ISSN (Print): 0974-6846 ISSN (Online): 0974-5645

Sentiment Analysis of Product Reviews using Support Vector Machine Learning Algorithm

Esha Tyagi* and Arvind Kumar Sharma

Career Point University, Kota – 324005, Rajasthan, India; eshatyagi26@gmail.com, drarvindkumarsharma@gmail.com

Abstract

Now days, there is a huge increase in number of peoples who have been accessing many social networking sites and microblogging websites which open a new door to the impression of today's generation. Various user reviews for a specific product, company, brand, individual, forums and movies etc. have been much helpful in judging the perception of people. Thus the analysts took the initiative to develop algorithms to automate the classification of distinctive reviews on the basis of their polarities particularly: Positive, and Negative. This automated classification mechanism is referred as Sentiment Analysis. The primary aim of this paper is to apply Support Vector Machine (SVM) machine learning algorithm to classify the sentiments and texts for product reviews that analyses different datasets used for classification of sentiments and texts. Furthermore, various data sets have been utilized for training as well as testing and implementing the Support Vector Machine learning algorithm to find the polarity of the ambiguous sentiments. **Objectives**: The primary objective of this paper is to implement Support Vector Machine (SVM) machine learning algorithm to classify the sentiments and texts for product reviews that analyses different datasets used for classification of sentiments and texts. Methods/Statistical Analysis: In this research study, several datasets have been applied for training as well as testing and simulating the support vector machine learning algorithm to compute the polarity of the ambiguous sentiments or reviews. Findings: At the outset, Support Vector Machine (SVM) classification algorithm gives higher accuracy 89.98% than the other ones. The obtained accuracy would be enhanced further by including more sentence forms. Finally, it concludes that the Support Vector Machine (SVM) algorithm behaves well. Application/Improvements: The performance resulting models are tested to measure accuracy of Support Vector Machine learning algorithm. Finally, the SVM classification algorithm has been achieved high accuracy and found better classification algorithm than others.

Keywords: Clustering, Product Reviews, SVM, Sentiment Analysis

1. Introduction

Machine learning is a logical train that investigates the development and investigation of calculations that can gain from data¹². Such calculation work by building a model in view of data sources and utilizing that to settle on forecasts or choices as opposed to following just expressly modified guidelines. Machine learning is closely related to computational statistics. Sentiment analysis approaches are classified into the following categories which are shown in Figure 1.

Technically, these machine learning algorithms can be classified into the following ways: Supervised,

Unsupervised, Semi-Supervised and Reinforcement Learning. If the different input objects are given with a labeled output value (also called supervisory signal) is called supervised machine learning, in contrast to unsupervised learning, where there is no such supervisory signal, between both supervised and unsupervised learning there is semi-supervised learning where there are large amount of input objects and only some of these objects are labeled with output value. However, in reinforcement learning there is a learning system to which the training information is provided by the environment based on which it has to discover which action will yield the best reward.

^{*}Author for correspondence

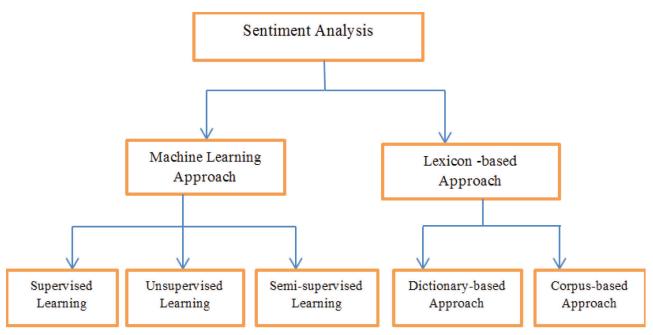


Figure 1. Classification of sentiment analysis approaches.

The rest of paper is organized as follows: Section 2 describes related work in the sentiment analysis domain. Section 3 covers the proposed work. Section 4 presents experimental evaluation along with simulation for sentiment analysis. Section 5 concludes the paper while references are mentioned in the last.

2. Related Work

This section provides a brief summary of many works performed in the field of sentiment analysis and opinion mining.

Tripathy et al.¹ used different supervised machine learning algorithm and by applying them he tried to classify movie reviews.

Maria et al.² given a hybrid approach for the prediction of sentiment in which context-sensitive coding offered by Word2Vec and sentiment/emotion information offered by a lexicon were combined and this was done to get the best results in terms of efficiency, accuracy and the process time.

Appel et al.³ proposed a hybrid system approach to the sentiment analysis problem at the sentence level and give a high level of accuracy of 88.02% and precision of 84.24%. Natural Language Processing (NLP) method is used. To three different data-sets hybrid method is proposed, and results achieved are compared to those using Naïve Bayes (NB) and Maximum Entropy (ME) technique.

Korkontzelos et al.⁴ presented work on the effect of sentiment analysis, on extracting adverse drug reactions from tweets and forum posts. Automatic identification of ADR was tough because of its large size, but through this method this was made easy.

Severyn et al.⁵ predicted polarities at both message and phrase level by deep learning approach to sentiment analysis of tweets, and for this, they used an unsupervised neural language model which trained initial word embedding, and further was tuned to find the polarities.

Nakov et al.⁶ presented a task on the evaluation of Compositional Distributional Semantics Models on full sentences organized for the first time within SemEval-2014. Its task had two sub-tasks, (i) semantic relatedness and (ii) entailment.

Gokulakrishnan et al.^Z analyzed tweets from micro blogging twitter site and classified them as positive, negative and irrelevant. And further, they studied the performance of various classifying algorithms.

Pak et al.⁸ proposed a work in which they create a method to train a sentiment classifier by automatic collection of a corpus. They used Tree Tagger for POS-tagging, through which difference between negative, positive and neutral sets of distributions was observed. For sentiment analysis and opinion mining purposes we used Twitter, the most popular micro blogging platform and used a method in which they automatically collected corpus.

Jansen et al.² presented a paper that describes if a company has to work on their branding strategy and then they should explore micro blogging websites because they are a potentially rich avenue for companies. In their paper the authors examined e-WOM branding. The authors further analyzed branding comments, sentiments, and opinions in more than 150,000 micro blog postings.

3. Proposed Work

The flow of proposed work is shown in Figure 2.

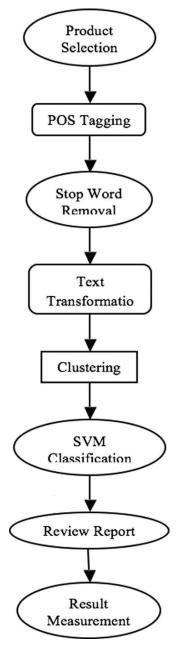


Figure 2. Flow of proposed work.

The proposed work contains the following key points:

- The proposed work discusses about applying sentiment analysis and machine learning algorithm to investigate the relationships among the online reviews for smart phone products and the revenue of performance.
- The algorithm is to be applied on product reviews and predict collection of the product based on the reviews and analyses how much effect the reviews have on the collection. The product collection for the next day is predicted based on online reviews of the present day.
- A prediction of high or low collection is also predicted.
 From the website, the detailed information containing the values for the following: brands, product date, rank of sale, user's reviews, etc. of a smart phone were obtained.
- Part of Speech (POS) model in which a sentiment or textual review is represented as a vector, whose entries correspond to individual terms of a vocabulary. Partof-speech information is supposed to be a significant indicator of sentiment expression.
- The score of each sentence in the dataset is calculated by sum of weight of each term in the corresponding sentences. Clustering of the review data based on the TF-IDF measure has been performed.
- Finally, the proposed work achieves high accuracy, the reviews are taken as appropriate and the success or failure of the smart phone product is predicted based on the reviews by using Support Vector Machine classification algorithm.

3.1 Simulation Work

The simulation work is performed by using Java programming language, WAMP Server and MySQL. The complete simulation is accomplished by using the following modules:

- Text preprocessing
- Transformation
- Clustering
- SVM classification
- Evaluation

3.1.1 Text Pre-Processing

Text pre-processing techniques are divided into two subcategories which are POS tagging and stop words removal. In POS, textual data comprises block of characters called tokens. The input reviews are separated as tokens and start the pre-processing. A stop-list is the name commonly given to a set or list of stop words. Some of the more frequently used stop words for English include "a", "of", "the", "I", "it", "you", and these are generally regarded as 'functional words' which do not carry meaning. Hence remove those words that support no information for the task.

3.1.2 Transformation

In the transformation process, the score for each sentence is calculated in the document. For that, first the weight of each term is calculated by the product of term frequency and inverse document frequency.

3.1.3 Clustering

Clustering of the document review is based on the TF-IDF measurement. Thus, points on the edge of a cluster, maybe in the cluster to a lesser degree than points in the center of cluster. It chooses the number of clusters and it finds centroid.

3.1.4 SVM Classification

After the removal of the outliers based on the clustering, the improved feature sets were used for sentiment classification. SVM is mainly used for the sentiment classification. It classifies the positive and negative reviews.

4. Evaluation

The performance of algorithm depends on several parameters such as Precision, Recall, F-Measure and Accuracy. This can be understood using True Positives, False Positives, True Negative and False Negative, where, True Positive (TP) is items correctly labeled as belonging to the positive class.

False Positives (FP) are items incorrectly labeled as belonging to the positive class True Negative (TN) are items correctly labeled as belonging to the negative class False Negative (FN) are items incorrectly labeled as belonging to the negative class ¹⁰.

Here, the precision is a measure of relevant instances among the retrieved instances; it is the ratio of number of elements correctly labeled as positive to total number of positively classified elements.

Precision =
$$\frac{TP}{TP + FP}$$

Whereas, Recall is a measure of how many truly relevant results are returned; it is the ratio of total number of positively labeled elements to total elements which are truly positive.

Recall =
$$\frac{TP}{TP + FN}$$
 13

F-Measure is a measure that combines precision and recall and is the harmonic mean of precision and recall.

F-Measure =
$$2.\frac{\text{Precision.Recall}}{\text{Precision} + \text{Recalll}}$$

Accuracy is also used as a statistical measure of how well a classification test correctly identifies or excludes a condition. That is, the accuracy is the proportion of true results (both true positives and true negatives) among the total number of cases examined.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}^{15}$$

Accuracy is a common measure for classification performance. It is the proportion of correctly classified examples to the total number of examples, while error rate uses incorrectly classified instead of correctly 16.

5. Experimental Evaluation

In this section, the experimental work has been simulated by using Java NetBeans IDE 7.2.1 open source software to perform Support Vector Machine classification algorithm. The complete experimental evaluation as well as simulation has been performed by using various modules.

The product reviews are now transferred to a Part of Speech Tagger (POS Tagger) which reads the words from the reviews and assigns each one as Noun, Verb, or Adjective etc.

Now, in the next module we do the complete process of POS tagging which is shown in Figure 4.

Now, the product reviews have been pre-processed by searching for many predefined Stopwords from product reviews as shown in Figure 5.



Figure 3. User interface of proposed work.

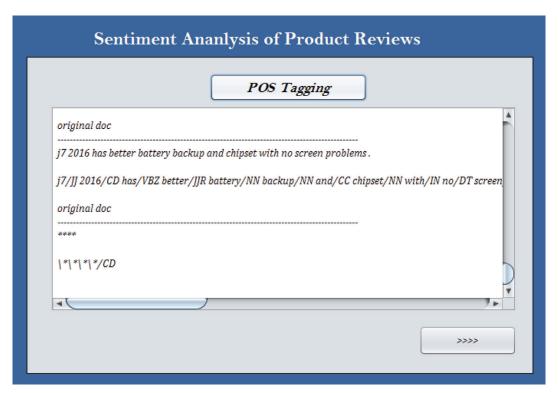


Figure 4. POS tagging process.



Figure 5. Data pre-processing of stopwords.

The found stopwords are now removed from the original product reviews which do not contribute to the product reviews are being positive or negative. The product reviews are shown in Figure 6.

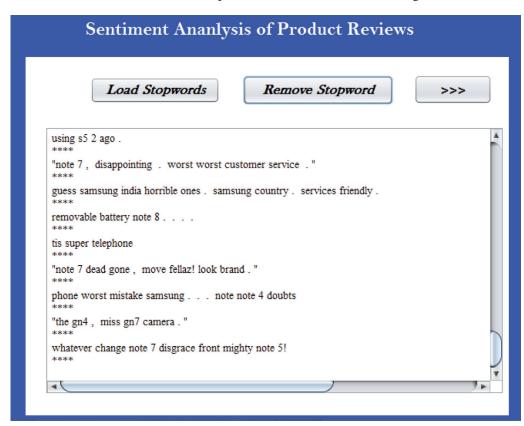


Figure 6. Product reviews.

After pre-processing the text has been transformed through Term Frequency (TF) that assigns a weight to each word based on how frequently a word occurs in the document. This is computed by using this expression:

$$TF(t) = \frac{\text{Number of times term (t) appears in a document}}{\text{Total number of terms in the document}}$$

After computing TF Inverse Document Frequency (IDF) measures how important a term is by weighing down frequent terms and scaling up the rare ones. IDF is calculated by using this expression:

IDF(t) =
$$\log_e e \frac{\text{Total number of documents}}{\text{Number of documents with term (t) in it}}$$
 10

This module refers the validating feature by using clustering is shown in Figure 7.

With the help of hyper plane in Support Vector Machine the data is divided into two classes such as Positive and Negative referred in Figure 8.

Now, Support Vector Machine classification data is measured into Negative shown in Figure 9.

In this section, the results are shown in Table 1 the reviews which is based on three datasets. Hence, the number of reviews in Table 1 is categorized into two classes one is of the positive reviews and second one negative review. Positive review shows the good results and the negative shows poor results.

Table 1. Product reviews

Product Name	No. of Reviews	Positive Reviews	Negative Reviews
HtcDesire.csv	250	156	94
htcOneMe.csv	180	102	78
HtcOne-X9.csv	190	125	65

Moreover, the Support Vector Machine (SVM) classification algorithm provides higher accuracy 89.98% than the other ones. The obtained accuracy would be enhanced further by including more sentence forms. Finally, it concludes that the SVM algorithm is better one.



Figure 7. Clustering process.

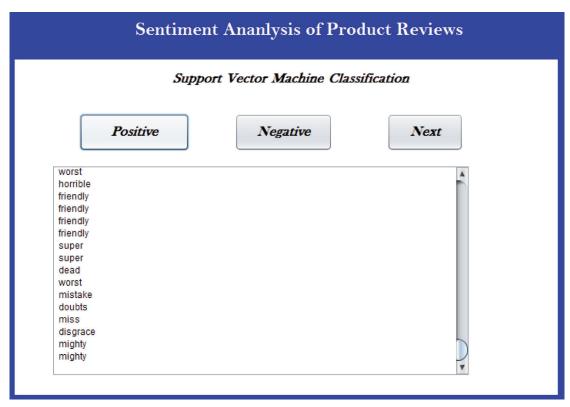


Figure 8. SVM Classification of positive words.

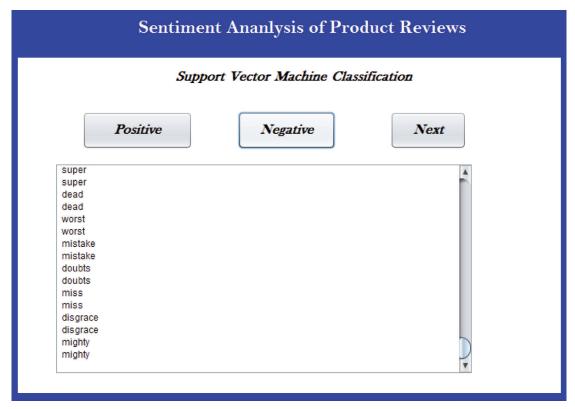


Figure 9. SVM classification of negative words.

6. Conclusion

This paper presents an experimental study along with a proposed model through SVM algorithm on different datasets of Product reviews to measure the polarity of reviews whether positive or negative and words related to the products such as good, bad, excellent, super hit. The performance resulting models are tested to measure accuracy of Support Vector Machine learning algorithm. Finally, the Support Vector Machine classification algorithm is achieved high accuracy and found better one than others.

7. References

- Tripathy A, Agrawal A, Rath SK. Classification of sentiment reviews using N-gram machine learning approach. Expert Systems with Applications International Journal. 2016 Mar; 17(1).
- Giatsogloua M, Vozalis MG, Diamantaras K, Vakali A, Sarigiannidis G, Ch. Chatzisavvas K. Sentiment analysis leveraging emotions and word embeddings. Expert Systems with Applications International Journal. 2016 Oct 18:214–24.
- Appel O, Chiclana F, Carter J, Fujita H. A hybrid approach to the sentiment analysis problem at the sentence level. Knowledge-Based Systems. 2016 May 19; 108:110–24. Crossref.
- 4. Korkontzelos I, Nikfarjam A, Shardlow M, Sarker A, Ananiadou S, Gonzalez GH. Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. Journal of Biomedical Informatics. 2016 Aug; 62:148–58.
- Severyn A, Moschitti A. Twitter sentiment analysis with deep convolutional neural networks. SIGIR '15 Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval; 2015. p. 959–62.
- Nakov P, Zesch T. Computational semantic analysis of language: SemEval-2014 and beyond. Language Resources and Evaluation. 2016 Mar 1; 50(10):1–4.

- Gokulakrishnan B, Priyanthan P, Ragavan T, Prasath N, Perera A. Opinion mining and sentiment analysis on a twitter data stream. 2012 International Conference on Advances in ICT for Emerging Regions (ICTer), 2012 Dec 12–15, Colombo, Sri Lanka; 2013.
- 8. Pak A, Paroubek P. Twitter as a corpus for sentiment analysis and opinion mining. Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010, 2010 May 17–23, Valletta, Malta; 2010.
- Jansen BJ, Zhang M, Sobel K, Chowdury A. Twitter power: Tweets as electronic word of mouth. Journal of the American Society for Information Science and Technology. 2009 Nov; 60(11):2169–88. Crossref.
- 10. Available from: https://discuss.analyticsvidhya.com.
- 11. Kohavi R, Provost F. Machine learning. Kluwer Academic Publishers, Boston. Manufactured in the Netherlands. 1998; 30(2/3):271–4.
- 12. Han J, et al., Classifier evaluation metrics: precision and recall, and F-measure.
- 13. Pang B, Lee L, Vaithyanathan S. Thumbs up? sentiment classification using machine learning techniques. Proceedings of the ACL-02 conference on Empirical methods in natural language Processing. 2002; 10:79–86. Crossref.
- 14. Mullen T, Collier N. Sentiment analysis using support vector machines with diverse information sources. Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, EMNLP 2004, A meeting of SIGDAT, a Special Interest Group of the ACL, held in conjunction with ACL 2004, 25-26 July 2004, Barcelona, Spain. 2004; 4:412–18.
- Matsumoto S, Takamura H, Okumura M. Sentiment classification using word sub-sequences and dependency sub-trees. Advances in Knowledge Discovery and Data Mining. 2005:301–11. Crossref.
- Luo F, Li C, Cao Z. Affective-feature-based Sentiment Analysis using SVM Classifier. 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in design; 2016. Crossref.
- 17. Mertiya M, Singh A. Combining Naïve Bayes and adjective analysis for sentiment detection on Twitter. IEEE; 2016.