University of Ibadan
Journal of Science and Logics in ICT Research

A Model for Tracking Sentiment in Reviews at Aspect Level Using Support Vector Machines

Obayagbona I. E. and Adeyemo A. B.

Department of Computer Science, University of Ibadan, Nigeria.

Abstract

In this era of internet and online services, opinionated data are generated at an incredible amount. Wading through these huge amounts of opinions to glean relevant insight or information for decision making is a daunting task. Sentiment Analysis is a vital tool in Natural Language Processing (NLP) that has made it possible to understand people's opinions on different topics. In this study, a level of sentiment analysis known as Aspect Level Sentiment Analysis was used to identify what exactly people are talking about in reviews (such as aspects or features of a subject). A model for Aspect Level Sentiment Analysis using Support Vector Machines was developed. Datasets from a customer review area (Laptop products) was used to train and evaluate the developed model. Results obtained gave an appreciable performance showing concise reviews and accuracy in categorizing opinions on polarity scale (positive or negative).

Keywords: *Sentiment Analysis, Support Vector Machines, Aspects (Features), Polarity Scale.*

1. INTRODUCTION

The overwhelming growth and importance of the internet has made large volume of data available. Essentially, these data are generated by users who meet, express and share views on various issues on platforms provided. The volume of views or opinions about products, services, politics, etc. on these different platforms (social networking sites, blogs, discussion groups, review sites) grows very rapidly making it near impossible to parse and organize. As a result, mining opinions from data online has seen an increasing attention in recent times.

The task of analysing huge amount of opinionated data is very difficult and time consuming. Wading through numerous opinionated data to glean relevant insight or information is daunting both for the reader and targeted individual or company. For example, a potential customer might be discouraged having to read through hundreds of product reviews in order to make decision on whether to buy the product while the manufacturer of the product, keeping track and managing customer reviews is a tedious job. To tackle this problem, it will suffice to build a system that will ease the processing or the mining of extensive reviews reliably into categories of interest.

Generally, one way of mining or extracting opinion from reviews is to utilize sentiment analysis (also known as opinion mining). Sentiment analysis is the field of study that analyses people's opinions, sentiments, appraisals, attitudes, and emotions toward entities and their attributes expressed in written text. The entities can be products, services, organizations, individuals, events, issues, or topics [1]. Simply put, it identifies and

Obayagbona I. E. and Adeyemo A. B. (2020). A Model for Tracking Sentiment in Reviews at Aspect Level Using Support Vector Machines, *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR),* Vol. 4 No. 1, pp. 96 - 105

understands the feelings of user towards an entity through the use of Natural Language Processing (NLP).

2. LITERATURE REVIEW

Sentiment analysis mainly studies opinions that express or imply positive or negative sentiment. According to [2], sentiment analysis can be done at three different levels:

- Document Level: In this level, the overall opinion about the document is found and classified as positive, negative or neutral.
- Sentence level: Here, each sentence in the document is analysed, classifying them as positive, negative or neutral.
- Feature or Aspect level: At this level, fine grained analysis is performed for each feature present in the document or sentence and classified the opinion expressed as positive, negative, or neutral.

Document level and sentence level only classify the whole document or sentence; it does not identify the aspect present in the document or sentence. If the polarity of the document is positive or negative that does not mean document possess positive or negative opinion for each aspect. To determine the opinion on every aspect, opinion mining at aspect level is performed.

Most common technique for detecting aspect is Frequency-based approach. It identifies the frequency of word in the review text. These words are nouns and noun phrases in sentences of the given review. The most frequent nouns are counted and considered as potential aspect.

This approach was used in extracting aspects or aspect words from customer reviews [3]. Frequent nouns and noun phrases were regarded as aspects and identify using Part-of-Speech (POS) tagger. Occurrence of these noun and noun phrase is counted, if more than the predefined threshold value then they are considered to be frequent aspects. Further improvement on Hu and Liu [3] system remove noun phrases that are not aspect. Using Pointwise Mutual Information (PMI) score, the semantic relation between the phrase and the aspect was computed to obtain a true aspect [4].

An approach that uses models to extract aspects from reviews as well as sentiment polarity is the Model-based approach. The approach learns patterns from data which is then applied to new dataset. Most commonly used models are based on supervised learning techniques such as the Conditional Random Field (CRF) and Hidden Markov Model (HMM) and those based on unsupervised learning (or Topic Modeling) techniques are Probabilistic Latect Semantic Analysis (pLSA) and Latent Dirichlet Allocation (LDA).

Conditional Random Field (CRF) was employed to extract aspects from sentences which contain an opinion expression [5]. Features like token, Part-of-speech, short dependency path, word distance and opinion sentence were use as input for the CRF-based approach. Token is the current token string in sentence represented as potential aspect, Part-of-speech represents tag of the current token as identified by the Stanford POS Tagger, and Short Dependency Path captures relationship between sentiment expressions and their aspects syntactically. This is achieved by recognizing and labelling all the tokens which have direct dependency link to sentiment expression in a sentence. Word Distance captures opinion expressions and their aspects not connected by short paths in the dependency parse tree of the above mentioned feature. This feature locate the closest tokens or noun phrases to each sentiment expression in a sentence and finally Opinion Sentence, a feature that help label all tokens that occurred in a sentence with sentiment expression and the feature allow CRF algorithm to differentiate between the certain token that occurred in a sentence containing sentiment from the sentence without a sentiment. At the end, aspects were extracted from sentences containing sentiment expression which is then model as sequence segmentation and labelling task. This CRF algorithm receives a sequence of tokens $t_1...t_n$ for which it has to predict a sequence of labels $l_1...l_n$ as sentiment aspects.
A Lexicalized Hidden Markov Model was used to learn patterns to extract aspects and sentiment expressions [6]. They integrated linguistic features such as part-of-speech,

phrases formation patterns and surrounding contextual clues of words or phrases into the learning. Two major tasks were carried out in this process: they are Recognition and Classification. Recognition identifies sentences expressing sentiments while Classification categorize these sentences into sentiment words or phrases-aspects and also determine the polarity of the sentiment word whether it is positive or negative.

The process of discovering topics from text documents is known as Topic Modeling. This method output a set of word clusters and a topic distribution for each document. Each word cluster is called a topic and is a probability distribution over words (also called topical terms) in the document collection. This kind of method is an unsupervised learning method and pLSA (Probabilistic Latent Semantic Analysis) and LDA (Latent Dirichlet allocation) are the two basic models. In the context of sentiment analysis, discovered topics from Topic models are aspects. Hence, Topic Modeling can be used for aspect extraction. However, Topic Modeling help in aspect grouping and cover both aspect and sentiment words but in sentiment analysis, they need to be separated. Such separations can be achieved by extending the basic model (e.g. LDA) to jointly model both aspects and sentiments.

Mei et al [7] proposed a joint model for sentiment analysis. Specifically, they built an aspect-sentiment mixture model, which was based on an aspect (topic) model, a positive sentiment model, and a negative sentiment model learned with the help of some external training data. Their model was based on pLSA.

Zhao et al. [8] combined the Maximum Entropy and Latent Dirichlet Allocation (MaxEnt-LDA) to form a hybrid model which jointly discovers both aspect words and aspect specific opinion words. The model used syntactic (word context) features to help separate aspects and sentiment words. Aspect and opinion word separation in MaxEnt-LDA is achieved through an indicator variable (also called a switch variable) drawn from a multinomial distribution controlled by a set of parameters. This indicator variable determines whether a word in a sentence is an aspect word, an opinion word, or a background word. It uses Maximum Entropy on some labelled training data to learn the parameters of the distribution from which the indicator variable's value is drawn. A second indicator variable is also used to determine general and specific types of aspect or opinion. For example, in a restaurant review, each word in a sentence of the review can be one of the few types. The word may be a specific aspect word (e.g., waiter for the staff aspect), a general aspect word (e.g., restaurant), an opinion word specific to the aspect (e.g., friendly), a generic opinion word (e.g., great), or a commonly used background word (e.g., know).

In this work, an Aspect level sentiment analysis that classifies the user reviews for different aspect of a product on polarity scale (positive or negative) depicting the actual sentiments of user toward product is done and then generate summary of opinions. This is very useful because a lot of user opinions can be considered within a very short amount of time which is important for human.

3. MATERIALS AND METHODS

The developed model was based on supervised machine learning technique. The model is divided into three modules, each one addressing the different process of solving Aspect Level Sentiment Analysis problem: (a) Text Data Collection module where the review data or text are collected and transformed into appropriate form suitable for analysis, (b) Aspect/Sentiment Extraction module that focuses on the extraction of aspect and sentiment words from review text or sentences and (c) Sentiment Classification module in charge of computing the sentiment polarity of the discovered aspects. The schematic diagram of the model is show in Figure 1.

3.1 Text Data Collection

Laptop review datasets were taken from the research which was conducted for 2015 and 2016 SemEval Challenge. These downloaded datasets contain 4239 labelled review sentences that were used for training and testing the system.

3.2 Text Normalization and Pre-processing

This involves preparing the review text for processing. The review text were transform removing irrelevant documents (such as a period, hash tags, a newline character, semi-colon, etc.), converting all words to lower case, removing numbers and word contractions like *it's, doesn't to "it is and does not*".
The text pre-processing helps standardizes text for analytical task using variety of techniques like tokenization (splitting text into smaller meaningful components or words called tokens), lemmatization and stemming (reducing words to their base form and removing prefixes and suffixes from words), etc. These techniques in text pre-processing will be performed along with the process of extracting aspect and sentiment words in section B. The schematic diagram of Text normalization workflow is shown in Figure 2.

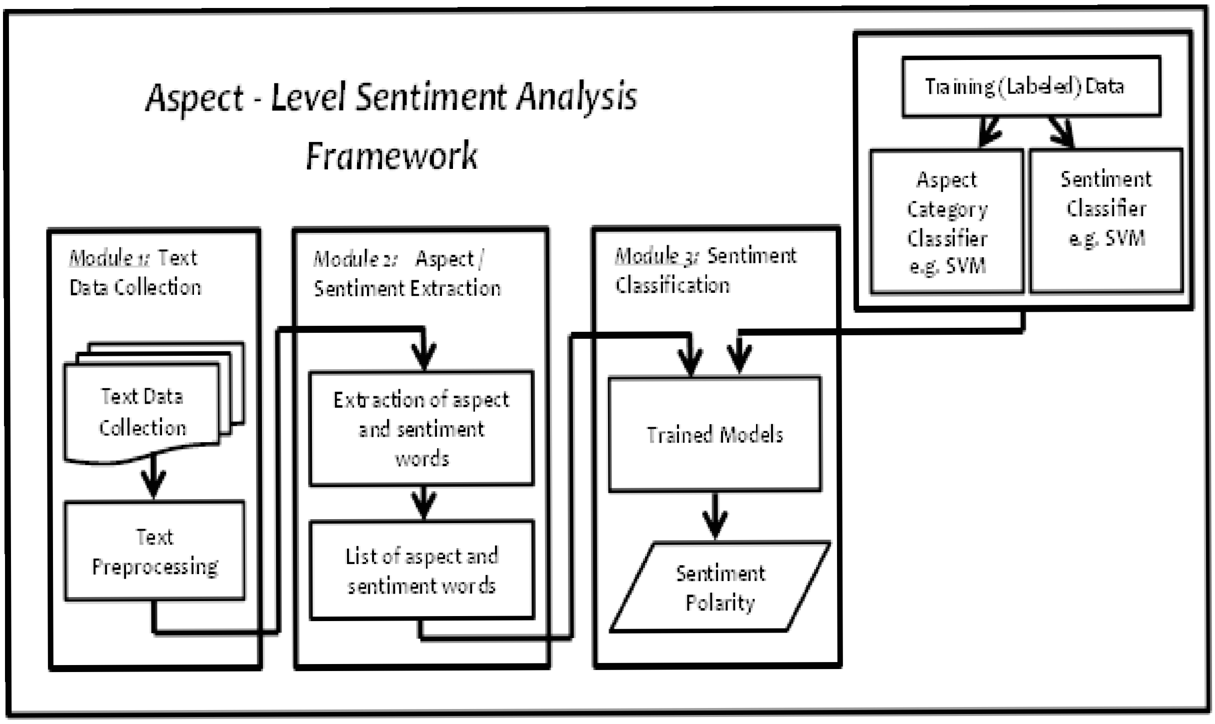

Figure 1: Framework for Aspect-Level Sentiment Analysis

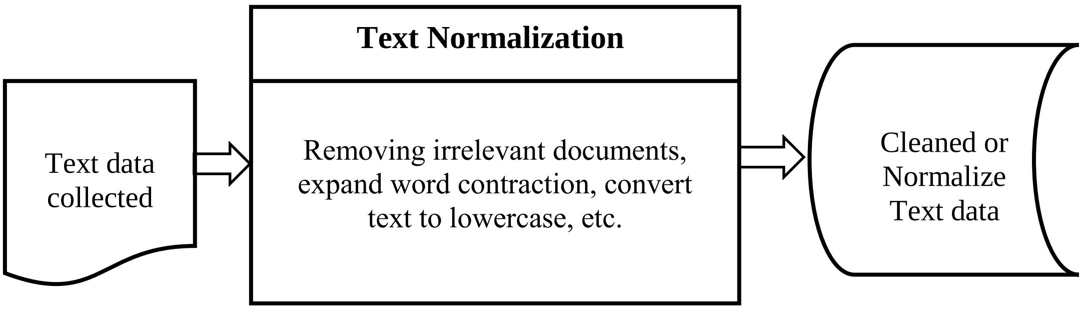

Figure 2: Text Normalization Workflow

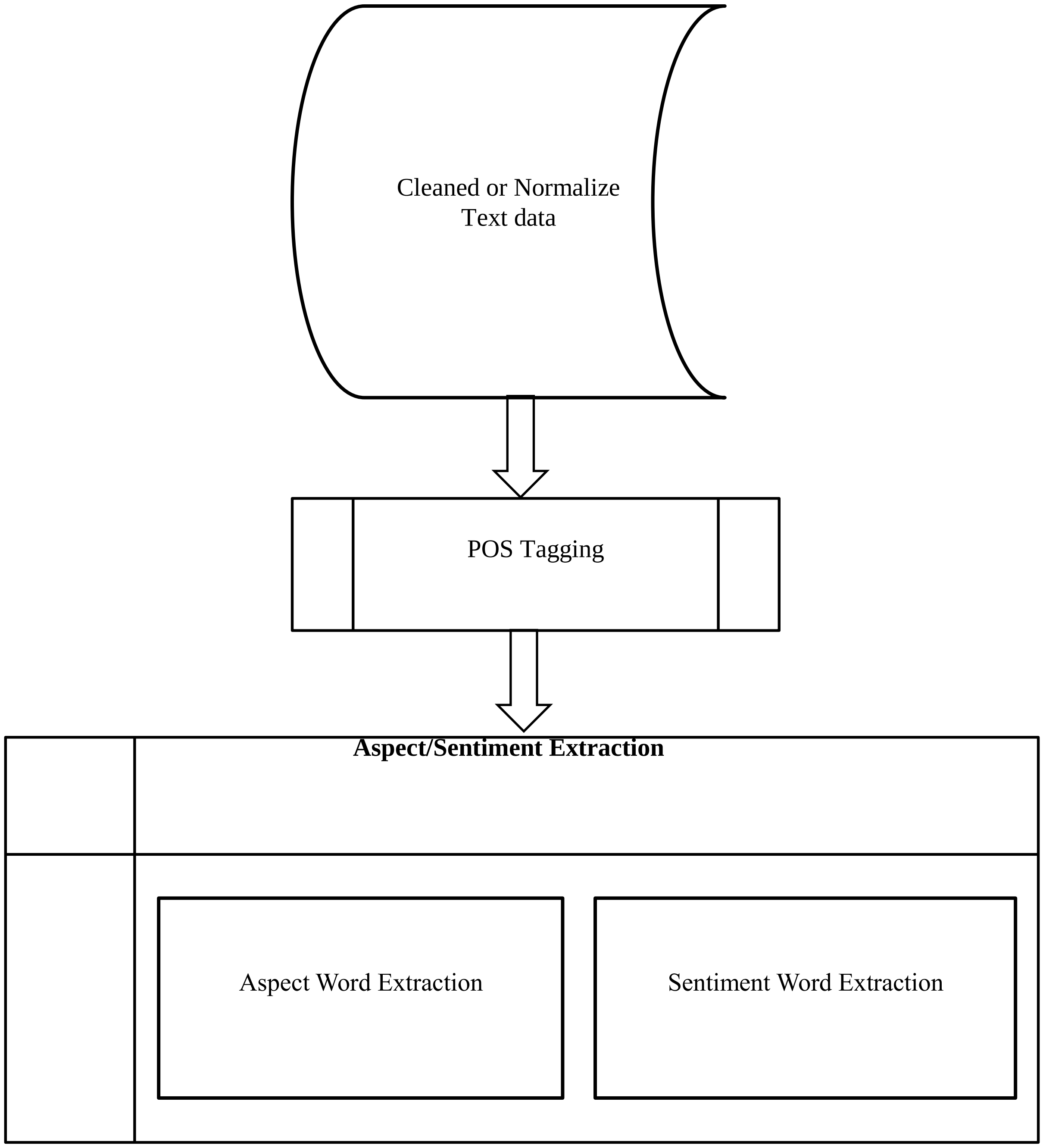

Figure 3: Aspect/Sentiment Extraction Workflow

3.3 Aspect/Sentiment Extraction

The decomposition of the review text or sentences into features is a focal task of any Aspect Level Sentiment Analysis. That is feature extracted from given customer review sentences will serve as input to the system. The feature extraction is in two categories:

1. Aspect Word Extraction
2. Sentiment Word Extraction

3.3.1 The Aspect Word Extraction: This is vital to analysing aspect level sentiment. Aspect words are specific words customers express opinion about and are usually nouns and noun phrases. Part of Speech (POS) tagging is used to determine the aspects and sentiment words from the reviews and the POS tagger of a natural language toolkit known SpaCy is use to tag all the words of the sentences to their appropriate part of speech. A list of potential aspect words for each review was obtained.

Consider the review, "*The processor is great but the battery life is pathetic."*
Here, processor and battery life which are noun and noun phrase in the review pertaining to the laptop or computer system domain are the aspects.

3.3.2 The Sentiment Word Extraction: The objective of this phase is to detect the sentiment words expressed toward a given aspect in a given review. Sentiment words are usually adjectives and sometime verbs. The same Part of Speech (POS) tagger of SpaCy is used to list of potential sentiment words for each review.

Considering the same review, *"The processor is great but the battery life is pathetic."*

The adjectives in this review are great and pathetic and these are regarded as the sentiment word expressed towards the aspects extracted the previous phase. Figure 3 show the Aspect/Sentiment Extraction workflow.

3.4 Sentiment Classification

The goal of Sentiment Classification module is to assign sentiment polarity label (Positive, Negative or Neutral) against a single input or review. Each review is transformed to list of aspect terms and sentiment terms, generated from Aspect/Sentiment Extraction module, which is passed to trained classifiers (Aspect Category and Sentiment Polarity classifiers) and then classified as positive, negative or neutral.

This sentiment classification phase was made possible by prior training of two classifiers namely Aspect Category model and Sentiment Polarity model. The Aspect Category model categorize the extracted aspects of a review into a particular category while the Sentiment Polarity model finally detect the sentiment expressed towards a given aspect category in the customer review under consideration. This training of classifier is further explained the next section.

3.5 Training Classifiers

Laptop review datasets were taken from the research which was conducted for 2015 and 2016 SemEval Challenge were merged to give 4239 training reviews. The dataset includes annotations of aspect categories and sentiment classes of laptop reviews. Here is a review sample:

```
<sentence id="79:2">
   <text>10 plus hours of battery...</text>
    <Opinions>
      <Opinion
category="BATTERY#OPERATION_PERFO
RMANCE" polarity="positive"/>
    </Opinions>
</sentence>
```

From this sample, the aspect category and sentiment attributes were extracted. The xml text tag holds the customer review and Opinion tag holds the aspect (BATTERY) and its category (OPERATION_PERFORMANCE) together and also its sentiment in this case polarity is positive. In the laptop review, 82 unique aspect categories were annotated with LAPTOP#GENERAL in majority.

3.6 Building Aspect Category Model

From each of the review text, aspect terms are identified or extracted. These aspects along with the aspect categories is transform and then pass to a machine learning algorithm e.g. a linear support vector classier as variant of Support Vector Machines (SVM). The algorithm learns the relationship between the aspects and the aspect categories so that it can appropriately classify new review (not previous seen review) into the right class. Below is a pseudo-code for building this aspect category classifier:

PSEUDO CODE:

Step 1: Get extracted aspect words and aspect categories.

Step 2: Transform aspect words and aspect categories to ML acceptable form.

Step 3: Finally, train ML algorithm to build a model.

3.7 Building Sentiment Polarity Model

The same extracted data from the xml file explained in the previous section is used. But this time instead of aspect categories, sentiment polarity (positive, negative and neutral) were extracted. These sentiment polarities along with the aspect terms extracted from review text are passed to support vector machine in order to learn the relationship between them for classification sentiment present in review. The following is a pseudo-code for building this sentiment polarity classifier:

PSEUDO CODE:

Step 1: Get extracted sentiment words and sentiment polarity.

Step 2: Transform sentiment words and sentiment polarity to ML acceptable form.

Step 3: Finally, train ML algorithm to build a model.

3.8 Building Aspect-Level Sentiment Model

Finally, the Aspect level Sentiment Model is built by combining the Aspect Category Model and Sentiment Polarity Model together to track sentiment express towards an aspect in a review text. A new dataset or some customer reviews - prepared by going through the text pre-processing and aspect/sentiment extraction as discussed in previous sections - is passed to this combined models to give the sentiment or sentiment polarity associated with each review.

4. RESULTS AND DISCUSSION

The Aspect Level Sentiment Analysis is a classification problem and the evaluation metrics criteria used for measuring how well the machine learning models perform are precision, recall, f1-score and accuracy. For these evaluation criteria, confusion matrix is used which help to summarized how well the models have perform in classifying given instances. It shows the number of correct and incorrect classification made by the model compared to the actual outcomes (target values) in the data. Based on this, classification reports are computed and presented.

Table 1 shows the confusion matrix for Aspect Category classifier or model where 164 of LAPTOP#GENERAL, 23 of BATTERY#OPERATION_PERFORMANCE and 53 of LAPTOP#QUALITY have been classified correctly. Based on this selected subset used for the overall confusion matrix, the classification report is shown in Table 2

Table 1: Confusion Matrix for Aspect Category Classifier

Confusion Matrix for a Subset of Aspect Category

Actual Class \ Predicted Class	LAPTOP#GENERAL	BATTERY#OPERATION_PERFORMANCE	GRAPHICS#GENERAL	HARD_DISC#DESIGN_FEATURES	LAPTOP#USABILITY	CPU#OPERATION_PERFORMANCE	LAPTOP#OPERATION_PERFORMANCE	LAPTOP#MISCELLANEOUS	LAPTOP#QUALITY	OS#DESIGN_FEATURES
LAPTOP#GENERAL	164	1	0	0	1	2	6	6	0	0
BATTERY#OPERATION_PERFORMANCE	0	23	0	0	0	0	0	1	0	0
GRAPHICS#GENERAL	0	0	2	0	0	0	0	0	0	0
HARD_DISC#DESIGN_FEATURES	0	0	0	2	0	0	0	0	0	0
LAPTOP#USABILITY	8	0	0	0	27	0	3	3	0	0
CPU#OPERATION_PERFORMANCE	0	0	0	0	0	3	0	0	0	0
LAPTOP#OPERATION_PERFORMANCE	16	0	0	0	2	0	48	0	0	0
LAPTOP#MISCELLANEOUS	9	2	0	0	0	1	0	61	0	0
LAPTOP#QUALITY	20	0	0	0	0	0	0	0	53	0
OS#DESIGN_FEATURES	0	0	0	0	0	0	0	0	0	0

Table 2: Classification Report for Aspect Category Classifier

Classification Report for a Subset of Aspect Category

	precision	recall	f1-score	support
LAPTOP#GENERAL	0.66	0.85	0.74	192.00
BATTERY#OPERATION_PERFORMANCE	0.74	0.82	0.78	28.00
GRAPHICS#GENERAL	0.50	1.00	0.67	2.00
HARD_DISC#DESIGN_FEATURES	1.00	1.00	1.00	2.00
LAPTOP#USABILITY	0.84	0.52	0.64	52.00
CPU#OPERATION_PERFORMANCE	0.50	1.00	0.67	3.00
LAPTOP#OPERATION_PERFORMANCE	0.77	0.66	0.71	70.00
LAPTOP#MISCELLANEOUS	0.78	0.79	0.79	77.00
LAPTOP#QUALITY	0.95	0.68	0.79	78.00
OS#DESIGN_FEATURES	0.00	0.00	0.00	0.00
micro avg	0.74	0.76	0.75	504.00
macro avg	0.67	0.73	0.68	504.00
weighted avg	0.76	0.76	0.74	504.00

Accuracy

```
# compute classifiction accuracy for aspect extraction classifiers
from sklearn import metrics

# Compare actual response values (y_test) with predicted response values (y_pred)
print("Accuracy: {0:.4f}".format(metrics.accuracy_score(y_test, y_pred)))
```

```
Accuracy: 0.7649
```

Figure 4: Overall Accuracy for Aspect Category Classifier

Overall, the Aspect Category Classifier has yielded an accuracy of 76.49% (Figure 4) in classifying the aspects into different classes with considerable high values of weighted average for recall and precision (Table 2). However, the weighted average for precision (0.76) indicates that the classifier has misclassified about 24% of the categories.

Table 3 show the confusion matrix for the Sentiment Polarity Classifier. The classifier predicted 431 positive, 317 negative and 38 neutral reviews correctly. It achieved an overall accuracy of 86.75% (Figure 5), and the evaluation metrics detailed in Table 4. According to the values of precision and recall, the results show that the model performs well in analysing either positive or negative reviews with weighted average of 87%.

Table 3: Confusion Matrix for Sentiment Polarity Classifier

Confusion Matrix for Sentiment

Actual Class \ Predicted Class	positive	negative	neutral
positive	431	35	6
negative	49	317	4
neutral	13	13	38

Table 4: Classification Report for Sentiment Polarity Classifier

Classification Report for Sentiment

	precision	recall	f1-score	support
positive	0.87	0.91	0.89	472.00
negative	0.87	0.86	0.86	370.00
neutral	0.79	0.59	0.68	64.00
micro avg	0.87	0.87	0.87	906.00
macro avg	0.84	0.79	0.81	906.00
weighted avg	0.87	0.87	0.87	906.00

Accuracy

```
# compute classifiction accuracy for sentiment extraction classifiers
from sklearn import metrics

#print('Accuracy:', metrics.accuracy_score(y_test, y_pred))
print("Accuracy: {0:.4f}".format(metrics.accuracy_score(y_test, y_pred)))

Accuracy: 0.8675
```

Figure 5: Overall Accuracy for Sentiment Polarity Classifier

In summary, taking the accuracies of both classifiers (76.49% + 86.75%) into consideration, an average accuracy of 81.62% is achieved which show a promising performance or result during the evaluation phase. This has provided supporting evidence for the accuracy and reliability of the Aspect Level Sentiment Analysis framework.

5. CONCLUSION

The objective of this work is to determine the polarity of customer reviews at Aspect level. A model which performs the Aspect based sentiment analysis on reviews and generates the summarized results was developed. The developed model will be helpful for the user in taking the right decision amidst huge volume of reviews users are bombarded with online nowadays, since users are usually busy and do not have a enough time to read all positive or negative reviews.

REFERENCES

[1] Liu B., (2015). Sentiment Analysis (Mining Opinions, Sentiments, and Emotions), Cambridge University Press, pp. 381.

[2] Thellaamudhan C, Suresh R, Raghavi P., (2016). A Comprehensive Survey on Aspect Based Sentiment Analysis. International Journal of Advanced Research in Computer Science and Software Engineering, Vol.6, Issue 4, pp. 442 – 447.

[3] Hu, M. and Liu, B. (2004). Mining and Summarizing Customer Reviews, Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), pp. 168-177.

[4] Popescu, Ana-Maria and Oren Etzioni (2005). Extracting Product Features and Opinions from Reviews. In Processings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2005).

[5] Jakob, N. and Gurevych, I. (2010). Extracting Opinion Targets in a Single and Cross-Domain Setting with Conditional Random Fields, Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, pp. 1035–1045.

[6] Jin, W. and Ho, H. (2009). A Novel Lexicalized HMM Based Learning Framework for Web Opinion Mining, Proceedings of the 26th International Conference on Machine Learning, pp. 465-472.

[7] Mei, Q., Xu L. and Zhai C. (2007). Topic Sentiment Mixture: Modeling Facets and Opinions in Weblogs, Proceedings of International Conference on World Wide Web (WWW-2007).

[8] Zhao, W., Jiang, J., and Li, X. (2010). Jointly Modeling Aspects and Opinions with a MaxEnt-LDA hybrid, Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2010), pp. 56–65.