

Website Evaluation System Using Opinion Mining

Suyesh Pandey¹, Vinamra Rathi², Tilak Badaya³, Bindu Bhargavi⁴

^{1,2,3}UG Student, Department of Information Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India

⁴Assistant Professor, Department of Information Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India

Abstract - There are many websites available online today. People use different websites for different purposes. With the large number of websites available to do the same job, it becomes difficult for a user to know which website is right for their purpose. Many websites are unreliable and their users are often victims of fraud. Rating these websites based on user feedback is a way to make sure users know which website to choose. This paper proposes a method for processing the comments made by the users and generating ratings according to those comments using Python libraries such as textblob to process the comments and Naive-bayes as the classifier to classify the comments.

Key Words: TextBlob, Opinion Mining, Review, Comment, Fraud, Naive-bayes

1. INTRODUCTION

There has been a growing increase in the number of online websites. These websites have a variety of products and services that are used by a large number of users. As the number of the users and also the number of websites available online is increasing day by day, we have all been experiencing that the navigation of these websites is just satisfactory and are missing important information. Taking into account all the available websites online every user wishes to find and access useful information. By the help of this information user may choose which website to opt for according to their need. Trust is the most important factor between these users and websites. By simply establishing trust, users can become more confident about the websites they use. However, most of the time, websites do not make their credibility and users often face losses as a result. This can be avoided by using a process known as website evaluation.

Website evaluation is a test performed to remove unwanted content that is on the website with great ease. In order to establish trust, a website test was proposed using opinion polls. The process involves taking input from users in relation to the websites they have used and their information on those websites. After that, user views are processed with a model that provides a rating such as output. Based on the criteria provided, it is determined whether the content provided on a web page is trustworthy or not, thus establishing trust.

1.1 Motivation

Because of the huge quantity and variety of current Web sites, it is very easy for a common user to get caught in a trap of fraud. These numerous websites need mechanised analysis of websites. The popularity of sentiment analysis and opinion mining is skyrocketing. This inspired us to build a website evaluation system using opinion mining which can not only suggest better websites to opt for but can also decrease fraud.

2. LITERATURE REVIEW

[1] Title: "A Language Model Approach for Retrieving Product Features and Opinions from Customer Reviews" Lisette Garcia-Moya, University Jaume I proposed the paper in 2013. This paper proposes a feature-based summary function by introducing a novel approach to retrieving product features from a collection of free customer reviews by text about a product or service. The basic approach used behind this paper is divided into three modules. The first module consists of sentence separation, analysis of noun phrases and dependency analysis. The second module consists of training the ME model and the latter module consists of extracting product idea ideas from unwritten reviews. In the pre-processing module the pipeline review includes processing, noun analysis, and dependency analysis First, we analyze the review sentences with Stanford lexicalized parser. Syntactic parse trees are automatically converted into dependable presentations. The ME training model has three steps. First, prepare training data that includes pre-processing and annotation of product views. Second, subtract the learning features of each pair in the training data. Lastly, train the model using the advanced entropy model. The effect of the training model is the weight of each feature function. In order to exclude product features and to identify ideas related to these features (product of a two-dimensional idea), we rely on the realization that there are feature names used to describe the product feature and the name of the idea. We have found that many expressions that reflect words are adjectives and nouns form product features.

The system provides a method that only allows those users to write a review about a product they purchased on a website. Some users were not allowed to provide updates. This reduces negative reviews of products and customers find reliable products.

[2] Title: "Effective Product Ranking Method based on Opinion Mining" Madhavi Kulkarni and Mayuri Lingayat proposed the paper in 2015. This paper proposes a system that makes the classification of updates provided by customers followed by the sensitivity of the reviews provided. The proposed system takes data from a random online shopping website where reviews of two products are taken. Both product reviews are operational and according to the polarity of the terms used in the reviews and comments, products are counted. In general, a high-quality product may not have the best qualities, or it may not be the best, but after product use, the updates provided to the product are also the most important factors while providing updates about the product. An official release of a given product feature is performed. Their system displays a sense of revision that is then presented in the form of a chart. The division of reviews and emotional analysis increased the accuracy of the system which also provides accurate reviews for the user.

[3] Title: "Review on Techniques and Tools used for Opinion Mining" Asmita Dhokrat, Sunil Khillare, C. Namrata Mahendar proposed the paper in 2015. Opinion mining is an emerging data mining platform used to summarize information from large amounts of data which can be customer feedback, feedback and reviews on any product or topic etc. which summarizes updates and other realtime applications. There are many other applications that can be talked about. It is designed that sensory separator rely heavily on courses or topics. From the above work it is evident that no model of continuous separation surpasses another, different face types have distinguished the distribution. It also indicates that the various types of features and algorithms of the categories are properly integrated to overcome their individual shortcomings and benefit from each other's merits, and at least improve the functioning of the emotional category.

This paper outlines the basic requirements for the excavation to explore current strategies used to create a complete plan. Highlights the opportunities or deployments and research of these programs. The tools available to build those programs are presented with fairness and limitations. In their study, they found some strategies such as Naive Bayes, Maximum Entropy, and SVM etc. They are widely used in opinion testing and emotional analysis.

[4] Title: "Twitter opinion mining framework using hybrid classification scheme 2014" Farhan Hassan Khan, Saba Bashir, Usman Qamar proposed the paper in 2014. In the proposed paper during the pre-processing steps each word was tested for that are available in the dictionary and the missing words are taken as abbreviations or slangs. Abbreviations are then replaced by expansions. Lemmatization is then applied and the spelling is finally checked to correct the lemmatizer effect. This proposed paper suggests a new framework for Twitter Opinion Mining to predict the polarity of words into good or negative words in tweets. Improves phase accuracy level. It is built using various stages.

[5] Title: "Opinion Mining and Analysis: A survey" Arti Buche, Dr. M.B. Chandak, Akshay Zadgaonkar proposed the paper in 2013. This proposed paper provides information on data sources used for excavation, study equipment and sensory analysis activities for sensory analysis, textual classification and available emotional classification tools, and performance evaluation. Natural Language Processing (NLP) works by processing real text objects. The text object is converted to machine format by NLP. Artificial Intelligence (AI) uses data provided by NLP and works with a number of statistics to determine if something is healthy or not negative. There are many ways to determine an author's point of view on a topic from a natural language text detail. A different, more effective, machine-based learning method is used. One of the forms of natural language analysis is to draw attention to people's perceptions of a particular product or topic. This software provides automatic extraction of ideas, feelings and emotions in the text and tracks attitudes and feelings on the web. People express their opinions by writing blog posts, comments, reviews and tweets about all sorts of different topics. Tracking products and products and determining whether they are viewed positively or negatively can be done using the web. Mind mining has slightly different functions and more words, e.g., emotional analysis, emission of ideas, emotional mines, subjectivity analysis, touch analysis, emotion analysis, mining reviews, etc.

The current study focuses on an emotional analysis called Opinion Mining because of the wide range of ideas for web resources such as forums, review sites and blogs available in terms of prices. One important problem in the emotional analysis of product reviews is to produce a summary of ideas based on product features. Various strategies designed for important mining activities were examined and analyzed in the proposed system. The proposed paper provides comprehensive overview of what is involved in the development of data mining software on the basis of our survey and analysis.

The Red Opal tool helps users identify products that rely on their characteristics. Product ideas are provided with a graph format by feature.

[6] Title: "A review of website evaluation studies in the tourism and hospitality fields from 1996 to 2009" Rob Law, Hee 'Andy' Lee proposed the paper in 2011. Proposed course review of website test papers in the tourism and hospitality industry published between January 1996 and September 2009. This study aims to review articles published in website evaluation in hospitality and tourism in the period 2000-2015 to provide a comprehensive review, as well as to provide results to academic and industry researchers. Content analysis was adopted by this study to review articles found in website reviews and hospitality. The essays were then analyzed from the perspective of consumers, from the perspective of suppliers, and from the opinions of consumers and suppliers using a systematic approach. Major findings of this study have shown that most articles focus mainly on hospitality or tourism. In addition, many articles often discuss user interaction, advertising

performance and website quality. However, these articles did not discuss in detail the results of the website test and ignored the interaction between providers and consumers to some extent. This study reviews the development of website testing in hospitality and tourism in the new millennium. The findings of this study provide important findings for hospitality and tourism researchers to promote consumer and consumer engagement.

A website evaluation framework that includes phased evaluation, feature testing and effectiveness is conducted. The strengths and weaknesses of each test method are analyzed, taking into account future research spaces and indicators for future research.

3. IMPLEMENTATION

When the application is launched, it opens up on the home page which shows all the posts created.

- If the user wants to post or comment, the user can log in or sign up.
- In, Login Page The user is prompted to enter their username and password on the login screen.
- After a successful login, the menu page opens, and displays all the available posts and if posted by the logged in user they can edit or delete.
- On clicking create, the user will be sent to the create Page, where they will be asked to enter the Title and URL of the website. The post is successfully stored in the DB and shown on the Home Screen.
- The user will be led to the home screen, and the user can again create, view or delete posts.
- LOGIN: This module will prompt the user for their account and password. In discourse, this will be acceptable. The home page will be displayed to the user after a successful login.
- Create: To create a new post, select this option. The user will just have to provide a title and url as input. After the input of all the content, the system will verify and store in DB. Following completion of all needed inputs, the system will redirect to the home page.
- Edit: This inbox will be displayed when the user selects Edit. The user will be able to update the post title or url.
- Delete: This component will delete the blog if made by the user from the DB and home page and all the data related to the post will be deleted.
- Comments: Users will be able to comment on the particular post.
- Rating: User will be able to see the rating of the website posted
- LOGOUT: Allows the user to log out of the app and returns them to the login screen.

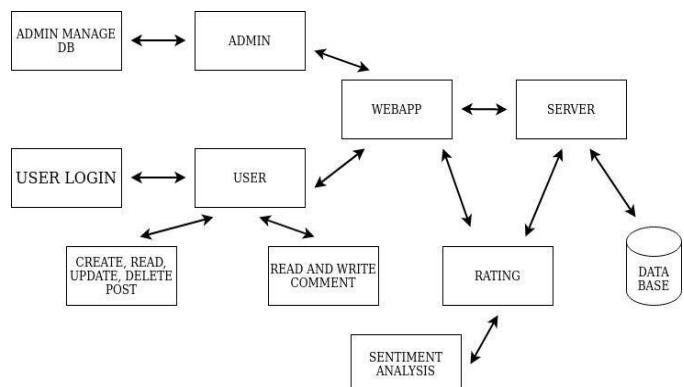


FIG -1: System Architecture

A. User Interface Design: HTML, CSS3, and Django were used to create the user interface. It correctly verifies if the user is authenticated and only allow them to make changes on their posts and comment

B. System Design: The system is designed as shown in the Architecture diagram below, which depicts the system's whole comprehensive sequence of events. As seen in the figure below, all activities are carried out and rating of the post is displayed

3.1 TECH STACK AND RESULTS

Python is the programming language used in the initial phase, to build all the modules and functionality. The frontend of the website, which has functionality such as user login, post construction and add comment is done using Django which is an open-source based platform based on Python. To perform the opinion mining and sentiment analysis, the TextBlob library is used for preprocessing and then Naive-bayes, which is a classification algorithm, is used to classify ideas and generate websites.

A product developed using the tools, technologies and operating methods mentioned above, is capable of producing the ratings of different websites that allow the trust of the users of those websites. Hence, by looking at the ratings, the user can now know which website to opt for.

4. CONCLUSIONS

The Proposed system gives the user a better understanding about websites. It prevents the user from being a victim of fraud. It will give them a detailed survey about the service or customer feedback regarding those websites. The proposed system could be a criticism for the sites hence driving them towards improvement on their drawbacks. On the basis of data availability our system can be expanded to different sectors. It is also open to customer's views based on their frequent experience with those websites.

The result expected from the proposed system seems to be efficient, convenient, time saving and trustworthy for the users. The proposed given system helps the user to choose the best available website from the various available websites on the web.

REFERENCES

- [1] Gautami Tripathi, S Naganna, "Opinion Mining: A Review", International Journal of Information & Computation Technology.
- [2] P Haseena Rahmath, "Opinion Mining and Sentiment Analysis - Challenges and Applications", International Journal of Application or Innovation in Engineering & Management (IJAEM), vol. 3, no. 5, May 2014.
- [3] Nidhi R. Sharma, Vidya D. Chitre, "Opinion Mining Analysis and its Challenges", International Journal of Innovations & Advancement in Computer Science, vol.3, no. 1, April 2014.
- [4] S. Chandra Kala, C. Sindhu, "Opinion Mining And Sentiment Classification: A Survey", IACT Journal on Soft Computing, vol. 3, no. 1, October 2012.
- [5] G.Vinodhini, R.M. Chandrasekaran, "Sentiment Analysis and Opinion Mining: A Survey", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2, no. 6, June 2012.
- [6] Bing Liu, 2012, Sentiment analysis and opinion mining, Morgan and Claypool publishers.
- [7] "Opinion Mining and Sentiment Analysis in Information Retrieval Vol. 2, Nos. 1 - 2 (2008), Bo Pang and Lillian Lee.
- [8] Ananta Arora et al, "Opinion Mining: An Overview", International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 11, November 2015.