Analysis of Behavior Extraction on Social Life Issues Using Sentiment Analysis: A Review

Adnan Beg
B.Tech. Student, Computer Science and Engineering Department, Integral University, Lucknow, India

Abstract: At present time most of the world can be found on internet. The popularity of the social networking sites has been growing fast, parallel to emerging technologies, with increase in the number of users to the social networking sites that actively express their opinions on these sites. To forecast the sentiment analysis we have used the data stored on social sites stockpiling. Our goal is to retrieve data from social sites and analyze the emotion of a particular person on certain topic. For our analysis we have used the data available on Twitter, as Twitter has a large amount of data that people post, which gives the output beyond the polarity but those polarities can be used in product profiling, trend analysis and forecasting. This paper presents an overview of past and current research on twitter sentiment analysis and presents better ideas for future work.

1. Introduction

People make judgments about the world around them where they are living. They create positive and negative attitudes about people, products, places and events according to their understandings. These types of attitudes can be considered as sentiments. Sentimental analysis also known as opinion mining is the process to gain an understanding of the attitudes, opinions and emotions expressed by people on social networking sites. Sentimental analysis is found extremely useful in social media monitoring as it allows us to gain an overview of the opinions of wider public behind certain topics. Twitter has 310 million monthly active users, who exchange information through tweets of length approx. 140 characters. Tweets on Twitter are delivered in real time from any geographical area and by anyone irrespective of the user backgrounds such as religion, political inclination or the level of education or civilization. In this study, the author has used these tweets tweeted on Twitter on a certain topic, to analyse whether the sentiments of the people are positive, negative or neutral. As web data like tweets on twitter is large amount of data generated by people so it’s become important to work for an efficient intelligent systems that can do refinement of data and analysis of tasks intelligently and efficiently. A significant research work can be seen since 1980 on different aspects of twitter sentiment analysis. Now we will see major work of these researches.

2. Literature Survey

2.1 Sentence based analysis [1]

In this paper the author had studied text to speech analysis of data. They study about processing of data and ability to render natural expressive speech. The work is mainly about the techniques that are used in natural expressive speech and how data of natural expressive speech is being used. Next we will see the process of natural expressive speech using the diagram shown by author. This work basically focuses on production of synthetic speech and adaption of sentiment analysis [11] procedure that can then be used as input feature for expressive speech synthesis. So, we can say that the work of this paper revolves around data to speech synthesis.

This paper is mainly concerned with the identification that how text and web data can effect in different aspects and tries to categories data on the basis of its effects. This work was the first attempt to adapt conventional sentiment analysis work. In this work author had used small training data for the study so the classification can be improved in terms of speed and accuracy when applied to large data sets. The conclusion to this research paper is using a fast scripting language the functionalities the author wants to introduce can be achieved up to a level.

2.2 Automated Topic Modeling and Sentiment Analysis of Tweets on Spark R [2]

In this paper the author had proposed an automated topic modeling technique LDA which helps in identifying the attracted matter of arguments from huge amount of tweets co-related to any two leading and famous political leaders of India. In this paper, the author takes a spark R framework on which the topic modeling technique is performed to upgrade to the speed and performance for vast social data processing and its analysis in real time. At that point sentiment analysis, those tweets are finished by utilizing dictionary based ways to deal with arrange the general population’s sentiment against the two leaders. The result shows that the proposed method has improved the speed compared to the normal R tool.

2.3 RA summarization of short comments [3]

Along a timeline (TASC-t) for dynamic tweets and experimented on 6 topics from published corpuses demonstrating that usefulness of the TASC algorithm. Their future work emphasized on accurate labeling of data and a good

In this paper author proposes a framework to summarize the opinions for some product and item based on user votes. Basically, it is the study of generating a RA summary of short comments, which is a decomposed view of the overall ratings for the major aspects so that a new user can make a best choice according to the experience and perspective of the old user.

The output is that overall rating is decomposed into several small aspects each having support information showing the
信心。未来的研究将侧重于对产品评论和卖家服务的比较研究，以提高市场营销。这种排名机制可以应用于各种产品以产生更准确的输出。

2.4 Field adaptive classification of tweets [4]

在这个论文中，作者研究了推特词的分类，这是一种将表现推特词不同的方法。由于推特领域非常多样化，因此不可能构建通用分类器，推特词缺乏数据标注，使它们更难以分类。TASC算法是该作者使用的更新主题-适应性特征方法，基于协作性对数据的未知值，这有助于从更可靠的推特中选择更多的话题，以提高性能。

算法中所采用的技术用于不同领域字段的推特词分类。表1列出了示例数据中不同特征字段的分类结果。它们也设计了适应模型

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Topics</td>
<td>Positive</td>
<td>Negative</td>
<td>Neutral</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Apple</td>
<td>1911</td>
<td>377</td>
<td>581</td>
<td>1,149</td>
<td></td>
</tr>
<tr>
<td>Microsoft</td>
<td>93</td>
<td>138</td>
<td>671</td>
<td>902</td>
<td></td>
</tr>
<tr>
<td>Google</td>
<td>218</td>
<td>61</td>
<td>604</td>
<td>883</td>
<td></td>
</tr>
<tr>
<td>President Debate</td>
<td>1,465</td>
<td>729</td>
<td>1,109</td>
<td>3,213</td>
<td></td>
</tr>
<tr>
<td>Twitter</td>
<td>68</td>
<td>78</td>
<td>647</td>
<td>793</td>
<td></td>
</tr>
<tr>
<td>Taco Bell</td>
<td>902</td>
<td>596</td>
<td>2,099</td>
<td>3,597</td>
<td></td>
</tr>
</tbody>
</table>

3. Other Related Works

其他研究工作包括跨领域情感分类[5]。使用这种特征值的扩充和选择，根据信息获取标准对跨领域情感分类工作的表现进行评估，较之其他方法，表现更好。

研究工作也显示了在某些实时主题领域的情感分类工作在描述辩论性能[6]。它们表明视觉和数据可视化可以用于识别社交媒体事件的视觉分析系统。

辩论分类也应用于Twitter[7]。Twitter作为一个新领域对象，是独特的提案用于情感分析。它们的短文档长度表明，任何情感表示方式都是紧凑而明确。一些研究发现Twitter中的情感分类在Twitter上的表现更好，这是由表情符号在Twitter上的使用。这些表情符号的辅助语言模型正在构建数据清洗和分类。这个模型被称为ESLAM。情感分类是[9]中的一种情感分类，其中大量的工作是用于抽取不同行为表现的推文。“这为这种多引擎监督分类器方法提供了依据，作者用于识别情绪话题来评估英文博客的输出。”

4. Conclusion

作为社交媒体平台，Twitter是一个适用于获取所有信息的平台，但可能的步骤可以被采取以获得有用信息。这可以通过通过跟踪博客领域的sentiment分析来实现，这可以实现对sentiment分析的理解，从而可以实现对更多领域Twitter的sentiment分析。

References

[1] Alexandre Trilla and Francesc Aliash;“Sentence-Based Sentiment Analysis for Expressive Text-to-Speech” IEEE transaction on audio, speech and language processing, February 2013, Page 223-233.