International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 3 Issue 6, October 2019 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Sentiment Analysis on Twitter Dataset using R Language

B. Nagajothi[1], Dr. R. Jemima Priyadarsini[2]

[1]Research Scholar, [2]Associate Professor

[1,2]Department of Computer Science, Bishop Heber College, Tiruchirappalli, Tamil Nadu, India

ABSTRACT

Sentiment Analysis involves determining the evaluative nature of a piece of text. A product review can express a positive, negative, or neutral sentiment (or polarity). Automatically identifying sentiment expressed in text has a number of applications, including tracking sentiment towards Movie reviews and Automobile reviews improving customer relation models, detecting happiness and well-being, and improving automatic dialogue systems. The evaluative intensity for both positive and negative terms changes in a negated context, and the amount of change varies from term to term. To adequately capture the impact of negation on individual terms, here proposed to empirically estimate the sentiment scores of terms in negated context from movie review and auto mobile review, and built two lexicons, one for terms in negated contexts and one for terms in affirmative (non-negated) contexts. By using these Affirmative Context Lexicons and Negated Context Lexicons were able to significantly improve the performance of the overall sentiment analysis system on both tasks. This thesis have proposed a sentiment analysis system that detects the sentiment of corpus dataset using movie review and Automobile review as well as the sentiment of a term (a word or a phrase) within a message (term-level task) using R language.

KEYWORDS: *opinion mining*

How to cite this paper: B. Nagajothi | Dr. R. Jemima Priyadarsini "Sentiment Analysis on Twitter Dataset using R Language" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-6, October 2019, pp.199-204, URL: https://www.ijtsrd.com/papers/ijtsrd28071.pdf

IJTSRD28071

1. INTRODUCTION

Sentiment Analysis is the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. It represents a large problem space. There are also many names and slightly different tasks, e.g., Opinion mining, opinion mining, opinion extraction, sentiment mining, subjectivity analysis, affect analysis, emotion analysis, review mining, etc. However, they are now all under the umbrella of opinion mining. While in industry, the term opinion mining is more commonly used, but in academia both Opinion mining and opinion mining are frequently employed. Blogs, online forums, comment sections on media sites and social networking sites such as Facebook and twitter all can be considered as social media. These social media can capture millions of peoples' views or word of mouth. Communication and the availability of these real time opinions from people around the world make a revolution in computational linguistics and social network analysis. Social media is becoming an increasingly more important source of information for an enterprise. On the other hand people are more willing and happy to share the facts about their lives, knowledge, experiences and thoughts with the entire world through social media more than ever before. They actively participate in events by expressing their opinions and stating their comments that take place in society. This way of sharing their knowledge and emotions with society and social media drives the businesses to collect more information about their companies, products and to know how reputed they are among the people and thereby take decisions to go on with their businesses effectively. Therefore it is clear that sentiment analysis is a key component of leading innovative Customer Experience Management and Customer Relationship Marketing focused enterprises. Moreover for businesses looking to market their products, identify new opportunities and manage their reputation. As businesses look to automate the process of filtering out the noise, understanding the conversations, identifying the relevant content and take appropriate action upon it. Many are now looking to the field of sentiment analysis. In the era which to live today, sometimes known as information age, knowledge society; having access to large quantities of information is no longer an issue looking at the tons of new information produced everyday on the web. In this era, information has become the main trading object for many enterprises. If to can create and employ mechanisms to search and retrieve relevant data and information and mine them to transfer it to knowledge with accuracy and timeliness, that is where to get the exact usage of this large volume of information available to us.The widespread growth of social networks throughout the past decade has opened up entirely new possibilities for researchers when it comes to collecting large amounts of data. With millions of conversations taking place on social networks every day they offer a rich data source that can be accessed in a comparatively effortless way.

The micro blogging service Twitter allows developers to connect to its Streaming API to receive a real-time data stream containing Twitter posts ("tweets") and user information.

The brevity of the posts as well as their mostly text-based nature facilitate the analysis using data mining methods and have, therefore, made Twitter a popular data source for scientific research. As data mining and text mining techniques become more advanced, information can be retrieved on an increasingly fine-grained level. An area that has received considerable attention throughout the past years is Sentiment Analysis. The method has been widely applied to capture individuals' sentiment towards products or to assess the overall sentiment expressed in a piece of text. Besides the more general classification of comments or reviews as positive, neutral or negative, it also allows researchers to identify the type and intensity of more distinct emotions, such as fear, joy or surprise, in written text. However, being able to detect emotional content in social network data does not necessarily imply that useful knowledge can be derived from it. Despite its benefits in regards to wealth and accessibility, social network data is naturally unstructured and noisy. It is therefore not a trivial task to filter and collect large amounts of data relevant to a specific research question and to choose appropriate tools for further analysis in order to obtain meaningful results. In this work an approach for acquiring, analyzing and interpreting suitable data from social networks using Sentiment Analysis is developed. Social network data from a research project on gender and cultural diversity in global software engineering is then used to verify the approach by applying it to two example research questions.

1.1. Different Levels of Analysis

I now give a brief introduction to the main research problems based on the level of granularities of the existing research. In general, Opinion mining has been investigated mainly at three levels:

A. Document level

The task at this level is to classify whether a whole opinion document expresses a positive or negative sentiment. For example, given a product review, the system determines whether the review expresses an overall positive or negative opinion about the product. This task is commonly known as document-level sentiment classification.

B. Sentence level

The task at this level goes to the sentences and determines whether each sentence expressed a positive, negative, or neutral opinion. Neutral usually means no opinion. This level of analysis is closely related to subjectivity classification, which distinguishes sentences (called objective sentences) that express factual information from sentences (called subjective sentences) that express subjective views and opinions

C. Entity and Aspect level

Both the document level and the sentence level analyses do not discover what exactly people liked and did not like. Aspect level performs finer-grained analysis. Aspect level was earlier called feature level (feature-based opinion mining and summarization). Instead of looking at language constructs (documents, paragraphs, sentences, clauses or phrases), aspect level directly looks at the opinion itself. It is based on the idea that an opinion consists of a sentiment (positive or negative) and a target (of opinion).

1.2. Opinion Lexicon and Its Issues

Not surprisingly, the most important indicators of sentiments are sentiment words, also called opinion words. These are words that are commonly used to express positive or negative sentiments. For example, good, wonderful, and amazing are positive sentiment words, and bad, poor, and terrible are negative sentiment words. Apart from individual words, there are also phrases and idioms, e.g., cost someone an arm and a leg. Sentiment words and phrases are instrumental to Opinion mining for obvious reasons. A list of such words and phrases is called a sentiment lexicon (or opinion lexicon). Over the years, researchers have designed numerous algorithms to compile such lexicons.

Although sentiment words and phrases are important for Opinion mining, only using them is far from sufficient. The problem is much more complex. In other words, to can say that sentiment lexicon is necessary but not sufficient for Opinion mining. Below, to highlight several issues:

1. A positive or negative sentiment word may have opposite orientations in different application domains. For example, "suck" usually indicates negative sentiment, e.g., "This camera sucks," but it can also imply positive sentiment, e.g., "This vacuum cleaner really sucks."
2. A sentence containing sentiment words may not express any sentiment. This phenomenon happens frequently in several types of sentences. Question (interrogative) sentences and conditional sentences are two important types, e.g., "Can you tell me which Sony camera is good?" and "If I can find a good camera in the shop, I will buy it." Both these sentences contain the sentiment word "good", but neither expresses a positive or negative opinion on any specific camera. However, not all conditional sentences or interrogative sentences express no sentiments, e.g., "Does anyone know how to repair this terrible printer" and "If youare looking for a good car, get Toyota Camry."

1.3. Different Types of Opinions

The type of opinions that to have discussed so far is called regular opinion. Another type is called comparative opinion. In fact, to can also classify opinions based on how they are expressed in text, explicit opinion and implicit (or implied) opinion.

A. Regular and Comparative Opinions

Regular opinion: A regular opinion is often referred to simply as anopinion in the literature and it has two main sub-types:

Direct opinion: A direct opinion refers to an opinion expressed directly on an entity or an entity aspect, e.g., "The picture quality is great."

Indirect opinion: An indirect opinion is an opinion that is expressed indirectly on an entity or aspect of an entity based on its effects on some other entities. This sub-type often occurs in the medical domain. For example, the sentence "After injection of the drug, my joints felt worse" describes an undesirable effect of the drug on "my joints", which indirectly gives a negative opinion to the drug.

B. Explicit and Implicit Opinions

I. Explicit opinion

An explicit opinion is a subjective statement that gives a regular or comparative opinion, e.g., "Coke tastes great," and "Coke tastes better than Pepsi."

II. Implicit (or implied) opinion

An implicit opinion is an objective statement that implies a regular or comparative opinion. Such an objective statement usually expresses a desirable or undesirable fact, e.g., "I bought the mattress a week ago, and a valley has formed," and "The battery life of Nokia phones is longer than Samsung phones." Explicit opinions are easier to detect and to classify than implicit opinions. Much of the current research has focused on explicit opinions.

1.4. Scope of the thesis

In this work aim to use Opinion mining on a set of Twitter movie review and Twitter Auto mobile reviews given by reviewers and try to understand what their overall reaction to the movie was, i.e. if they liked the movie or they hated it. This aim to utilize the relationships of the words in the review to predict the overall polarity of the review.

2. RELATED WORK

Lisa Branz [1] an approach towards the retrieval, analysis, and interpretation of social network data for research purposes is developed. The data is filtered according to pertinent criteria and analyzed using Sentiment Analysis tools modified specially to the data source. The approach is established by applying it to two example research questions, confirming history findings on cultural and gender differences in sentiment appearance. Hypothesis 1: The quantity of positive sentiment in tweets about sports differs considerably between male and female software engineers; with male's presentation more positive sentiment towards sports associated topics than females. Hypothesis 2: The amount of sentiment expressed in tweets differs considerably among software engineers from collective cultures and software engineers from free spirit cultures, with users from collective cultures expressing less sentiment.

Prabhsimran Singh, Ravindra Singh, and Karanjeeet Singh Kalhon, [2] they have examined this government policy the demonetization from the ordinary person's viewpoint with use of the approach of sentiment analysis and using Twitters data, Tweets are collected using the certain hashtag (#demonetization). Analysis based on geolocation (State wise tweets are collected). The sentiment analysis API used from meaning cloud and classified the states into six categories, they are happy, sad, very sad, very happy, neutral, and no data.

Vamshi Krishna [3] discusses a new topic model-based approach for opinion mining and sentiment analysis of text reviews posted in web forums or social media site which are mostly in unstructured in nature. In recent years, opinions are exchanged in clouds about any product, person, event or an interesting topic. These opinions help in decision making for choosing a product or getting feedback about any topic. Opinion mining and sentiment analysis are related in a sense that opinion mining deals with analyzing and summarizing expressed opinions whereas sentiment analysis classifies opinionated text into positive and negative. Aspect extraction is a crucial problem in sentiment analysis. The model proposed in the paper utilizes a topic model for aspect extraction and support vector machine learning technique for sentiment classification of textual reviews. The objective is to mechanize the process of mining attitudes, opinions and hidden emotions from the text.

Xing Fang, Justin Zhan, [4] they have solved the issue of sentiment polarity categorization, and it is one of the basic problems of sentiment analysis. Online product reviews data is used in this study, collected from Amazon.com. In this paper, Investigation for both sentence-level categorization and review-level categorization are achieved. Scikit-learn software is used for this study. Scikit-learn is an open source machine learning software package in Python. Naïve Bayesian, Random Forest, and SVM: These classification techniques selected for categorization.

Geetika Gautam, Divakar Yadav, [5] they contribute to the sentiment analysis for customers' review classification. Already labeled twitters data is used in this task. They have used three supervised techniques in this paper: naïve-Bayes, Max-entropy, and SVM followed by the semantic analysis which was used along with all three methods to calculate the similarity. They have used Python and NLTK to train and classify the: naïve-Bayes, Max-entropy, and SVM. Naïve-Byes approach gives a better result than the Max-entropy and SVM with unigram model gives a better result than using SVM alone. Then the correctness is then increased when the Word-Net of semantic analysis is applied after the above procedure.

Neethu M S, Rajasree R, [6] in this paper, they analyze the twitter data related to Electronic products using Machine Learning approach. They existent a new Feature-Vector for classification of the tweets and extricate peoples' opinion about Electronic products. Thus Feature-Vector is created from 8 relevant features. The 8 features used are a special keyword, presence of negation, pos tag, and number of positive keywords, emoticon, and number of negative keywords, number of negative hashtags and number of positive hashtags. Naïve-Bayes and SVM classifiers are implemented using built-in functions of Matlab. Max-Entropy classifier is implemented using Maximum-Entropy software. All the used classifiers have almost equal performance.

Blitzer et al. [7] proposed an approach called structural correspondence learning for domain adaptation where it used pivot features to bridge the gap between source and target domain. Automatic sentiment classification has been extensively studied and applied in recent years. However, sentiment is expressed differently in different domains, and annotating corpora for every possible domain of interest is impractical. They investigate domain adaptation for sentiment classifiers, focusing on online reviews for different types of products. First, they extend to sentiment classification the recently-proposed structural correspondence learning (SCL) algorithm, reducing the relative error due to adaptation between domains by an average of 30% over the original SCL algorithm and 46% over a supervised baseline. Second, they identify a measure of domain similarity that correlates well with the potential for adaptation of a classifier from one domain to another. This measure could for instance be used to select a small set of domains to annotate whose trained classifiers would transfer well to many other domains

3. METHODOLOGY

3.1. Introduction

The tweets posted by users in software engineering jobs were then classified for tweet topic. For the hypotheses proposed, a precise and faceted classification of tweet topics was not necessary.It was considered sufficient to classify a tweet as Movie-related and automobile related.

3.2. CORPUS

In linguistics, a corpus or text corpus is a large and structured set of texts (nowadays usually electronically stored and processed). They are used to do statistical analysis and hypothesis testing, checking occurrences or validating linguistic rules within a specific language territory. A corpus may contain texts in a single language (monolingual corpus) or text data in multiple languages (multilingual corpus).Multilingual corpora that have been specially formatted for side-by-side comparison are called aligned parallel corpora. There are two main types of parallel corpora which contain texts in two languages. In a translation corpus, the texts in one language are translations of texts in the other language.

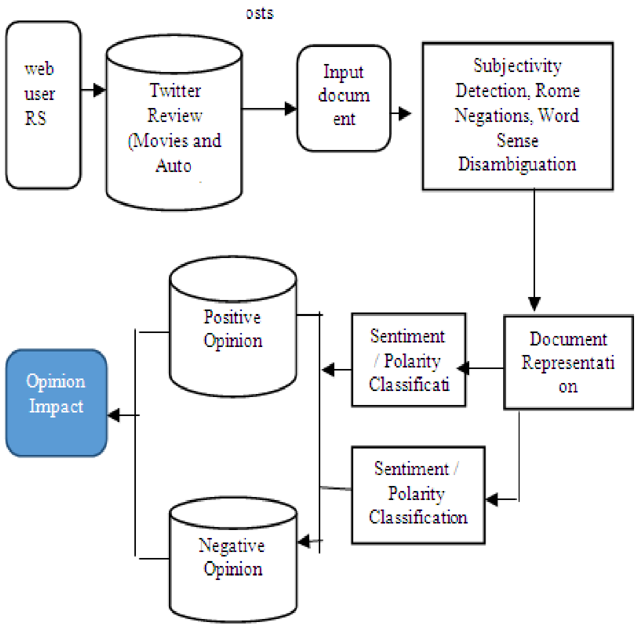

Fig 3.2 Proposed Methodology

Fig. 3.2 shows the main important steps in instruction to achieve a view impact. The web users post their views, comments and feedback about a particular product or a thing through some blogs, forums and social networking sites. Data is composed from such opinion sources in such a technique that only the appraisals related to the topic, that is inspected is selected. The input text is then preprocessed. Preprocessing, in this setting, is the removal of the fact based sentences, thus choosing only the opinionated sentences. Further refinements are made by removing the negations and by sensing the word disambiguation. Then, the process of extracting relevant features is done. Feature selection can potentially improve classification accuracy, narrow in on a key feature subset of sentiment discriminators, and provide greater insight into important class attributes. The extracted features contribute to a document vector upon which various machine learning techniques can be applied in order to classify the polarity (positive and negative opinions) using the obtained document vector and finally the opinion impact is obtained based on the sentiment of the web users.

3.3. R Language

R is a language and environment for statistical computing and graphics. It is a GNU project which is similar to the S language and environment which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. R can be considered as a different implementation of S. There are some important differences, but much code written for S runs unaltered under R.R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, …) and graphical techniques, and is highly extensible. The S language is often the vehicle of choice for research in statistical methodology, and R provides an Open Source route to participation in that activity.R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.

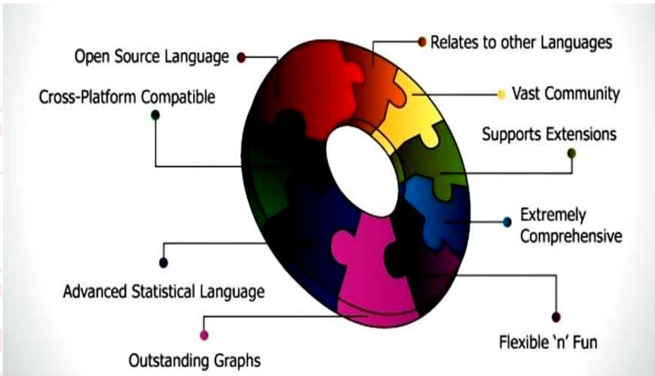

Fig 3.3 R language

In fig 3.3 defines the various element of the R language where the features of the language is defined. The features like Open Source language, Related to other language, Cross platform compatible, Comprehensive, Advanced statistical language and outstanding Graphs are defined.

3.4. Pre-processing

Data pre-processing is an often neglected but important step in the data mining process. The phrase "Garbage In. Garbage Out" is particularly applicable to data mining and machine learning Data pre-processing includes cleaning, normalization, and transformation. Feature extraction and selection. Etc. The product of data pre-processing is the final training set.

1. Data Cleaning:

The data can have many irrelevant and missing parts. To handle this part, data cleaning is done. It involves handling of missing data, noisy data etc.

A. Missing Data:

This situation arises when some data is missing in the data. It can be handled in various ways. Some of them are:

1. Ignore the tuple:

This approach is suitable only when the dataset we have is quite large and multiple values are missing within a tuple.

2. Fill the Missing values:

There are various ways to do this task. You can choose to fill the missing values manually, by attribute mean or the most probable value.

B. Noisy Data:
Noisy data is a meaningless data that can't be interpreted by machines. It can be generated due to faulty data collection, data entry errors etc. It can be handled in following ways :

3. Binning Method:
This method works on sorted data in order to smooth it. The whole data is divided into segments of equal size and then various methods are performed to complete the task. Each segmented is handled separately. One can replace all data in a segment by its mean or boundary values can be used to complete the task.

4. Regression:
Here data can be made smooth by fitting it to a regression function. The regression used may be linear (having one independent variable) or multiple (having multiple independent variables).

5. Clustering:
This approach groups the similar data in a cluster. The outliers may be undetected or it will fall outside the clusters.

C. Data transformation
In data transformation, the data are transformed consolidated into forms appropriate for mining. Data transformation can involve the following:
1. Normalization, where the attribute data are scaled so as to fall within a small specified range, such as -1.0 to 1.0 or 0 to 1.0.
2. Smoothing works to remove the noise from data. Such techniques include binning, Clustering, and regression.
3. Aggregation, where summary or aggregation operations are applied to the data. For example, the daily sales data may be aggregated so as to compute monthly and annual total amounts. This step is typically used in constructing a data cube for analysis of the data at multiple granularities.
4. Generalization of the dam, where low level or 'primitive' (raw) data are replaced by higher level concepts through the use of concept hierarchies. For example, Categorical attributes. like street, can be generalized to higher level concepts, like city or county.

Similarly, values for numeric attribute, like age, may be mapped to higher level concepts, like young. middle-aged, and senior.

D. Data Reduction
Complex data analysis and mining on huge amounts of data may take a very long time, making such analysis impractical or infeasible. Data reduction techniques have been helpful in analyzing reduced representation of the dataset without compromising the integrity of the original data and yet producing the quality knowledge. Feature selection (FS) extraction (FE) and construction (FC) can be used in combination. In many cases feature construction expands the number of features with newly constructed ones that are more expressive but they may include useless features. Feature selection can help automatically reduce those excessive features.

4. Twitter Movie Reviews
The twitter dataset contains 50,000 training examples collected from IMDb. Where each review is labeled with the rating of the movie on scale of 1-10. As sentiments are usually bipolar like good/bad or happy/sad or like/dislike, the categorized these ratings as either 1 (positive) or 0 (negative) based on the ratings. If the rating was above 5, here deduced that the person liked the movie otherwise he did not. Initially the dataset was divided into two subsets containing 25,000 examples each for training and testing. Here found this division to be sub-optimal as the number of training examples was very small and leading to under-fitting. Here then tried to redistribute the examples as 40,000 for training and 10,000 for testing. While this produced better models, it also led to over-fitting on training examples and worse performance on the test set. This improved the accuracy of our models across the boards. A typical review text looks like this:

```
# Basic Opinion mining Model
library(tm) ; library(RWeka) ;
# Read the file containing Postive and Negative terms
positive_terms = read.csv("PositiveMovie.csv")
positive_terms = as.character(positive_terms$Positive)
negative_terms = read.csv("NegativeMovie.csv")
negative_terms = as.character(negative_terms$Negative)
```

The sentiment analysis model found 14 positive words and 4 negative words, and the final sentiment score was 10. This tells us that the quarterly result for data set was good from the management's perspective. The word cloud below shows some of the positive/negative words that were picked from the text document on which to run the model.

Table4.1 Comparison table for Twitter Movie Review and Twitter Automobile Review

	Twitter Movie review	Twitter Automobile Review
Positive	14	8
Negative	4	4

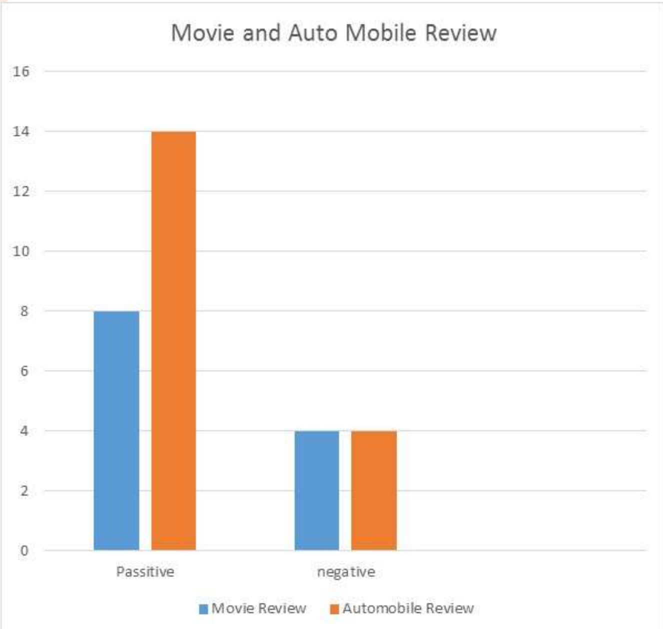

Fig 4.11 A graph for Comparison for Movie Review and Automobile Review

5. CONCLUSION

This thesis have proposed a Opinion mining system that detects the sentiment of corpus dataset using twitter movie review and twitter auto mobile review as well as the opinion of a term (a word or a phrase) within a message (term-level task).

From the results of the experimentation, it has been observed that when positive terms are negated, they tend to convey a negative opinion. In contrast, when negative terms are negated, they tend to still convey a negative opinion. Furthermore, the evaluative intensity for both positive and negative terms changes in a negated context, and the amount of change varies from term to term. To adequately capture the impact of negation on individual terms, here proposed to empirically estimate the sentiment scores of terms in negated context from movie review and auto mobile review, and built two lexicons, one for terms in negated contexts and one for terms in affirmative (non-negated) contexts. By using these Affirmiative Context Lexicons and Negated Context Lexicons to were able to significantly improve the performance of the overall Opinion mining system on both tasks. In particular, the features derived from these lexicons provided gains of up to 6.5 percentage points over the other feature groups.

5.1. Future work

In future the work could be extended by using an hybrid approach using both corpus based method and dictionary based method to determine the semantic orientation of words in tweets. Hybridization may improve the quality and enhance the output.

REFERENCES

[1] Lisa Branz and Patricia Brockmann. “Poster: Sentiment Analysis of Twitter Data: Towards Filtering, Analyzing and Interpreting Social Network Data”. In DEBS ’18: The 12th ACM International Conference on Distributed and Event-based Systems, June 25–29, 2018

[2] Singh, Prabhsimran, Ravinder Singh Sawhney, and Karanjeet Singh Kahlon. "Sentiment analysis of demonetization of 500 & 1000 rupee banknotes by the Indian government." ICT Express 4, no. 3 (2018): 124-129.

[3] Vamshi, Krishna B., Ajeet Kumar Pandey, and Kumar AP Siva. "Topic Model-Based Opinion Mining and Sentiment Analysis." In 2018 International Conference on Computer Communication and Informatics (ICCCI), pp. 1-4. IEEE, 2018.

[4] Fang, Xing, and Justin Zhan. "Sentiment analysis using product review data." Journal of Big Data 2, no. 1 (2015).

[5] Gautam, Geetika, and Divakar Yadav. "Sentiment analysis of Twitter data using machine learning approaches and semantic analysis." In 2014 Seventh International Conference on Contemporary Computing (IC3), pp. 437-442. IEEE, 2014.

[6] Neethu, M. S., and R. Rajasree. "Sentiment analysis in twitter using machine learning techniques." In 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), pp. 1-5. IEEE, 2013.

[7] J.Blitzer, M. Dredze, F. Pereira et al., Biographies, Bollywood, Boom-boxes and Blenders: Domain Adaptation for Sentiment Classification," in Proceedings of the 45th Annual Meeting of the Association of the Computational Linguistics, ACL, vol. 7, pp. 440 - 447, 2007.

[8] S. J. Pan, X. Ni, J.-T. Sun, Q. Yang, and Z. Chen, Cross-domain Sentiment Classification via Spectral Feature Alignment," in Proceedings of the 19th International Conference on World wide web. ACM, pp. 751-760, 2010.

[9] D. Bollegala, D. Weir, and J. Carroll, Using Multiple Sources to Construct a Sentiment Sensitive The saurus for Cross-domain Sentiment Classification," in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, vol. 1, pp. 132-141, 2011.

[10] Y. Choi and C. Cardie, Adapting a Polarity Lexicon using Integer Linear Programming for Domain Specific Sentiment Classification," in Proceedings of the Conference on Empirical Methods in Natural Language Processing, ACL, vol. 2, pp. 590 - 598, 2009.

[11] E. Agirre and D. Martinez,Exploring Automatic Word Sense Disambiguation with Decision Lists and the web,” in proceedings of the COLING-2000 work shop on semantic Annotation and Intelligent Content. Association for Computational Linguistic,pp.11-19,2000

[12] S. Fujita and A. Fujino, Word Sense Disambiguation by Combining Labeled Data. Expansion and Semi-supervised Learning Method," ACM Transactions on Asian Language Information Processing (TALIP), vol. 12, no. 2, p. 7, 2013

[13] R. Xia, F. Xu, C. Zong, Q. Li, Y. Qi, and T. Li, Dual Sentiment Analysis: Considering Two Sides of One Review," IEEE Transactions on Knowledge and Data Engineering, vol. .27, no. 8, pp. 2120-2133, 2015.

[14] Gupta, B., Negi, M., Vishwakarma, K., Rawat, G. and Badhani, P., 2017. Study of Twitter sentiment analysis using machine learning algorithms on Python. International Journal of Computer Applications, 165(9), pp.0975-8887.

[15] Anupkant, S., PVM Seravana Kumar, Nayani Sateesh, and D. Bhanu Mahesh. "Opinion mining on author's citation characteristics of scientific publications." In 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), pp. 348-351. IEEE, 2017.