International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 3 Issue: 9 5574 - 5577

User Feedback Analysis System using Natural Language Processing and Artificial Intelligence

Sameer N. Kolhe,
Dept.of Computer,
BVCOERI,Nashik
sameerkolhe77@gmail.com

Swati V. Darade,
Dept. of Computer,
BVCOERI, Nashik
swatidarade243@gmail.com

Pratiksha B. Suryavanshi,
Dept. of Computer,
BVCOERI, Nashik
prati.suryavanshi31@gmail.com

Pramod S. Ahire,
Dept. of Computer,
BVCOERI, Nashik
pramod.ahire18@gmail.com

H. D. Sonawane,
Dept. of Computer,
BVCOERI, Nashik
hod.comp @engg.brahmavalley.com

***Abstract*—** The Web has dramatically changed the way that people express their views and opinions. Now if one wants to purchase a product, he/she is no longer limited to asking his/her friends and families because there are many product reviews on the Web which give opinions of existing users of the product. Here we present the system which provides us information about such products and services in summarization form. Finding opinion sources and monitoring them on the Web can still be a difficult task because there are a large number of diffrent sources, and each source may also have a huge volume of opinionated text (text with opinions or sentiments). In most cases, opinions are hidden in long forum posts and blogs. It is complicated for a human reader to find relable sources, extract related sentences with opinions, read them, summarize them, and manage them into usable forms. Thus, automated summarization systems are needed. Using this summarization we can recognize the importance, quality, popularity of product and services. In this system we make summarization for product. But, we can use this system anywhere, where text analysis is required. Sentiment analysis, also known as opinion mining, grows out of this need. It is a challenging natural language processing or text mining problem. Due to its tremendous value for practical applications, there has been an excessive growth of both research in academia and applications in the industry.

Keyword: *Natural language processing, Machine Learning, Sentiment Analysis, Opinion Mining.*

*****

I. INTRODUCTION

IN recent years, with the growing volume of online reviews available on the Internet, sentiment analysis and opinion mining, as a special text mining task for determining the subjective attitude (i.e., sentiment) expressed by the text, is becoming a hotspot in the field of data mining and natural language processing. Sentiment classification is a basic task in sentiment analysis, with its aim to classify the sentiment (e.g., positive or negative) of a given text. The general practice in sentiment classification follows the techniques in traditional topic-based text classification, where the bag-of-words (BOW) model is typically used for text representation. In the BOW model, a review text is represented by a vector of independent words. The statistical machine learning algorithms (such as na€ıve Bayes, maximum entropy classifier, and support vector machines) are then employed to train a sentiment classifier. Although the BOW model is very simple and quite efficient in topic-based text classification, it is actually not very suitable for sentiment classification because it disrupts the word order, breaks the syntactic structures, and discards some 1semantic information. Consequently, a large number of researches in sentiment analysis aimed to enhance BOW by incorporating linguistic knowledge. However, due to the fundamental deficiencies in BOW, most of these efforts showed very slight effects in improving the classification accuracy. One of the most well-known difficulties is the polarity shift problem. Polarity shift is a kind of linguistic phenomenon which can reverse the sentiment polarity of the text. Negation is the most important type of polarity shift. For example, by adding a negation word “don’t” to a positive text “I like this book” in front of the word “like”, the sentiment of the text will be reversed from positive to negative. However, the two sentiment-opposite texts are considered to be very similar by the BOW representation. This is the main reason why standard machine learning algorithms often fail under the circumstance of polarity shift. "

II. OBJECTIVE

The purpose of this document is to present a detailed description of the product rating & Review Summarization. It will explain the purpose and features of the system, the interfaces of the system, whatever the system will do, the constraints under which it must operate and how the system will react to external stimuli.

III. PROBLEM STATEMENT

A large multinational food corporation is assessing consumer preferences for its fast food product line comprising of instant noodles & other rice and wheat based products. The questionnaire is focused on how to increase consumption frequency based on pack size, positioning it as snack between meals, ready-to-eat capability, variety of available flavors & promotional tie-ins with other products. The Product Manager has commissioned a market research agency to do a consumer study to rank people’s preferences for increased

consumption with respect to all these variables. The agency has come back with data from 1000 participants.

Checking Text:

The basic question asked in Sentiment Analysis is whether a given piece of text contains any subjective content (opinions, emotions, etc.) or not. This task aims to tackle this problem of differentiating between subjective and objective content.

Verifying discrete polarities:

Once the subjective part is determined, the next step is to determine if the content is positive or negative. This problem can be looked upon as a classification problem.

Identifying an ordinary value:

Some applications require not just the type of polarity but the intensity as well. For example, movies are typically rated on a 5 point scale. Thus, this step aims at identifying such an ordinal value.

Identifying subjective portions of text: The same word can be treated as subjective in one context, while it might be objective in other things. This makes it difficult to identify the subjective (sentiment-bearing) portions of text. Example:

- The language of the author was so crude.
- Crude oil is extracted from the sea .

The same word “crude" is used as an opinion in first sentence, while it is completely objective in the second sentence.

Associating sentiment with specific keywords: Many sentences indicate an extremely strong opinion, but it is difficult to pinpoint the source of these sentiments. Hence an association to a keyword or phrase is mostly difficult. For example:

- Each time I read `Pride and Prejudice' I Need to dig her up and beat her over the skull with her own shin-bone.

In this example, “her" refers to the character in the book “Pride and Prejudice", which is not externally mentioned. In these cases the negative sentiment must be associated with the character in book.

IV. LITERATURE SURVEY

We have been gathering information about the sentiment analysis and opinion mining techniques available and how does they work. The phenomenon that they have come across consists in special challenges associated with mining on e-learning reviews and studying e-learning blogs which make it very challenging task and adds to its complexity, and it is this factor that was at origin of our loss of accuracy. In this future work is: combining some of these feature selections, processing refinement and taking into account emotions and misspelled words and employing different linguistic techniques.

V. EXISTING SYSTEM

Finding opinion sources and monitoring them on the Web can still be a difficult task because there are a large number of diverse sources, and every source may also have a big volume of opinionated text (text with opinions or sentiments). Selecting an attributes for sentiment classification using feature relation networks[1]. In many cases, opinions are hidden in long forum posts and blogs. It is complex for a human reader to find relational sources, extract relational sentences with opinions, read them, understand them, and organize them into usable forms. Thus, automated summarization systems are needed. Using this summarization we can recognize the importance, quality, popularity of product and services. In this system we make summarization for movie. But, we can use this system anywhere, where text analysis is required[2]. Sentiment analysis, also known as opinion mining, grows out of this need. It is a challenging natural language processing or text mining problem[3]. Due to its tremendous value for practical applications, there has been an explosive growth of both research in academia and applications in the industry[4].

VI. PROPOSED SYSTEM

With the flourish of the Web, online review is becoming a more and more useful and important information resource for people. As a result, review mining and summarizing has become a hot research topic recently. Different from traditional text understanding, review mining and summarizing aims at extracting the features on which the reviewers express their opinions and determining whether the opinions are positive or negative. In this paper, we focus on a specific domain movie review. A multi knowledge based approach is proposed, which integrates WorldNet, statistical analysis and movie knowledge. The experimental results show the effectiveness of the proposed approach in movie review mining and summarizing.

We are designing one system that can help any organization for improving their products, analyze thousands of feedback and provide generalized opinion for the product. The design can also be extended to other product-review domains easily.

This System is specifically made for-When an organization wanted to find opinions of the general public about its products and services, it conducted surveys and focus groups. Such that, with the explosive growth of the social media content on the Web in the past few years, the world has been transformed. People can now post reviews of products at merchant sites and express their views on almost anything in discussion forums and blogs, and at social network sites. Now if one wants to buy a product, one is no longer limited to asking one’s friends and families because there are many user reviews on the Web. For a company, it may no longer need to conduct surveys

or focus groups in order to gather consumer opinions about its products and those of its competitors because there is a plenty of such information publicly available

Our Approach to do Sentiment Analysis:

1. Fetch comments provided by the user for processing.
2. Using dictionary approaches to determine the product, user is speaking about.
3. Create dictionaries for weak and strong sentiment related patterns.
4. Apply strong negative sentiment patterns to the input in relative to the product.
5. If not found try searching for weak negative patterns.
6. Search for positive sentiment patterns in the comments with relative to the product.
7. If positive sentiment pattern is found make sure that it does not have negative pattern preceding it. If found just flip the polarity of the sentiment to negative.

E.g.:

✓ Apple iPhone 4 is not made me happy at all.

✓ Here we need to flip the polarity, as positive pattern is preceded by negative one.

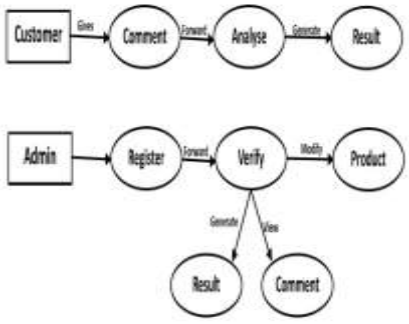
✓ Samsung Y is indeed a **smart** phone.

✓ Here the term 'Smart' indicates that the sentiment is positive for Samsung Ace.

✓

VII. ACTIVITY FLOW

Data Flow Diagram–Level 0



Data Flow Diagram–Level 1

VIII. MATHEMATICAL MODULE

1) Let S = { I, P, O, S }

Where I ⟶ Inputs
P ⟶ Process
O ⟶ Output
S ⟶ System

2) Let,

I Stands for comment data set
P stands for Process on each Input
O stands for Output
S system to be used

3) So,

$S_i = \{ I_1, I_2, I_3 \ldots\ldots I_n \}$ …set of inputs

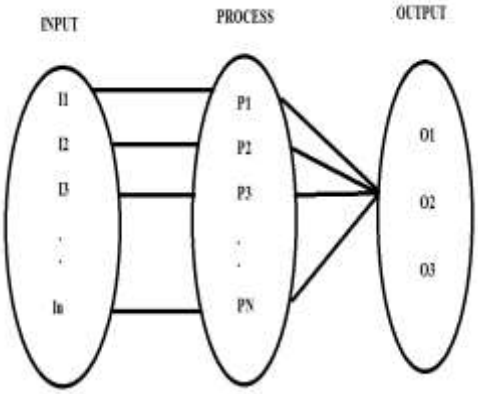
$S_P = \{P_1, P_2, P_3, \ldots\ldots P_n\}$ ..set of processes

$S_o = \{O_1, O_2, O_3\}$ ...set of outputs

$S_s = \{NLP, JAVA\}$ ..set of systems

4) So, $S = \{ (I_1, I_2, I_3 \ldots\ldots I_n), (P_1, P_2, P_3, \ldots\ldots P_n), (O_1, O_2, O_3), (NLP, JAVA)\}$

5) Execution Flow :



IX. ALGORITHM

1. Fetch comments provided by the user for processing.

2. Using dictionary approaches to determine the product, user is speaking about.

3. Create dictionaries for weak and strong sentiment related patterns.

4. Apply strong negative sentiment patterns to the input in relative to the product.

5. If not found try searching for weak negative patterns.

6. Search for positive sentiment patterns in the comments with relative to the product.
7. If positive sentiment pattern is found make sure that it does not have negative pattern preceding it. If found just flip the polarity of the sentiment to negative.

X. MODULES

END-USER MODULE:
1. Login
2. Select product
3. Check comments.
4. Check graphical rating related to product.
5. Exit

MAIN-SYSTEM MODULE:
1. User Authentication
2. Retrieving product details
3. Retrieving comments of such product from dataset.
4. Generate Graphical rating based on comments over product.
5. Display graphical rating.
6. Stop

XI. CONCLUSION

On real-life applications, to provide a completely automated solution is nowhere in sight. However, it is possible to devise effective semi auto generated solutions. The key is to fully understand the whole range of issues and pitfalls, catastrophically manage them, and determine what portions can be done automatically and what portions need human assistance. In the continuum between the whole manual solution and fully automated solution, we can increase more and more toward automation.

XII. ACKNOWLEDGMENT

It is our pleasure to express our knowledge to our respected sir Prof. H. D. Sonawane, Computer Engineering Department, BVCOE&RI, Nashik for his valuable guidance, inspiration and continues support. This paper could not be success without protocol study done which help to understand the necessity for this paper.

XIII. REFERENCES

[1] A. Abbasi, S. France, Z. Zhang, and H. Chen, “Selecting attributes for sentiment classification using feature relation networks,” IEEE Trans. Knowl. Data Eng., vol. 23, no. 3, pp. 447–462, Mar. 2011.
[2] E. Agirre, and D. Martinez, “Exploring automatic word sense disambiguation with decision lists and the web,” in Proc. COLING Workshop Semantic Annotation Intell. Content, 2000, pp. 11–19.
[3] J. Cano, J. Perez-Cortes, J. Arlandis, and R. Llobet, “Training set expansion in handwritten character recognition,” in Proc. Struct., Syntactic, Statistical Pattern Recognit., 2002, pp. 548–556.
[4] Y. Choi and C. Cardie, “Learning with compositional semantics as structural inference for subsentential sentiment analysis,” in Proc. Conf. Empirical Methods Natural Language Process., 2008, pp. 793–801.