Customer Sentiment Classification and Rating System based on Products Reviews

Monika[1] Er. Sandeep Garg[2]

[1]Research Scholar [2]Assistant Professor

[1,2]RPIIT Bastara, Karnal, India

Abstract— In the present competitive business scenario vast amount of consumer reviews are written on Web about any product or service. WWW contains an overwhelming volume of customer reviews about different categories of commodities avail. The advent of social media and ecommerce has brought the era of a new age business and its customer base is growing exponentially every year. In today's world, the online market is increasingly getting popular and it becomes more and more important to help the customer get the best product by all parameters. The quality of a product is best confirmed by taking the customer reviews from those who are already using that. All popular shopping websites like Amazon, flipkart, ebay etc. allow customer reviews once the product has been purchased. These reviews are such huge in numbers on these websites that it is not possible for a customer to consider them all. This paper focuses on extracting the features from product reviews taken from amazon.com or ebay.in sites given by reviewers to state their opinions. This is done at aspect level of analysis using ontology. Then it determines whether they are positive or negative thereby giving a scaling system to identify the effectiveness of a product. So, we propose a technique here, Sentiment Analysis for Product rating through Ontology (SAPRO). This technique performs text mining techniques on the customer reviews obtained from any of these websites and calculates the rating for this product (out of 5), to establish the like-ability of the product by the existing customers. SAPRO considers the end users' perspective while addressing the Sentiment Analysis problem. The research uses a combination approach of domain ontology and Stanford dependency relation which intends to enhance the sentiment classification. The proposed technique is programmed in Java 8 programming language taking review data from the Amazon website as the dataset which is in JSON format. The results show that the review rating calculated by the proposed technique has 91% similarity with the existing ratings from the customers. The difference is mainly because the customer gives discrete numeric ratings but explains his sentiment better in words. The results when compared to base research by Arindam Chaudharyet. al [24], clearly show that the proposed technique(91% accuracy) is better than base technique(80-88% accuracy) for reviews.

Key words: Customer Sentiment Classification, Customer Sentiment Rating System

I. INTRODUCTION

Ever increasing use of Internet and online activities (like chatting, conferencing, surveillances, ticket booking, online transactions, e-commerce, social media communications, blogging and micro-blogging, clicks streams, etc.) leads us to extract, transform, load, and analyze very huge amount of structured and unstructured data, at a fast pace, referred to as Big Data. Such data can be analyzed using a combination of Data Mining, Web Mining and Text Mining techniques in various real life applications. Huge amount of information related to customer opinions/reviews is quite cumbersome to analyze and needs extant approaches to get a generalized opinion summary.

Opinion Mining or Sentiment Analysis is the study that analyzes people's opinions or sentiments from the text towards entities such as products and services. It has always been important to know what other people think. With the rapid growth of availability and popularity of online review sites, blogs', forums', and social networking sites' necessity of analysing and understanding these reviews has arisen. The main approaches for sentiment analysis can be categorized into semantic orientation-based approaches, knowledge-based, and machine-learning algorithms.

II. TEXT MINING TECHNIQUES

A. Feature Extraction

While trying to extract the particular feature from a review or comment, the first thing is how to represent the review[6].One way to do that is the conventional VSM(Vector space model). The VSM does the job in a quantitative way by representing the documents with vectors. The review text or comment is first tokenized and then added to a vector. The vector is filtered for the stop words. Then, each word in the vector is converted to its corresponding word stem. The effectiveness of these words is evaluated by using quantitative methods and their statistical information including the frequency of each word stem, or another similar measure, in the comment as the corresponding element in the vector used to represent that text. The following steps are performed for text classification.

1) Parsing the documents and case-folding

This step removes all abbreviations and non-alpha characters from the comment. Case-folding means converting all the characters in a text into the same case. In this research, we use Stanford parser to parse our sentences to disintegrate them into different parts of grammar. Now we can exclude the unwanted words

2) Removing stopwords

There are words in English which are used to provide structure to the language like conjunctions, articles, pronouns and prepositions. Such words which occur very frequently and carry no useful information about the content are called stopwords. So, remove such words from the review comment.

3) Stemming

Stemming is the process for reducing derived words to their stem or root. Porter stemmer utilizes suffix stripping. Porter's stemming algorithm steps:

- Remove the plurals and suffixes like –ed or –ing from each token.
- Convert each existence of y to i when another vowel in stem.

– The double suffixes are mapped to single ones:-ization, -ational etc.
– The suffixes like -full, -ness etc. are either removed or properly dealt with.
– Remove the suffixes like –ant, -ence, etc.
– Gets rid of a final –e.

4) Term Weighting

In this step, weight is assigned to a word based on number of times it occurs in the comment. This method is called term frequency and inverse term frequency which is a traditional method to assign a weight to the words.

The explicit feature can be extracted as:
– Depending on frequency of nouns and noun phrases
– Based on the relations between Opinion and Target.

5) Sentiment Mining

Today, numerous customers and users share their experiences using various social media sites such as Twitter, Facebook and blogs. It has become a challenge for organizations to monitor and understand what people post on social media sites. Traditional content analysis methods are no longer able to meet organizations' needs to analyze the large amount of new content on a daily basis. Applying automatic methods to quickly analyze such content is increasingly needed by organizations. As users continue to post textual information on various social media sites, there is a growing interest in using text mining, sentiment analysis and social network analysis approaches to process large amounts of user-generated data and extract meaningful knowledge and insights. As an emerging technology, text mining aims to extract meaningful information from unstructured textual data. To glean useful information from a large number of textual documents quickly, it has become imperative to use automated computer techniques. Text mining is focused on finding useful models, trends, patterns, or rules from unstructured textual data. Different from traditional content analysis, the main purpose of text mining is to automatically extract knowledge, insights, useful patterns or trends from a given set of textual documents. Text mining techniques have been used to analyze large amounts of textual data. Morinaga et al. present a framework for mining public opinions related to product reputation on the Internet. They find that text mining techniques offer both a dramatically reduced cost and increased knowledge discovery from public opinion, compared with the conventional survey approach.

6) Stanford Parser

"I am doing my work and no one should uselessly disturb"
– I/PRP
– am/VBP
– doing/VBG
– my/PRP$
– work/NN
– and/CC
– no/DT
– one/NN
– should/MD
– uselessly/RB
– disturb/VB

The above example shows the working of Stanford parser. It parses the sentence and returns the grammar type of each word, This helps us to understand the NOUN words which can possibly form the table names. We generate the list of noun words and also add to it the synonyms of these words generated by WORDNET.

7) WordNet

WordNet is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. The resulting network of meaningfully related words and concepts can be navigated with the browser. WordNet's structure makes it a useful tool for computational linguistics and natural language processing.

III. BASE RESEARCH

Chaudharyet. al in [24] proposed and developed a hierarchical bidirectional recurrent neural network (HBRNN) in order to characterize sentiment specific aspects in review data available at DBS Text Mining Challenge. HBRNN is developed by extending RNN and BRNN so that accuracy and efficiency are improved. HBRNN predicted aspect sentiments vector at review level. HBRNN is optimized by fine tuning different network parameters and compared with methods like long short term memory (LSTM) and bidirectional LSTM (BLSTM). The methods are evaluated with highly skewed data. All models are evaluated using precision, recall and F1 scores. The results on experimental dataset indicate superiority of HBRNN over other techniques.

IV. PROPOSED METHODOLOGY

The customer reviews of different products for any particular enterprise are considered to extract entity level sentiments. Data acquisition and data pre-processing are most common sub-tasks required for text mining and SA. The product reviews are analyzed by (a) extracting most important features of product and (b) assigning an overall score for each of them. This allows us to structure information from reviews by summarizing them in a comprehensive and concise form. The problem can thus be formulated as: Given a review as form of sentence S_i, the sentiment scores $SS_{a,i}$ of relevant features or aspects a are to be identified.

A. Algorithm

– Step1: Input all reviews in form of JSON data
– Step2: Convert all reviews into vector of sentences where each element in vector is a review.
– Step3: Loop for i=1 to length (vector)
– Step4: Let s[i] be each review
– Step5: Split s[i] into sentences and final sentiment of each sentence by value= evaluate (sentence)
– Step6: Final total Sentiment of all sentences
– Step7: Scale value to order of 5
– Step 8: Scale value like very good, good, average, bad and worst.
– Step 9: Find mean of evaluated value summary evaluation and Amazon rating. This is exact rating for this review [end loop].

1) S -> sentence

Evaluate(S)

Initialize value = 0
– Step 1: Convert to lowercase
– Step 2: Filter unwanted words
– Step 3: Perform stemming of each word

− Step 4: Final sentiment carrying words and separate them to compare lists like positive words, negative words, domain specific positive and domain specific negative words.

− Step 5: Each match adds some numeric to final value

2) Match (Word1, Word2)

− Step 1: Return true if word1 = word2

− Step 2: x = soundex (word1) y = soundex (word2)

− Step 3: Return true if x=y

− Step 4: Match synonyms of these words by using WordNet and return true if successful match.

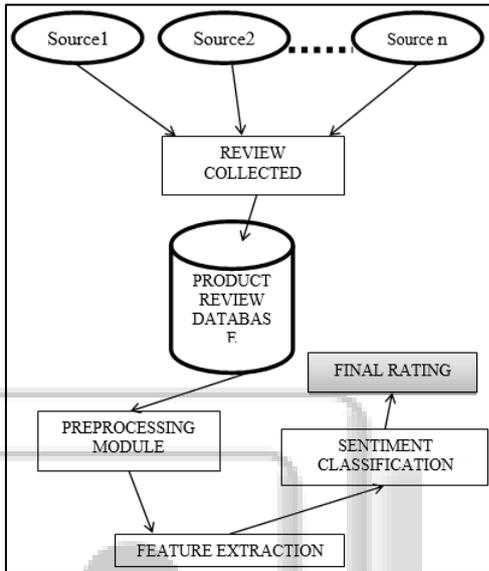The process flow diagram of our work is shown in the figure below.

Fig. 1: Demonstration of SAPRO process

B. Dataset Used

1) Sentiment Analysis for Product rating through Ontology. (SAPRO)

SAPRO is aimed at classification from multiple reviews on various products from any online source (amazon.com from 28-Feb-2014 to 15-Jul-2014 in this case). The data received from Amazon is in JSON format and needs to be parsed by our java application for implementing the proposed methodology. The JSON review format is as below:

2) Data Set format

Field Name	Datatype
ReviewerId	String
Asin	Numeric
ReviewerName	String
Helpful	Array
ReviewText	String
Overall	Double
Summary	String
UnixReviewTime	long number
ReviewTime	Timestamp

Table 1: Data Set format

The three most useful fields for our implementation are reviewText, Overall and summary. A part of the json data available to us from AMAZON.COM is listed below:

{"reviewerID": "A2IBPI20UZIR0U", "asin": "1384719342", "reviewerName": "cassandratu \"Yeah, well, that's just like, u...", "helpful": [0, 0], "reviewText": "Not much to write about here, but it does exactly what it's supposed to. filters out the pop sounds. now my recordings are much more crisp. it is one of the lowest prices pop filters on amazon so might as well buy it, they honestly work the same despite their pricing,", "overall": 5.0, "summary": "good", "unixReviewTime": 1393545600, "reviewTime": "02 28, 2014"}

{"reviewerID": "A14VAT5EAX3D9S", "asin": "1384719342", "reviewerName": "Jake", "helpful": [13, 14], "reviewText": "The product does exactly as it should and is quite affordable. I did not realized it was double screened until it arrived, so it was even better than I had expected. As an added bonus, one of the screens carries a small hint of the smell of an old grape candy I used to buy, so for reminiscent's sake, I cannot stop putting the pop filter next to my nose and smelling it after recording. :D If you needed a pop filter, this will work just as well as the expensive ones, and it may even come with a pleasing aroma like mine did! Buy this product! :]", "overall": 5.0, "summary": "Jake", "unixReviewTime": 1363392000, "reviewTime": "03 16, 2013"}

{"reviewerID": "A195EZSQDW3E21", "asin": "1384719342", "reviewerName": "Rick Bennette \"Rick Bennette\"", "helpful": [1, 1], "reviewText": "The primary job of this device is to block the breath that would otherwise produce a popping sound, while allowing your voice to pass through with no noticeable reduction of volume or high frequencies. The double cloth filter blocks the pops and lets the voice through with no coloration. The metal clamp mount attaches to the mike stand secure enough to keep it attached. The goose neck needs a little coaxing to stay where you put it.", "overall": 5.0, "summary": "It Does The Job Well", "unixReviewTime": 1377648000, "reviewTime": "08 28, 2013"}

{"reviewerID": "A2C00NNG1ZQQG2", "asin": "1384719342", "reviewerName": "RustyBill \"Sunday Rocker\"", "helpful": [0, 0], "reviewText": "Nice windscreen protects my MXL mic and prevents pops. Only thing is that the gooseneck is only marginally able to hold the screen in position and requires careful positioning of the clamp to avoid sagging.", "overall": 5.0, "summary": "GOOD WINDSCREEN FOR THE MONEY", "unixReviewTime": 1392336000, "reviewTime": "02 14, 2014"}

{"reviewerID": "A94QU4C90B1AX", "asin": "1384719342", "reviewerName": "SEAN MASLANKA", "helpful": [0, 0], "reviewText": "This pop filter is great. It looks and performs like a studio filter. If you're recording vocals this will eliminate the pops that gets recorded when you sing.", "overall": 5.0, "summary": "No more pops when I record my vocals.", "unixReviewTime": 1392940800, "reviewTime": "02 21, 2014"}

{"reviewerID": "A2A039TZMZHH9Y", "asin": "B00004Y2UT", "reviewerName": "Bill Lewey \"blewey\"", "helpful": [0, 0], "reviewText": "So good that I bought another one. Love the heavy cord and gold connectors. Bass sounds great. I just learned last night how to coil them up. I guess I should read instructions more carefully. But no harm done, still works great!", "overall": 5.0, "summary": "The Best Cable", "unixReviewTime": 1356048000, "reviewTime": "12 21, 2012"}

V. RESULTS

The proposed methodology was implemented in java programming language and taking JSON data from

Amazon.com as dataset for the customer reviews on the musical instruments sold via Amazon. The methodology has been implemented in such a way that it can be worked with any dataset no matter what format it is specified in. The only change needed is a change in connection details. Also, we can implement dataset from any other source and from any other domain, but a change in domain requires updating of domain specific positive and negative words.

```
Enter the json file name containing all reviews data ...
data2.json
The value of this review is 5.0
[good] = 3.0
NET value of this review is 4.3333333333333333
The value of this review is 5.0
[jake] = 3.0
NET value of this review is 4.3333333333333333
The value of this review is 5.0
[doe, job, well] = 5.0
NET value of this review is 5.0
The value of this review is 5.0
[good, windscreen, for, monei] = 5.0
NET value of this review is 5.0
The value of this review is 5.0
[no, more, pop, i, record, my, vocals.] = 5.0
NET value of this review is 5.0
The value of this review is 5.0
[best, cabl] = 3.0
NET value of this review is 4.3333333333333333
```

Fig. 2: Output from implementation for current Amazon dataset for musical instruments

Table 2: Results for calculated review rating from proposed methodology and User defined value.

Calculated Review Value	User Defined Value	Calculated Review Value	User Defined Value
4.33	5.00		
4.33	5.00		
5.00	5.00		
5.00	5.00		
5.00	5.00		
4.33	5.00		
3.33	5.00		
4.33	5.00		
3.78	5.00		
4.67	4.00		
4.33	3.00		
4.33	5.00		
3.33	5.00	5.00	5.00
3.79	2.00	5.00	5.00
4.00	4.00	5.00	5.00
3.84	5.00	4.00	4.00
5.00	5.00	4.33	5.00
4.11	4.00	4.00	4.00
5.00	5.00	4.85	5.00
3.00	4.00	5.00	5.00
3.33	5.00	2.33	2.00
5.00	5.00	4.33	3.00
1.33	4.00	5.00	5.00
3.78	5.00	4.33	5.00
2.50	5.00	5.00	5.00
4.67	4.00	4.67	4.00
2.22	5.00	4.33	5.00
4.33	5.00	4.33	5.00
3.22	5.00	4.58	5.00
4.00	4.00	3.00	4.00
3.92	5.00	4.33	5.00
4.33	5.00	4.33	5.00
3.67	3.00	4.38	5.00
3.81	5.00	3.78	5.00
4.33	5.00	4.00	4.00
4.00	4.00		
4.58	5.00		

Total Records = 274
Average accuracy is 90.8709939947569
Build Successful (total time: 5 seconds)

The data shown above clearly demonstrates that the accuracy achieved via proposed methodology is 90.87%

which is very perfect by any standards. The results show that calculated value varies at few places in comparison to user defined value mainly because users tend to give ratings numerically in discrete values but the words explain their experience better. The proposed method evaluates the sentiment on basis of text mining and then calculates the rating based on that. The two values when compared reveal that they are almost 91% similar.
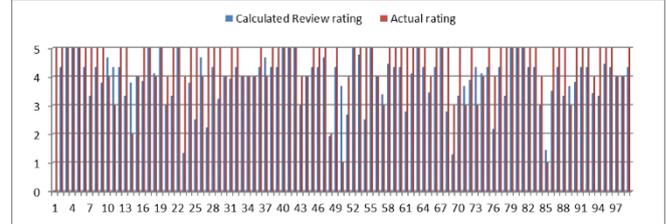
Fig. 3: Chart between calculate review rating and actual rating for 1st 100 reviews

The above chart between calculated review rating by proposed methodology and actual rating given by the customer indicates the similarity in both values. The values differ at few places where the customer rating doesn't match the feedback in words. The comparison reveals 91% accuracy in proposed methodology results.
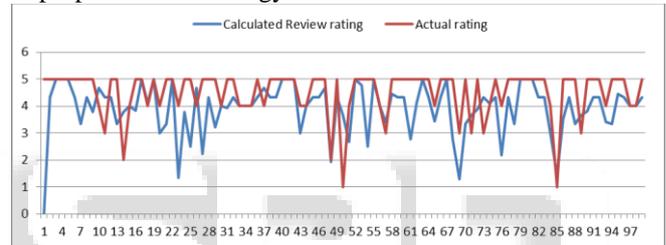
Fig. 4: Line graph between proposed technique results and customer given ratings.

The line graph shown above displays the variation between ratings calculated via proposed technique and the ratings provided by the customer. The values entered by the customer are discrete values which most of the time don't show the exact sentiment. So, the proposed technique calculates the rating based on the sentiment expressed by the user in his reviews. The variation at many a places is mainly due to the concerns expressed by user in words but still given ratings differently.

VI. CONCLUSION

In today's world, the online market is increasingly getting popular and it becomes more and more important to help the customer get the best product by all parameters. The quality of a product is best confirmed by taking the customer reviews from those who are already using that. All popular shopping websites like Amazon, flipkart, ebay etc. allow customer reviews once the product has been purchased. These reviews are such huge in numbers on these websites that it is not possible for a customer to consider them all. So, we propose a technique here, Sentiment Analysis for Product rating through Ontology (SAPRO). This technique performs text mining techniques on the customer reviews obtained from any of these websites and calculates the rating for this product (out of 5), to establish the like-ability of the product by the existing customers. SAPRO considers the end users' perspective while addressing the Sentiment Analysis problem. The research uses a combination approach of

domain ontology and Stanford dependency relation which intends to enhance the sentiment classification. The proposed technique is programmed in Java 8 programming language taking review data from the Amazon website as the dataset which is in JSON format. The results show that the review rating calculated by the proposed technique has 91% similarity with the existing ratings from the customers. The results when compared to base research by ArindamChaudharyet. al [24], clearly show that the proposed technique(91% accuracy) is better than base technique(80-82% accuracy) for reviews. The difference is mainly because the customer gives discrete numeric ratings but explains his sentiment better in words.

As future work the proposed method would be applied to other fine grained text and opinion mining tasks with increasing data volumes and also can be performed on different datasets. The future researchers can also use optimization techniques to increase the accuracy to better levels.. Also experiments can be performed in order to determine to what extent these tasks are jointly modeled in this multitasking framework by incorporating soft computing tools.In future, this technique can be used on various types of data from different sources like hotels, eshopping sites, banks, etc and the effectiveness of proposed methodology can be verified for better results in comparison to other contemporary techniques.

REFERENCES

[1] Fabio Ciravegna and SandaHarabagiu, "Recent advances in natural language processing",IEEE Computer society, 2003.Pp 12-13.

[2] MladenStanojevic and SanjaVranes, "A Natural Language Processing for Semantic Web Services", IEEE 2005.Pp 229-232.

[3] Xiaofei Le and XushengXiel, "Research of Intelligent Word Segmentation and Information Retrieval", ICETC IEEE 2010. Pp 411-414.

[4] Tian Xia in "Improved VSM text classification by title vector based document representation method" ICCSE Pg. 210-213, 2011.

[5] RuiXia, ChengqingZong and Shoushan Li, "Ensemble of feature sets and classification algorithms for sentiment classification", Information Sciences Elsevier 2011. Pp 138-1152.

[6] FisnikKastrati, Xiang Li, ChristophQuix and MohammadrezaKhelghati, "Enabling Structured Queries over Unstructured Documents", IEEE National conference on Mobile Data management, 2011.

[7] S. Samudaria and S. Sasirekha, "Improving the Precision Ratio Using Semantic Based Search", ICSCCN, IEEE 2011, Pp 465-470.

[8] AyuPurvarianti, "A Non Deterministic Indonesian Stemmer", ICEEI, IEEE 2011.

[9] Xueying Zhang, Chunju Zhang, Chaoli Du and Shaonan Zhu, "SVM based Extraction of Spatial Relations in Text", IEEE 2011.Pp 529-533.

[10] Simran Fitzgerald, George Mathews, Colin Morris and OlesZhulyn, "Using NLP techniques for file fragment classification", Digital Investigation Elsevier 2012.Pp S44-S49.

[11] MuktaTakalikar, ManaliKshirsagar and Gauri Dhopvakar, "Intuitive Approaches for Named Entity Recognition and Classification: A survey", IJCA 2013, Pp 35-38.

[12] Nie, Liu, Wang, Song in "The Opinion Mining Based on Fuzzy Domain Sentiment Ontology Tree for Product Reviews" Journal of SW Vol. 8 No.11, 2013.

[13] James Mountstephens, "Mnemonic phrase generation using genetic algorithms and Natural language processing", IEEE 2013, Pp527-530.

[14] Saani H and Reghu Raj P C, "Structured Information Extraction from On-line Advertisements- A Bayesian Approach", IJARCSSE Sep. 2013. Pp 581-586.

[15] Oliver Keszocze, Mathias Soeken, EugenKuksa and Rolf Drechsler, "lips: An IDE for Model Driven Engineering Based on Natural Language Processing", IEEE 2013. Pp 31-38.

[16] [Ibrahim EldesokyFattoh, AmalElsayedAboutabl and Mohamed Hassan, "Tappinginto the Power of Automatic Question Generation", IJCA Vol. 103, No. 1, Oct. 2014. Pp 1-6.

[17] M.Kasthuri and S.Britto Ramesh Kumar, "An Improved Rule based Iterative Affix Stripping Stemmer for Tamil Language using K-Mean Clustering", IJCA Vol. 94, No. 13, May 2014. Pp 36-41.

[18] Maryam, SeemaKoulkur in "Feature ranking in sentiment analysis", International Journal of Computer Applications, 2014.

[19] Sheetal Pereira, Uday Joshi in "Implementation of SVM technique in feedback analysis system", International journal of computer applications, 2014.

[20] DeepshikhaChaturvedi and Shalu Chopra, "Customers Sentiment on Banks", IJCA Vol. 98 No. 13, July 2014. Pp 8-13.

[21] Roul, Devanand, Sahay in "Web Document Clustering and Ranking using Tf-Idf based Apriori Approach", International Conference on Advances in Computer Engineering & Applications,2014.

[22] ChandhanaSurbhi in "Natural language processing future", International conference on optical imaging sensor, 2013.

[23] Kumar Ravi and VadlaMani Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications", Knowledge-Based Systems 89-91, Elsevier 2015.

[24] ArindamChaudhuri and Soumya K. Ghosh, "Sentiment Analysis of Customer Reviews Using Robust Hierarchical Bidirectional Recurrent Neural Network", Advances in Intelligent Systems and Computing 464, Springer 2016.