2016 IEEE International Conference on Computer and Information Technology

Supervised Learning Based Approach to Aspect Based Sentiment Analysis

Nipuna Upeka Pannala[1]
Department of Information Technology
Sri Lanka Institute of Information Technology,
New Kandy Road, Malabe, Sri Lanka
[1]nipuna.pannala@gmail.com

Chamira Priyamanthi Nawarathna[2]
Department of Information Technology
Sri Lanka Institute of Information Technology,
New Kandy Road, Malabe, Sri Lanka
[2]priyamanthi11@gmail.com

J.T.K.Jayakody[3]
Department of Information Technology
Sri Lanka Institute of Information Technology,
New Kandy Road, Malabe, Sri Lanka
[3]kosala.private@gmail.com

Lakmal Rupasinghe[4]
Department of Computer Science and Networking
Sri Lanka Institute of Information Technology,
New Kandy Road, Malabe, Sri Lanka
[4]lakmalr@gmail.com

Kesavan Krishnadeva[5]
Department of Information Systems Engineering
Sri Lanka Institute of Information Technology,
New Kandy Road, Malabe, Sri Lanka
[5]deva130989@gmail.com

***Abstract* - Aspect base sentiment analysis is a very popular concept in the machine learning era which is under the research domain still at the movement. This research mainly consist of the way of exploring the sentiment analysis based on the trained data set to provide the positive, negative and neutral reviews for different products in the marketing world. Most of the existing approaches for opinion mining are based on word level analysis of texts and are able to detect only explicitly expressed opinions. In aspect-based sentiment analysis (ABSA) the aim is to identify the aspects of entities and the sentiment expressed for each aspect. The ultimate goal is to be able to generate summaries listing all the aspects and their overall polarity.**

For this research mainly natural language and machine learning techniques are used. To train the application for the given data sets SVM (support vector machine) and ME (Maximum Entropy) classification algorithms have been used. Differentiation of the performance of the each algorithm will be analyzed through this research using the proven technologies available in the world like "Re call", "F-Measure" and Accuracy.

***Index Terms* - Aspect-based sentiment analysis (ABSA), SVM (support vector machine), ME (Maximum Entropy)**

I. INTRODUCTION

This research goes in to more detail of developing a product recommendation system based on the review analysis. Mainly this proposed system will be developed by using the machine learning algorithms with the more briefings of Clustering, Regression and classification algorithms. As a research objective, some classification algorithms will be reviewed against one another to identify the optimal use of it.

This total concept is known as Aspect Based Sentiment Analysis which is more popular in the domain of the machine learning technology. Aspect Based Sentiment Analysis (ABSA) systems receive as input a set of texts (e.g., product reviews or messages from social media) discussing a particular entity (e.g., a new model of a mobile phone). The systems attempt to detect the main (e.g., the most frequently discussed) aspects (features) of the entity (e.g., 'battery', 'screen') and to estimate the average sentiment of the texts per aspect (e.g., how positive or negative the opinions are on average for each aspect).

In aspect-based sentiment analysis the aim is to identify the aspects of entities and the sentiment expressed for each aspect. The ultimate goal is to be able to generate summaries listing all the aspects and their overall polarity [11]. For many application scenarios document level review classification is too coarse-grained and does not provide the desired information. Most of the existing approaches are based on word-level analysis of texts and are able to detect only explicit expressions of sentiment. Usually classifying opinion texts at the document level or the sentence level is often insufficient for applications because they do not identify opinion targets or assign sentiments to such targets (aspects).

A. Machine Learning Approach

This approach use well known machine learning techniques such as terms presence, frequency, parts of speech (POS), opinion words and negations. When going into more detail of the terms, following descriptions can be provided.

- **Terms presence and frequency**: These features are individual words or word n-grams and their frequency counts. It either gives the words binary weighting (zero if the word appears or one if otherwise) or uses term frequency weights to indicate the relative importance of features [5].
- **Parts of speech (POS):** finding adjectives, as they are important indicators of opinions.
- **Opinion words and phrases:** these are words commonly used to express opinions including good or bad, like or hate. On the other hand, some phrases express opinions without using opinion words. For example: cost me an arm and a leg.

978-1-5090-4314-9/16 $25.00 © 2016 IEEE
DOI 10.1109/CIT.2016.107

- **Negations**: the appearance of negative words may change the opinion orientation like not good is equivalent to bad.

Text Classification Problem Definition: which have a set of training records,

D = {X1;X2; :::;Xng} where each record is labeled to a class. The classification model is related to the features in the underlying record to one of the class labels. Then for a given instance of unknown class, the model is used to predict a class label for it.

B. Supervised Learning

Supervised learning is the machine learning task of inferring a purpose from labeled training data [6]. The training data involve of a set of training examples. Unsupervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal). A supervised learning algorithm studies the training data and produces an inferred function, which can be used for mapping new examples. An optimal scenario will allow for the algorithm to correctly determine the class labels for unseen instances. This requires the learning algorithm to generalize from the training data to unseen situations in a "reasonable" way.

C. Supervised Learning Mechanisms

- Support Vector Machine Classifier

This method can be used in supervised learning models with associated learning algorithms that analyze data and identify the patterns which are used for classification and regression analysis. The main principle of SVMs is to determine linear separators in the search space which can best separate the altered classes. That means for the given set of data, SVM tries to create a most suitable hyper lane which will separate between two classes with largest possible gap.

II. METHODOLOGY

A. System Architecture

This chapter outlines the overall design of the proposed solution. It illustrates the general structure for the approach which has taken to tackle each of the sub-tasks of Aspect Based Sentiment Analysis (ABSA) system. This proposed system mainly develops using the machine learning techniques, which incorporated with natural language processing. The main aim of this research is to investigate and develop a methodology to facilitate supervised learning based approach to aspect based sentiment analysis. This research is conducted exploring several related techniques used in opinion mining and sentiment analysis.

Simultaneously for this research, some subsets of features are selected and chose the algorithm with the highest precision, recall, the highest F-Measure and with a high degree of consistency. Based on the constraints identified in online customer reviews data and classification models, several mitigation mechanisms are introduced to overcome these limitations (i.e. domain adaption).

The proposed architecture is detailed in the below Figure. This proposed system contains 3 main modules.

- Preprocessing module
- Domain Specific Feature Generation module
- Sentiment Analyze module with aid of a General Purpose Sentiment Dictionary created based on publicly available data sets of reviews for a particular domain.

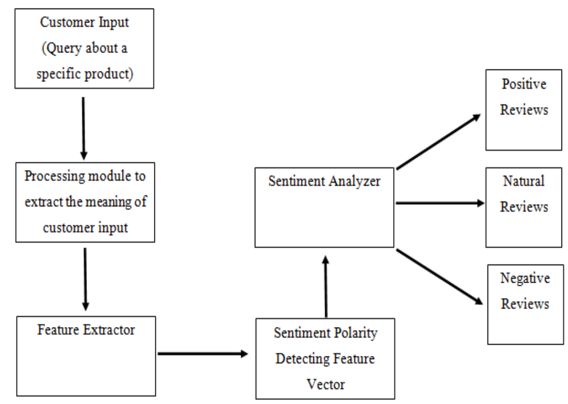

Fig 1. System Architecture

This processing module mainly concentrates about how to setup the sentence in order to get the extract meaning of the user entered inputs. Most of the time natural language sentences used as inputs for sentiment analysis, probable to contain malformed and incomplete text because of the less restricted, free format nature users were given at the time of generating the content. In order to prepare the data for further process, each sentence has to be normalized using some simple filters. Word boundaries: white space and punctuations are handled such that removing them will not affect the features of sentences. Short form words and playful terms like ‘Im’, ‘gotta’ are converted into suitable formal English words as in ‘I am’, ‘got to’. Emoticons contain rich clues about the sentiment of sentences. But those are removed from sentences since the scope of this research is to focus on aspect based opinion mining.

B. Feature Extractor

Decomposition of natural language sentences into feature vectors is a focal task of any supervised learning based sentiment analysis system. Feature Extraction module consists of two main modules namely General Purpose Sentiment Dictionary and Domain Specific Feature generation module. That means feature extraction module is responsible for extracting a set of feature vectors from sentence given as the input. This feature extractor can be divided into three main sub categories. Those are,

1. Lexical Feature Extraction
2. Domain Specific Feature Extraction
3. Sentiment Analyzer

Feature extraction module extract the verbal, dictionary and domain specific features associated with online customer reviews (Data available in data set) with the aid of domain specific feature module and general purpose sentiment dictionary. Two feature sets (including corresponding class labels) are generated as the resultant outcome of this module.

Domain specific feature generation module facilitates the proposed system to get and identify the exact meaning of the terms relevant to a specific product category. As an example, for the category of laptops, some terms like useful, fast, excellent and high performance should be identified as the positive comments. Domain Specific Feature Generation module aids the feature extractor module to extract relevant domain specific features. The figure depicts the architecture of the proposed domain specific feature generation module. As shown in figure, proposed Domain Specific Feature Generation module starts with un-labeled corpora and finally produces a domain specific sentiment and word clusters.

C. Sentiment Polarity Detecting Feature Extractor

The objective of the sentiment polarity detection is to detect the sentiment expressed towards a given aspect category in a given sentence. Given a set of pre identified aspect categories for a sentence, this phase aims at determining the polarity (positive, negative, or neutral) of each aspect category. Instead of multiple classifiers, such as implemented for the aspect category detection phase, but this research used to train a single classifier, to choose between positive, negative or neutral polarity.

D. Sentiment Analyzer

The goal of the Sentiment Analyzer Module is to detect the sentiment expressed towards a given aspect category in a given sentence. The output of this sentiment analyser is a single sentiment label, (Either Positive, Negative, Neutral or Conflict of any) against a single input pair (Sentence and Aspect Category). Each sentence is transformed to a feature vector containing features generated from feature extractor and then classified as positive, negative or neutral [11]. For the classification purposes some of the machines learning algorithms are used in this scenario since it allows differentiating the results accordingly.

E. Data and Methods

Mainly this research was done based on the common tools and languages which is available in the world at the moment. As an example, for the development purposes, mainly python and NLTK (Natural Language Training Kit) is used on top of the Linux platform. This is a known and popular language available for the natural language processing perspectives.

F. Data Used

Being a data oriented research, the accuracy, impartiality and volume of the dataset on which the research is based, contributes greatly towards the success of the overall outcome. This study was used 3 review datasets that were used in the 2015 SemEval Challenge.

G. Restaurant Review Dataset

Restaurant review dataset was taken front the research which was conducted for the 2015 semEval Challenge. The dataset contains 1,315 training and 663 test sentences. The dataset includes annotations of coarse aspect categories of restaurant reviews. Specifically the annotations were performed by experienced human annotators from the SemEval team. Format of the restaurant review dataset is as below [4].

```
<sentence id="1004293:0">
<text>Judging from previous posts this used to be a
good place, but   not any longer.</text>
<Opinions>
<Opinion target="place"
category="RESTAURANT#GENERAL" polarity="negative"
from="51" to="56"/>
</Opinions>
</sentence>
```

The restaurant reviews dataset contains 6 entities (Food, Drink, Restaurant, Service, Ambiance and Location) and 5 attributes (Price, Quality, General, Style and Options).

H. Laptop Review Dataset

This data set mainly retrieved from the Amazon.com which is belongs to the laptop domain and contains 450 English reviews. The data set contains 1739 training and 725 test sentences. Sample data capture is given below,

```
<sentence>
<text>
The retania is great, its amazingly fast when it
boots up because of the SSD  storage and the clarity
of the screen is amazing as well.
</text>
<Opinions>
<Opinionscategory="LAPTOP#GENERAL"
polarity="positive"/>
<Opinions category="LAPTOP#OPERATION_PERFORMACE"
polarity="positive"/>
<Opinions category="HARD_DISK#DESIGN_FEATURE"
polarity="positive"/>
</Opinions>
</sentence>
```

The laptop reviews dataset contains 22 entities (e.g., LAPTOP, SOFTWARE, SUPPORT), and 9 attributes (i.e., PRICES, QUALITY, STYLE OPTIONS, etc.).

The training dataset contains 1654 aspect category annotations, and the test dataset contains 845 aspect category annotations. In the laptops domain, 81 unique aspect categories were annotated in the training set and 58 in the test set. LAPTOP is the majority entity class in sets, 62.36% in training, and 72.81% in test data. The remaining 37.64% of the annotations in the laptops training data correspond to 72 categories with frequencies ranging from 6.53% to 0.05%. In the test set, the remaining 27.19% of the annotations correspond to 49 categories.

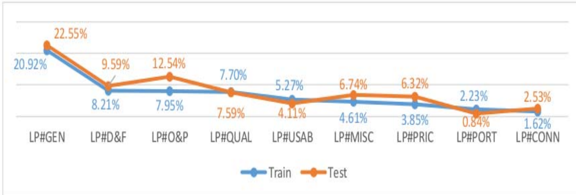

Fig 2: Data distribution in Laptop domain

Similar to previous dataset, each sentence contains aspect category specific polarities from a set P = {positive; negative; neutral}.

As a result of this research, mainly algorithms are evaluated using the much known evaluation models in the sentiment analysis domain. For this purpose precision, recall, F measure and accuracy used to identify the correctness of the classification.

Mainly recall criteria indicated the quality of the result and precision reflects the exactness of the quality of the results. F measure is the mean of precision and recall and also it is the indication of the accuracy [11]. For this evaluation criteria's, following indicated base figure is used.

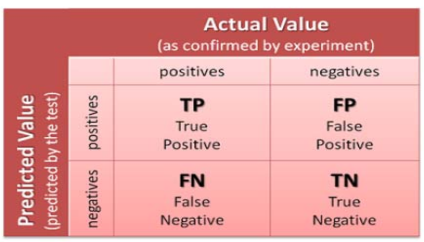

Figure 3: Evaluation Criteria

III. Results and Discussions

Mainly via this research following two subtasks will be analysed which owns to Aspect Based Sentiment Analysis.

A. Aspect Category Detection Module

As described in the Methodology section, this research mainly examines the binary SVM (Support Vector Machine) for each category with various parameters depending on the behaviour of the data set. Restaurant Review dataset is used as the first analysis for this research. For this evaluation procedure One VS All, SVM algorithm is used.

Features	**Precision**	**Recall**	**F-Score**	**Accuracy [%]**
Unigrams	0.6307	0.6012	0.6155	61.54
Unigrams+Stemmed	0.6402	0.6102	0.6248	61.98
Unigrams+Stemmed+Bigrams	0.6634	0.6214	0.6417	64.08
Unigrams+Stemmed+Bigrams+ HeadWords	0.6828	0.6466	0.6642	66.12
Unigrams+Stemmed+Bigrams+ HeadWords +WordCluster	0.7682	0.6987	0.7318	74.10
Unigrams+Stemmed+Bigrams+HeadWords +WordCluster+POS	0.7824	0.6993	0.7385	75.78
Unigrams+Stemmed+Bigrams+HeadWord +WordCluster+POS+ HeadWord POS	0.7615	0.6522	0.7026	72.89

Table 1: Measures of Restaurant domain

Same mechanism used to analyze the laptop review dataset as well.

Features	**Precision**	**Recall**	**F-Score**	**Accuracy[%]**
Unigrams	0.4867	0.4212	0.4515	45.94
Unigrams+Stemmed	0.4896	0.4304	0.4580	46.98
Unigrams+Stemmed+HeadWords	0.4902	0.4405	0.4640	47.54
Unigrams+Stemmed+HeadWords+POS	0.4912	0.4418	0.4651	47.64
Unigrams+Stemmed+HeadWords+POS+HeadWord POS	0.5084	0.4512	0.4780	48.18
Unigrams+Stemmed+HeadWords+POS+HeadWord POS+WordCluster	0.5452	0.4812	0.5112	52.84
Unigrams+Stemmed+HeadWords+POS+HeadWord POS+WordCluster +Bigram	0.5301	0.4766	0.5019	51.63

Table 2: Measures of Laptop domain

As described in the Methodology section, this research mainly examines the binary SVM (Support Vector Machine) for each category with various parameters depending on the behaviour of the data set. Restaurant Review dataset is used as the first analysis for this research. For this evaluation procedure One VS All, SVM algorithm is used.

	Positive	**Negative**	**Neutral**	**Accuracy [%]**
Bag Of Words	0.69	0.573	0	66.33
Elongated Words	0.723	0.583	0.001	68.34
Punctuation	0.74	0.589	0.001	69.28
Capitalization	0.768	0.603	0.012	72.36
Negation	0.783	0.645	0.102	75.02
POS	0.751	0.532	0.054	74.54
Lexicon	0.793	0.686	0.163	77.12
Domain Specific	0.844	0.702	0.196	82.32

Table 3: Results with SVM for SAM

	Positive	Negative	Neutral	Accuracy[%]
Lexical				
Bag Of Words	0.64	0.542	0.029	65.28
Elongated Words	0.683	0.563	0.055	67.34
Punctuation	0.73	0.589	0.78	70.22
Capitalization	0.768	0.611	0.102	72.00
Negation	0.773	0.645	0.132	76.92
POS	0.781	0.666	0.154	76.54
Lexicon	0.793	0.686	0.244	79.56
Domain				
Domain Specific	0.864	0.772	0.268	84.46

Table 4: Results with Logistic Regression for SAM

For the aspect category extraction task, 75.78 scored for the restaurant dataset compared to the state of art score of 83.98. Laptop domain archived five points higher than the art score. Overall all the measures of restaurants domain are relatively higher than the laptops domain. The main reason for these phenomena is that the classification schema in restaurants domain is fine grained and the laptops domain course grained.

In Sentiment Analyser module both classifiers performed well in the laptops domain than the restaurants domain. And both systems performed better in laptops domain than the restaurants domain. This is probably due to the fact that in the restaurants domain the positive polarity is significantly more frequent in the training than in the test data, which may have led to biased models. For aspect sentiment prediction, score of 72.39 for the laptops domain is using Logistic regression as the base classifier and score of 0.7318 using SVM as the base classifier.

IV. Conclusion

Aspect Based Sentiment Analysis focuses to go beyond a mere word level analysis of text and provide a more semantic analysis of text through the use of web on sources or semantic networks. This empowers novel approaches to sentiment analysis. In aspect-based sentiment analysis (ABSA) the aim is to identify the aspects of entities and the sentiment expressed for each aspect. The ultimate goal is to be able to generate summaries listing all the aspects and their overall polarity. The natural language processing tools can be used to facilitate the SA process. It gives better natural language understanding and thus can help produce more accurate results of SA.

The other major contribution of this study is the classification methodology. For the aspect category extraction, usage of SVM binary classification with selected lexical and semantic features is proposed and it is identified as an optimal feature of the machine learning technique. A mixture of lexical features such as unigrams, bigrams, POS, and head words with domain specific word clusters will be using for the performance evaluation. Sentiment polarity detection task is formulated as a multi-class classification problem. On top of features generated from, domain specific sentiment lexicons and general purpose sentiment lexicon, this research focus SVM and logistic regression algorithm to be analysed the performance of the system.

V. References

[1] Wikipedia. "Fundamental human needs". Internet: https://en.Wikipedia .org/wiki/Fundamental_human_needs [Apr. 29, 2016].

[2] Wikipedia."Maslow's hierarchy of needs". Internet: https://en.wikipedia. org/wiki/Maslow%27s_hierarchy_of_need s J [Apr. 29, 2016].

[3] Ay Gseuhensy. "Survey: 90% Of Customers Say Buying Decisions Are Influenced By Online Reviews". Internet: http://marketingland.com /survey-customers-more-frustratedby-how-long-it-takes-to-resolve-a-customer-service-issue-thanthe-resolution-38756, Oct. 25, 2015 [Apr. 29, 2016].

[4] Bo Wang and Min Liu. 2015. Deep Learning For AspectBased Sentiment Analysis. Stanford University report, https://cs224d.stanford .edu/reports/WangBo.pdf.

[5] Ahmed Hassan "Sentiment analysis algorithms and applications: A survey" [On- line]. Available:http://ac.elscdn.com/S2090447914000550 /1-s2.0-S2090447914000550main.pdf? [May. 01, 2016].

[6] Wikipedia. "Supervised learning. Internet: https://en.wikipedia.org/wiki /Supervised_learning Jan 2013[Apr. 29, 2016].

[7] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:127:27, 2011. Software available at http://www. csie.ntu.edu.tw/ cjlin/libsvm.

[8] Chin Chen Chien, Tseng You-De. Quality evaluation of product reviews using an information quality framework. Decis Support Syst 2011;50:75568.

[9] Ko Youngjoong, Seo Jungyun. Automatic text categorization by unsupervised learning. In: Proceedings of COLING-00, the 18th international conference on computational linguistics; 2000.

[10] Miller G, Beckwith R, Fellbaum C, Gross D, Miller K. WordNet: an on-line lexical database. Oxford Univ. Press; 1990.

[11] Saias, Jos, and Rua Romao Ramalho. "Sentiue: Target and Aspect based Sentiment Analysis in SemEval-2015 Task 12."

[12] Thorsten Joachims. Transductive inference for text classification using sup-port vector machines. In ICML, volume 99, pages 200209, 1999.

[13] Kushal Dave, Steve Lawrence, and DavidMPennock. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In Proceedings of the 12th international conference on World Wide Web, pages 519528. ACM, 2003.

[14] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In Proceedings of the ACL- 02 conference on Empirical methods in natural language processing-Volume 10, pages 7986. Association for Computational Linguistics, 2002.