discusses the proposed methodology in detail, Section 4 points to the datasets and evaluation metrics used. The experimental results and analysis are presented in Section 5. The findings of the work are concluded and possible future directions are discussed in Section 6.

2. Related Works

There have been some sincere efforts to explore aspect terms extraction which have been mostly supervised in nature. The current survey analyses the prominent works in the field which are unsupervised/semi-supervised/hybrid in nature. One of the earliest approaches in the field of aspect term extraction has been by Hu and Liu[10] where the features/aspects were found by extracting all nouns and noun phrases and then pruning them based on the frequency of occurrence. They also attempted to extract infrequent features by finding opinion words corresponding to the extracted aspects and in turn finding if they are acting as modifiers in other sentences and hence finding unidentified aspects. The proposers reported a precision and recall of 0.8 and 0.72 respectively on the dataset they have created from reviews on electronic products like camera, DVD player etc. Zhuang et al. specially designed a strategy[11] for the movie domain where they pruned feature words/aspects using statistics from the labelled data. The proposers created a list of all aspects from train data and pruned the ones which had a frequency of occurrence less than one percentage of all the feature words taken together. Their work proved that the remaining words could identify more than 90% of the features in the movie domain. Double propagation method[12], which was a state of art towards the end of the previous decade was experimented for aspect extraction which exploits syntactic relations between opinion words and uses this two way relationship to explore both the categories iteratively. An attempt to improvise the basic double propagation strategy[13] has been attempted by Bing Liu and team where the learnings from dependency relationships between features and opinion words is in turn used to extract more aspect terms which are part of such dependencies. The experimentations were performed on *Cars* and *mattress* domains, the dataset being borrowed from a commercial company.

Semi-supervised approaches[14] in this field included using a seed set for topic modelling and hence extract the aspect terms. A set of seed aspects from each category in the respective domain would guide the clustering approaches to identify specific terms related to seed words in each cluster. The experiments were performed in hotel domain and the proposers claimed improvement over traditional topic modelling approaches. A Conditional Random Field based approach [15] has been attempted which visualizes aspect term extraction as a sequence labelling problem. The experiments have been carried out on hotel domain reviews which have been self-annotated at the sentence level. They reported an average F-measure of 0.6 across different aspect categories considered. The proposers of [16] claimed to be the first attempt of applying a deep learning approach to the task of aspect extraction. A deep CNN architecture with input sentences represented as word embedding vectors augmented by a set of rules based on linguistic patterns formed the architecture backbone. They have reported the results on both Hu and Liu 2004 and SemEval 2014 datasets and have improvements over the state of art in aspect extraction task.

Improvised versions[17] of double propagation methods where rule based methods augmented by initial seed list of opinion words and aspect terms have been applied on SemEval 2014 datasets and they have reported an F-measure of 0.61 in Restaurants domain and 0.36 in Laptops domain. The same team attempted the task in SemEval 2015 too where aspect term extraction was open only in the Restaurant domain. They used a graph based model which modelled opinion words and aspect terms as nodes of the graph and the edges representing dependency relations. A random walk algorithm is used to rank the nodes and hence find the most reliable aspects. The model[18] reported an F-measure of 0.45 on the Restaurant domain. A hybrid approach[19] to aspect term extraction has been explored based on distributed representations of words and dependency paths learned in an unsupervised manner. The basic idea is to connect two words with the dependency path between them in the embedding space. These learned features are then used to design a CRF based model for aspect term extraction. The proposers have reported F-measure values of 0.84 and 0.75 on Restaurant and Laptop domains respectively. A methodology[20] which claims to be completely unsupervised has also been reported which performs in two phases. The first phase uses a high precision rule-based approach supported by lexicon resources. This phase creates annotated datasets automatically which is used as the train data for a robust supervised model based on a bidirectional long short-term memory (BLSTM) network. They have reported F-measures of 0.42 and 0.54 on Laptops and Restaurant domains on SemEval 2014 datasets. A hybrid strategy for aspect term extraction[21] has also been explored which initially uses a rule-based unsupervised strategy at the chunk level followed by pruning methods to generate pseudo labels for aspect terms. These pseudo labels are in turn used to train a deep GRU network for aspect term extraction designed as a sequence labelling task. They have reported an F-measure of 0.76 and 0.61 on Restaurant and Laptop domains respectively of SemEval 2014 datasets.

Most of the works in the field of aspect term extraction have been supervised in nature. The latter half of the current decade has witnessed the exploration of a few unsupervised and semi-supervised models too. The survey has also presented a few state-of-art models which have reported results on SemEval 2014, a standard dataset for aspect term extraction. The results reported by pure unsupervised models and the understanding that only minimal domains are explored is an eye-opener on that the area demands further exploration.

Aspect term extraction aims at identifying the aspects in the review sentence which could be single or multiword terms. To a certain extent they are found as nouns or part of noun phrases in a generic context. But the exceptions that deviate from such typical patterns are equally voluminous augmented by the errors of POS tagging systems. Therefore, it is almost practically impossible to create a robust model for aspect term extraction from POS tags alone. Hence the aspect extraction model needs to be strengthened by deriving inputs from various grammatical dependencies which exist among the words in the sentence. The proposed approach attempts to explore the strength of grammatical syntactic patterns, POS(Part of Speech) tags and NER(name Entity Recognition) tags together supported by heuristic rules to form a pure unsupervised aspect extraction model with high recall and hence use hierarchical pruning strategies to filter the false positives. The proposed hierarchical three layer model tries to extract single and multi-word aspects using grammatical dependencies in Layer1, extract more single and multi-word aspects using heuristic rules in Layer 2 and finally prune the false positives in Layer 3.

3. Proposed Methodology

The proposed model is a hierarchical rule based model which includes grammatical dependency based rules and heuristic rules augmented with pruning strategies. The proposed methodology for aspect term extraction is diagrammatically depicted in Fig.1. The model comprises of a pre-processing module followed by a rule based aspect term extraction module to extract single and multi-word prospective aspect term candidates. This rule based aspect term extraction module relies on the strength of POS tags, grammatical dependencies, NER tags and heuristic rules to extract aspect terms ensuring a high recall. This is followed by a pruning module which incorporates multiple strategies to filter the false positives.

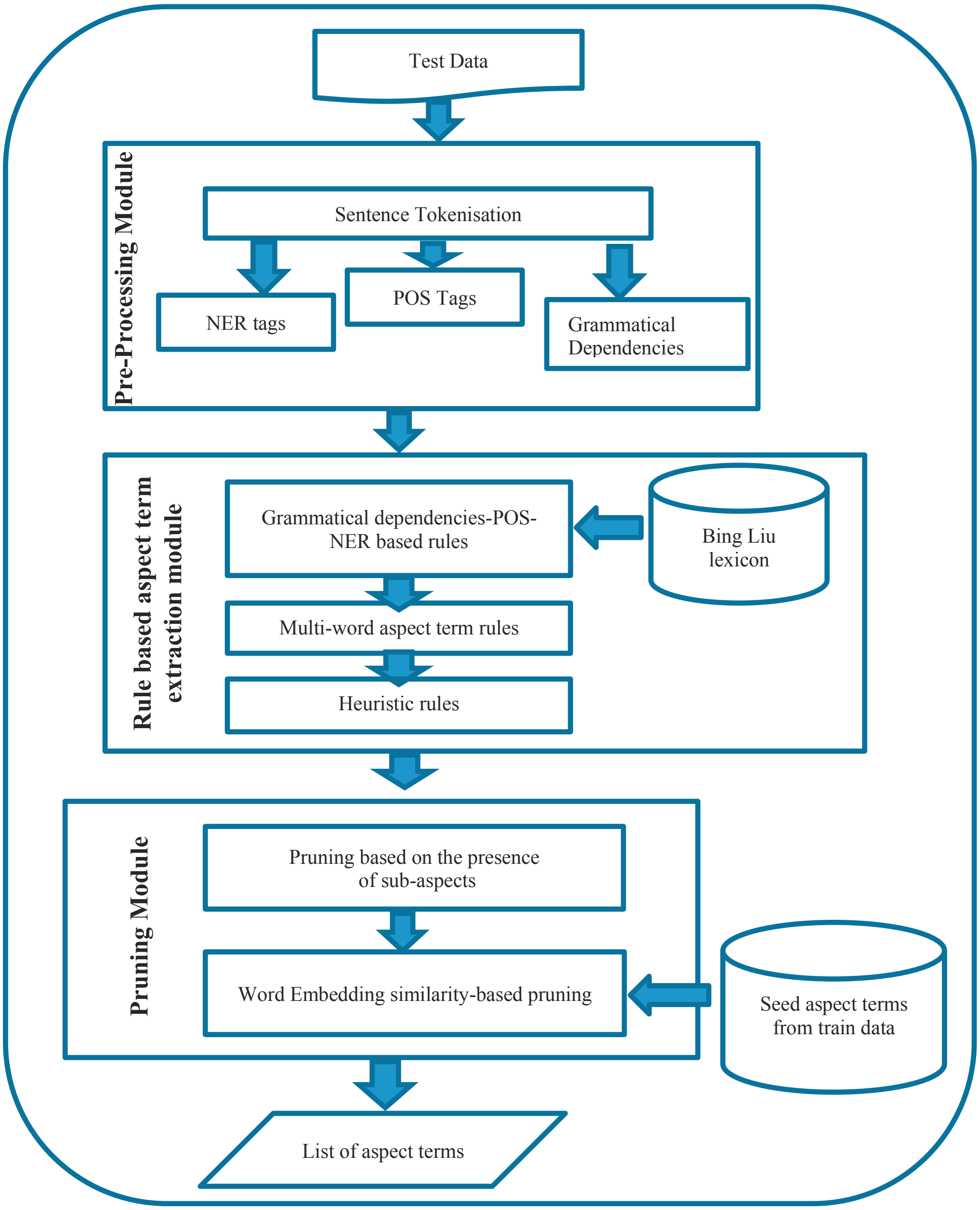
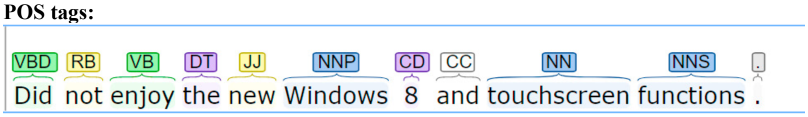


Fig.1 Flow diagram of the Hierarchical Rule based model for Aspect Term Extraction Augmented with Pruning Strategies

Table 1. Preprocessed outputs of a sample review sentence

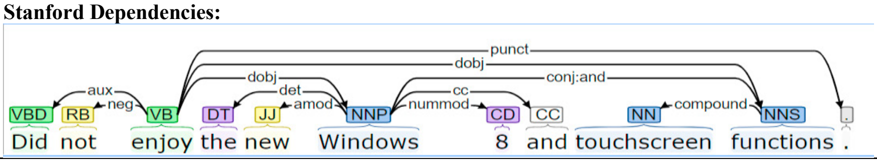
Input sentence:
Did not enjoy the new Windows 8 and touchscreen functions.

POS tags:



NER tags: Windows 8 : PRODUCT

Stanford Dependencies:



3.1 Pre-Processing module

Every review sentence in the dataset is subjected to a POS tagging process to derive the POS tag of every word in the sentence. All the grammatical dependencies between the words in the sentences are extracted using the Stanford Dependency Parser using Stanford CORENLP 3.9.2[*]. The NER tags of each word in the review is also extracted using spacy [†]package in Python 3.5. The outputs of the pre-processing stage are shown in Table.1. The parse tree of the sentence is showcased for visualization. For the sample input review sentence given in Table.1, the POS tags along with the index of the word is pre-processed output that would be utilized by the model. The NER tags often give an indication of the aspect terms or at least serve in cross verifying that the word is not an aspect. Complete reliability on NER tags is constrained by model accuracies. The grammatical dependencies for the sample sentence shown in Table.1 reveals that the word *Windows* is in a *dobj* relationship with the word *enjoy* which means that Windows is a direct object of the *verb* "*enjoy"* which actually helps us to identify *Windows* as an aspect. The raw input sentences, their POS tags, grammatical dependecies and NER tags are stored in datastructures suitable for retreival which forms the input for the next stage.

3.2 Hierarchical Rule Based Aspect Term Extraction Module (HRB-ATA)

This module (*HRB-ATA*) is designed using a set of eleven rules out of which the first eight are based on grammatical dependencies, POS and NER tags and the remaining are heuristic based. Each review sentence after the pre-processing stage is subjected to a set of rules depicted in Table.2 in the form of an algorithm. The first rule R1 in Level1 attempts to discover aspects of the form *USB stick, technical support, built-in camera* etc. which are two word aspects where the initial word is tagged as *adjective* by the POS taggers but the rules ensure them to be aspect terms by cross verifying that they are not opinion words which reflect sentiment. All affirmations of being an opinion word or not in the proposed methodology has been cross verified using the presence in Bing Liu lexicon[‡]. The second rule R2 tries to capture occurrences of two consequent nouns to form an aspect term. The third rule R3 tries to capture a single word noun aspect if it is preceded by an adjective which is an opinion word. The first three rules utilize only the POS tags and lexicons as input resources. All the consequent rules in Level1 try to capture single word aspects which are tagged as nouns and are identified based on its grammatical dependencies. For example, the rule numbered R6 tries to identify noun tokens which are in a *amod* relationship with an adverb or adjective. Fig.2 depicts the dependency relations in a

[**] https://stanfordnlp.github.io/CoreNLP/download.html
[†] https://spacy.io/usage
[‡] https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html

Table 2. Algorithm for Rule Based Aspect Extraction

***Inputs**: Pre-processed data for the respective domain S*
Liu : opinion word dictionary of Bing Liu
***Output** :aspect[][] a two dimensional array where each row aspects[i]corresponds to the aspects in each sentence of S*

LEVEL 1:
for i = all sentences in *S*
for every word in *S[i]*

R1: if word is a noun and preceded by an adjective which is not an opinion word then concatenate the word and its preceding word and append it to aspect[i]

R2: if word is a noun and preceded by another noun then concatenate the word and its preceding word and append it to aspect[i]

R3: if word is a noun and preceded by an adjective which is an opinion word then append the word to aspect[i]

R4: if word is a noun and is in a dobj relationship with a verb in the sentence then append the word in aspect[i]

R5: if word is a noun and is in a nsubj relationship with an adjective in the sentence then append the word in aspect[i]

R6: if the sentence S[i] contains a SUBJECT VERB and if the word has any adverbial or adjective modifier which is an opinion word append the word to aspect[i]

R7: if word is a noun and is in a modifier relationship with a copula verb in the sentence then append the word to aspect[i]

R8: if word is a noun and if the previous word is not a preposition of place or a word corresponding to an NER tag of {'TIME', 'ORDINAL', 'NUMBER', 'DATE', 'PERCENT'} and there is atleast one opinion word in the sentence , then append the word to aspect[i]

LEVEL 2 :
for i = length of aspect[][]
for j= length of aspect[i]
if aspects[i][j] is a single aspect term

R9: Find other nouns in the sentence which has direct dependency with aspect[i][j] and append them to aspect[i]

R10: if they are in a compound relationship with other nouns, append them to aspect[i]

LEVEL 3:
for each single word aspects in aspects[i]

R11: If any two retrieved aspects are in proximal indices, they are combined to produce multi-word aspects

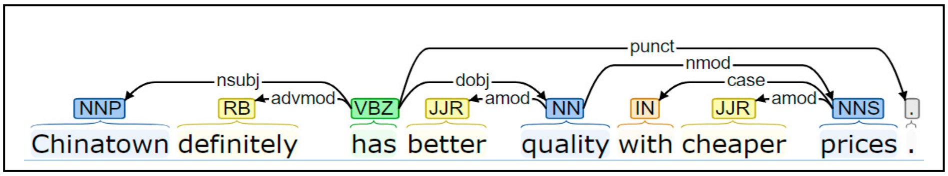


Fig.2 Example of dependency grammar and POS tag used to extract an aspect term based on rule R4

sample review sentence where by the application of R4, the term *quality* which is a *dobj* relationship with a verb is identified as an aspect term. Similarly, the term *prices* which is modified by the adjective *cheaper* is identified as an aspect term using rule R6. The final rule in Level1 i.e. R8 is more on the generic side trying to capture all single word nouns with multiple checks based on the presence of atleast one opinion word in the sentence, filtered using unlikely NER tags. The rules in Level 2, R9 and R10 attempts to discover the aspect terms unidentified by the rules in Level1 through their dependencies if any with the extracted prospective aspect terms from Level1. The rules in Level1 and Level2 have attempted to retrieve only single and two word aspects, hence Level3 attempts to combine noun phrases extracted in previous steps based on the proximal indices. The aspects extracted in the sentence are analyzed and if they occur in continuous indices they are merged to form a single aspect term. This step was designed mainly considering the Restaurant domain which has a larger count of multi-words aspects of length more than two which were mostly names of dishes like *mushroom barley soup*, *chili signed food items*, *ice blended mocha* and so on. This strategy contributes to improving both precision and recall. Either it ends up in extracting multi-words aspect terms or in the worst case combines the false positives and reduces their count.

3.3 Pruning Module

This module of the proposed approach decribes the different strategies adopted for pruning false positive aspect terms and are explained in the sebsequent subsections.

Pruning based on the presence of sub-aspects (PS1)

The rule-based extraction module would have retrieved words as single word aspects as well as part of multiword aspects. This pruning strategy analyses every single word aspect retrieved from the sentence and if it's already a part of a multi-word aspect retrieved then the former is removed from the list, provided it is not an aspect term derived from a reliable rule in Level1. An analysis on the reliability of the rules in Level1 and Level2 has been performed on the basis of ration of true positives and false positives produced by them on the train data. Based on this analysis rules R3, R5 and R6 have been rated the most reliable rules. For e.g. in the review "*We loved the pink pony"*, if both the aspect terms *pony* and *pink pony* have been retrieved as aspect terms, the aspect term *pony* would be removed from the list. This strategy is expected to have a high impact because in order to ensure that the proposed model has a high recall a few rules for single aspect term extraction are sketched in a generic manner. This strategy contributes to improving both precision and recall because it ends up in extracting multi-words aspect terms or in the worst case combines the false positives and reduces their count.

Word Embedding Similarity based Pruning (PS2)

The semantic similarity of the prospective aspect terms with the domain under consideration is used as a measure to prune them. Word Embedding is one of the most striking vector representations for measuring syntactic and semantic similarities by capturing both continuous distributed representations of the context. Word2Vec[22], a popular word embedding model based on a feed forward neural network has been adapted to implement this feature. In order to ensure that context information has been captured to the maximum, the proposed approach uses vector representations that have been trained from a large corpus of reviews from the respective domain rather than using pre-trained vectors. The challenge lies in the fact that resource chosen for training should be sufficiently large enough. For the restaurant domains reviews the input has been from Yelp restaurant reviews[§] which is around 1.25GB of data consisting of around 52 lakh sentences. For the Laptop domain, Amazon product reviews[**] from electronic appliances domain which consists of more than 16 lakh reviews have been used for training. The net result is that the Word2Vec model arrives at a domain specific dense vector representation for all the words in the training corpora of both the domains. (Yelp and Amazon data respectively). An algorithm to filter prospective aspect terms based on word embedding similarity is depicted in Table 3. In order to create a representative sample of each domain, the most frequent aspect terms in the respective SemEval train dataset is extracted to form a seed set which consists of 16 target terms. The similarity of the prospective candidate aspect terms with the domain is measured in terms of their cosine

[§] https://www.yelp.com/dataset

[**] http://jmcauley.ucsd.edu/data/amazon/

similarity with these seed words. The similarity of each prospective aspect term with the seed set is calculated using the Word2Vec vector representations and the maximum value among those is assigned as depicted in Step3 of the algorithm. For multiword aspect terms the word2vec similarity is calculated as the cosine similarity between the seed term and an averaged vector representation of all the words in the aspect term. A threshold is arrived based on the statistics of the word-similarity values and all the prospective aspect terms whose similarity with the seed set is less than the threshold value is pruned. Here the threshold has been assigned a value 0.1. The final aspect terms in the generated list form the output of the proposed model.

Table 3. Algorithm for Word Embedding Similarity based pruning

Inputs: *seed set P which is the set of high frequent aspects terms extracted from train data*
aspects[][] : a two dimensional array where each row contains the aspects in each sentence of S
Outputs: *filtered_asp[], the set of filtered aspects which pass the domain similarity test*

```
Step1:  Create a 1D list asp[] from the unique words of aspects[][]
Step2:  filtered_asp=[]
Step3:  for i in len(asp)
            max=0
            for j  in len(P):
                if asp[i] is a multi-word aspect:
                    avg=averaged word vector for asp[i]
                    sim=word2vec_similarity(P[j],avg)
                else:
                    sim=word2vec_similarity(P[j],asp[i])
                if sim>max:
                    max=sim
            if max>0.1 :
                filtered_asp.append(asp[i])
```

4. Datasets and Evaluation Metrics

The proposed approach has been experimented and evaluated on SemEval-2014 datasets. SemEval is a series of semantic evaluation workshops organized since 2010 and one of the tasks in 2014 was Aspect Based Sentiment Analysis. The dataset released by SemEval-2014 for aspect term extraction task contains data from Restaurant and Laptop domains. The dataset contains reviews annotated at the sentence level with the aspect terms contained in each sentence. The train data in both domains contain more than 3000 sentences and 800 sentences as test data. The model performance on the test data has been evaluated using the standard metrics Precision, Recall and F-measure.

5. Experimental Results and Analysis

The proposed model has been implemented, evaluated and the results of evaluation have been reported in Table.4. As observed the results of the hierarchical rule based aspect term extraction model has been better on restaurant domain than laptop domain . This owes to the fact that the laptop domain is more challenging with aspect terms which are very technical and specific to operating systems and electronic appliances. The first pruning strategy based on pruning sub-aspects achieves a considerable achievement, an improvement of 11-12 points in precision across both the domains without compromising significantly on the recall. The second pruning strategy based on word embedding based similarities achieves an improvement of more than 3 points in both the domains with a slight dip in the recall. A further analysis to understand the model strength in terms of recall percentage with respect to single and multi-word aspect terms is presented in Table 5. As observed in the restaurant domain, more than 90% of single aspect terms have been retrieved whereas only 71% of the two-aspect terms and only 25% of multi word aspect terms with more than two terms have been retrieved. Whereas in the laptop domain 78% of single aspect terms, 67% of the two-word aspect terms and 26% of multi-word aspect terms have been retrieved. The model has performed well on retrieving single word aspect terms even though the results in laptop domain are not in par with the Restaurant domain. A couple of reasonings could be that the Laptop domain is not that expressive as Restaurant and hence the reachability of aspect terms through adverbial and adjectival modifiers would have been significantly less. The quality of POS tagging in the Laptop domain takes a hit owing to the presence of technical terms related to gadgets and operating systems. The major drawback is in dealing with multi-word aspects of length more than two as observed from the results across both the domains. The restaurant domain has a lot dish names as multi-word aspect terms in French or Italian which posed challenges to POS tagging and hence the aspects extracted. These constraints the performance of the rule based extraction model.

The model has been compared with baseline models representing the state of art and results presented in Table.6. Aitor et.al 2014 [17] which is the first considered baseline is based on a double propagation algorithm, which starts from a seed set and then propagates in both directions iteratively creating aspect terms and opinion words. Athanasios, et al. 2017 [20]is the second baseline which experiments a hybrid approach where the first phase uses a high precision rule-based approach supported by lexicon resources. This phase creates annotated datasets automatically which is used as the train data for a robust supervised model based on a bidirectional long short-term memory (BLSTM) network. Chuhan et al. 2018[21] explores a hybrid strategy for aspect term extraction which uses a rule-based unsupervised strategy at the chunk level followed by pruning methods to generate pseudo labels for aspect terms. These pseudo labels are in turn used to train a deep GRU network for aspect term extraction designed as a sequence labelling task. This hybrid model is considered as the third baseline and has showcased the best performance in both the domains. The proposed approach inspite of being a pure unsupervised model is the second performer in Laptop domain and Restaurants domain too. The proposed hierarchical rule based aspect extraction model has performed well considering the fact that it is purely unsupervised in nature. This model can serve as a strong base upon which further hybrid models can be constructed.

Table 4. Results across domains for Aspect Term Extraction on SemEval 2014 dataset

	Restaurant		Laptop	
Model	Precision	Recall	Precision	Recall
HRB-ATA	39.6	83.1	29.4	71.8
+PS1	50.1	82.2	41.9	70.9
+PS2	**53.0**	**81.9**	**45.5**	**68.7**

Table 5. Analysis of proposed model performance in terms of the recall of single and multi-word aspect terms

Domain	Recall % of single word aspect terms	Recall % of two-word aspect terms	Recall % of multi-word(>2) aspect terms
Restaurant	91.6	71.3	25.0
Laptop	78.1	67.4	26.6

Table 6. Comparison of proposed approach with baselines

Domain	Model Type	Restaurant	Laptop
Metric		F-measure	F-measure
Aitor et.al 2014 [17]	Unsupervised	0.61	0.36
Athanasios, et al. 2017[20]	Hybrid	0.54	0.42
Chuhan et al. 2018[21]	Hybrid	0.76	0.61
Proposed Approach	Unsuprvised	0.64	0.55

6 Conclusion

The work proposes a hierarchical rule-based model for aspect term extraction, augmented by pruning strategies used to filter the false positives. The model performance has been compared with three unsupervised/hybrid baselines reported on the same dataset. The model has reported an F-measure of 0.64 and 0.55 in the Restaurant and Laptop domain respectively of SemEval 2014 dataset. Even though Baseline 2 has showcased the best performance in both the domains, the proposed approach is the second performer in Laptop domain and Restaurants domain too. The proposed work has contributed an unsupervised model for aspect term extraction which is in par with the state of art. The model could be used for generating the seed base on which further models could be explored. Topic modelling is an unsupervised approach which helps to discover the latent topics in a document. Incorporation of topic modelling approaches to identify aspect terms could be explored. Existing model improvisation focused primarily towards multi-term aspect retrieval is another direction in demand.

References

[1] Liu, Bing. (2012) "Sentiment analysis and opinion mining." *Synthesis lectures on human language technologies* **5**(**1**): 1-167.

[2] Moraes, Rodrigo, JoaO Francisco Valiati, and Wilson P. GaviãO Neto. (2013) "Document-level sentiment classification: An empirical comparison between SVM and ANN." *Expert Systems with Applications* **40**(**2**): 621-633.

[3] Venugopalan, Manju, and Deepa Gupta. (2015) "Sentiment classification for Hindi tweets in a constrained environment augmented using tweet specific features." *International Conference on Mining Intelligence and Knowledge Exploration.* Springer, Cham :664-670

[4] Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. (2002) "Thumbs up?: sentiment classification using machine learning techniques." *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10.* ACL :79-86

[5] Venugopalan, Manju, and Deepa Gupta. (2015) "Exploring sentiment analysis on twitter data." *Eighth International Conference on Contemporary Computing (IC3).* IEEE: 241-247

[6] Pak Alexander, and Patrick Paroubek. (2010) "Twitter as a corpus for sentiment analysis and opinion mining." *LREc*. Vol. 10 :1320-1326

[7] Sanagar, Swati, and Deepa Gupta. (2016) "Roadmap for Polarity Lexicon Learning and Resources: A Survey." *The International Symposium on Intelligent Systems Technologies and Applications.* Springer, Cham:647-663

[8] Sanagar, Swati, and Deepa Gupta. (2015) "Adaptation of multi-domain corpus learned seeds and polarity lexicon for sentiment analysis." 2015 *International Conference on Computing and Network Communications (CoCoNet).* IEEE:50-58

[9] Pontiki, Maria, et al. (2016) "Semeval-2016 task 5: Aspect based sentiment analysis." *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)*:19-30

[10] Minqing Hu and Bing Liu, (2004) “Mining and summarizing customer reviews.” *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*: 168-177

[11] Zhuang, Li, Feng Jing, and Xiao-Yan Zhu. (2006) "Movie review mining and summarization." *Proceedings of the 15th ACM international conference on Information and knowledge management*: 43-50

[12] Qiu, Guang., Bing, Liu., Jiajun Bu and Chun Chen. (2009) “Expanding Domain Sentiment Lexicon through Double Propagation.” *Proceedings of IJCAI:* 1199-1204

[13] Zhang, Lei, et al. (2010) "Extracting and ranking product features in opinion documents.” *Proceedings of the 23rd international conference on computational linguistics: Posters.* Association for Computational Linguistics*:*1462-1470.

[14] Mukherjee, Arjun, and Bing Liu. (2012). "Aspect extraction through semi-supervised modeling.” *Proceedings of the 50th annual meeting of the association for computational linguistics: Long papers-volume 1.* Association for Computational Linguistics :339-348

[15] D. Marcheggiani, O. Tackstrom, A. Esuli, and F. Sebastiani, (2014) “Hierarchical Multi-label Conditional Random Fields for Aspect-Oriented Opinion Mining,” *Proceedings of the 36th European Conference on Information Retrieval (ECIR 2014).* Springer : 273–285.

[16] Poria, Soujanya, Erik Cambria, and Alexander Gelbukh. (2016) "Aspect extraction for opinion mining with a deep convolutional neural network." *Knowledge-Based Systems 108* : 42-49.

[17] Pablos, Aitor García, Montse Cuadros, and German Rigau. (2014) "V3: Unsupervised generation of domain aspect terms for aspect based sentiment analysis." *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)* :833-837

[18] Pablos, Aitor García, Montse Cuadros, and German Rigau. (2015) "V3: Unsupervised aspect based sentiment analysis for semeval2015 task 12." *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*: 714-718

[19] Yin, Yichun, et al. (2016) "Unsupervised word and dependency path embeddings for aspect term extraction. *Proceedings of the Twenty fifth International Joint Conference on Artificial Intelligence:* 2979-2985

[20] Giannakopoulos, Athanasios, et al. (2017) "Unsupervised aspect term extraction with b-lstm & crf using automatically labelled datasets." *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* :180-188

[21] Wu, Chuhan, et al.(2018)”A hybrid unsupervised method for aspect term and opinion target extraction" *Knowledge-Based Systems* 148: 66-73.

[22] Mikolov, Tomas, et al. (2013) "Efficient estimation of word representations in vector space." *Proceedings of the International Conference on Learning Representation.*