IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2024

DOI: 10.17148/IJARCCE.2024.13590

Fake Product Review Monitoring and Removing Using Opinion Mining

Prof. Nandini G R[1], Leena M S[2], Arusitha R K[3], Arshiya Banu[4], Kommineni Bavyasree[5]

Assistant Professor, Information Science and Engineering, SSIT, Tumakuru, India.[1]

Students, Information Science and Engineering, SSIT, Tumakuru, India. [2-5]

Abstract: Data mining and opinion mining used by organization in gaining customers and increasing sales. Online Customers take keen interest in creating system can take the comments and review of people on a certain product as input, and after applying properly developed mathematical model for predicting the reliability rate of a product, gives the anticipated results on the sale of one to ten. The mathematical model requires a properly designed methodology that can search in the historical data of each component and the newly collected data, the number of fake feedbacks created by competitors and general public by searching for the IP address of each of these reviews that influences the success and failure rate of the product. If same type of comments (negative feedback) comes from an IP address, it is blocked from the website. The final output is displayed based on multiple threshold calculations of the already given parameters and the results obtained from public platform on social media. The admin is responsible for adding new word to the database, adding information about the product and the specifications.

I. INTRODUCTION

What is Data Mining?

Generally, data mining (sometimes called data or knowledge discovery) is the process of analysing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyse data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

How Data Mining Works?

While large-scale information technology has been evolving separate transaction and analytical systems, data mining provides the link between the two. Data mining software analyses relationships and patterns in stored transaction data based on open-ended user queries. Several types of analytical software are available: statistical, machine learning, and neural networks.

Generally, any of four types of relationships are sought:

- **Classes**: Stored data is used to locate data in predetermined groups. For example, a restaurant chain could mine customer purchase data to determine when customers visit and what they typically order. This information could be used to increase traffic by having daily specials.
- **Clusters**: Data items are grouped according to logical relationships or consumer preferences. For example, data can be mined to identify market segments or consumer affinities.
- **Associations**: Data can be mined to identify associations. The beer-diaper example is an example of associative mining.
- **Sequential patterns**: Data is mined to anticipate behaviour patterns and trends. For example, an outdoor equipment retailer could predict the likelihood of a backpack being purchased based on a consumer's purchase of sleeping bags and hiking shoes.

II. LITERATURE SURVEY

1) Addressing complex and subjective product-related queries with customer reviews: AUTHORS: J. McAuley and A. Yang:

Online reviews are often our first port of call when considering products and purchases online. When evaluating a potential purchase, we may have a specific query in mind, e.g. `will this baby seat fit in the overhead compartment of a 747?' or `will I like this album if I liked Taylor Swift's 1989?'. To answer such questions, we must either wade through huge volumes of consumer reviews hoping to find one that is relevant, or otherwise pose our question directly to the

DOI: 10.17148/IJARCCE.2024.13590

community via a Q/A system. In this paper we hope to fuse these two paradigms: given a large volume of previously answered queries about products, we hope to automatically learn whether a review of a product is relevant to a given query. We formulate this as a machine learning problem using a mixture-of-experts-type framework---here each review is an `expert' that gets to vote on the response to a particular query; simultaneously we learn a relevance function such that `relevant' reviews are those that vote correctly. At test time this learned relevance function allows us to surface reviews that are relevant to new queries on-demand. We evaluate our system, Moqa, on a novel corpus of 1.4 million questions (and answers) and 13 million reviews. We show quantitatively that it is effective at addressing both binary and open-ended queries, and qualitatively that it surfaces reviews that human evaluators consider to be relevant.

2) Image based recommendations on styles and substitutes

AUTHORS: J. J. McAuley, C. Targett, Q. Shi, and A. van den Hengel

Humans inevitably develop a sense of the relationships between objects, some of which are based on their appearance. Some pairs of objects might be seen as being alternatives to each other (such as two pairs of jeans), while others may be seen as being complementary (such as a pair of jeans and a matching shirt). This information guides many of the choices that people make, from buying clothes to their interactions with each other. We seek here to model this human sense of the relationships between objects based on their appearance. Our approach is not based on fine-grained modelling of user annotations but rather on capturing the largest dataset possible and developing a scalable method for uncovering human notions of the visual relationships within. We cast this as a network inference problem defined on graphs of related images, and provide a large-scale dataset for the training and evaluation of the same. The system we develop is capable of recommending which clothes and accessories will go well together (and which will not), amongst a host of other applications.

3) Diffusion of innovations revisited: from social network to innovation network

AUTHORS: X. Rong and Q. Mei

The spreading of innovations among individuals and organizations in a social network has been extensively studied. Although the recent studies among the social computing and data mining communities have produced various insightful conclusions about the diffusion process of innovations by focusing on the properties and evolution of social network structures, less attention has been paid to the interrelationships among the multiple innovations being diffused, such as the competitive and collaborative relationships between innovations. In this paper, we take a formal quantitative approach to address how different pieces of innovations socialize with each other and how the interrelationships among innovations affect users' adoption behaviour, which provides a novel perspective of understanding the diffusion of innovations. Networks of innovations are constructed by mining large scale text collections in an unsupervised fashion. We are particularly interested in the following questions: what are the meaningful metrics on the network of innovations? What effects do these metrics exert on the diffusion of innovations? Do these effects vary among users with different adoption preferences or communication styles? While existing studies primarily address social influence, we provide a detailed discussion of how innovations interrelate and influence the diffusion process.

4) The early-adopter graph and its application to web-page recommendation

AUTHORS: I I. Mele, F. Bonchi, and A. Gionis

In this paper we present a novel graph-based data abstraction for modelling the browsing behaviour of web users. The objective is to identify users who discover interesting pages before others. We call these users *early adopters*. By tracking the browsing activity of early adopters we can identify new interesting pages early, and recommend these pages to similar users. We focus on news and blog pages, which are more dynamic in nature and more appropriate for recommendation. Our proposed model is called *early-adopter graph*. In this graph, nodes represent users and a directed arc between users *u* and *v* expresses the fact that *u* and *v* visit similar pages and, in particular, that user *u* tends to visit those pages *before* user *v*. The weight of the edge is the degree to which the temporal rule "*v* visits a page before *v*" holds. Based on the early-adopter graph, we build a recommendation system for news and blog pages, which outperforms other out-of-the-shelf recommendation systems based on collaborative filtering.

5) Studies of independence and conformity: I. a minority of one against a unanimous majority

AUTHORS: A. S. E

The investigations described in this series are concerned with the conditions of independence and lack of independence in the face of group pressure. The abstract temper of present-day theory and investigation in this region rests to a considerable degree on a neglect of the cognitive and emotional experiences that are part of the individual's psychological field. The understanding of social influences will require the study of a wide range of conditions and of the interrelated operations of different psychological functions. A group of seven to nine individuals was gathered in a classroom to take part in what appeared to be a simple experiment in visual discrimination. The subjects were all male, white college students, ranging in age from 17 to 25; the mean age was 20. For certain purposes a large number of critical subjects was

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2024

DOI: 10.17148/IJARCCE.2024.13590

required for the present experiment. The present report is based on a total of 123 subjects. The task consisted of the comparison of a standard line with three other lines, one of which was equal in length to the standard. We investigated some of the conditions responsible for independence and lack of independence in the face of arbitrary group pressure. To this end we produced a disagreement between a group and one individual member about a clear and simple issue of fact. The interview, which followed the experimental session, provided qualitative evidence concerning the effects produced by the majority, The particular properties of the experimental situation and their relation to more usual social contradictions were described.

III. SYSTEM DESIGN

1. The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.

2. The data flow diagram (DFD) is one of the most important modelling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.

3. DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.

4. DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.

IV. METHODOLOGY

System works as follows: -

- Admin will add products to the system.
- Admin will delete the review which is fake.
- User once access the system, user can view product and can post review about the product.
- System will track the IP address of the user.
- If the system observes fake review coming from same IP address many a times this IP address will be tracked by the system and will inform the admin to remove this review from the system.

Features: -

- Admin Login: - Admin login to the system using his admin ID and password.
- Add product: - Admin will add product to the system.
- Delete Review: - Admin will remove the review which tracked by the system as fake.
- User Login: - User will login to the system using his user ID and password.
- View product: User will view product.
- Post Review: - User can post review about the product.
- Tracks IP Address: - If the system finds a review is fake it will inform the admin to remove the fake review.

MODULES:

- Online Shopping.
- Admin.
- Users.
- Fake Review Monitoring.

MODULE DESCRIPTIONS:

Online Shopping: Online Shopping is a form of e-commerce which allows consumers to directly buy goods from a seller over the internet using a web browser. Consumers find a product of interest by visiting the retailer directly or by searching among alternative vendors using a shopping search engine. Online stores typically enable shoppers to use "search" features to find specific models, brands or items. Online customers must have access to the Internet and a valid method of payment in order to complete a transaction, such as a credit card, an Interaction enabled debit card, or a service such as PayPal.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2024

DOI: 10.17148/IJARCCE.2024.13590

Admin: In this Module, Admin can manage their website such as add/remove the products from their Website. They can also check the Orders have been ordered by Consumers and also checks whether the products are added with right specifications. The important feature of the admin is to monitor the Fake Reviews Posted on the Website about Product.

Users: In terms of Users, they can view the various products of their desired one on the shopping website and also they can search for particular brands. Users can check the specifications of the product and Order it if the product is suitable for them. Rating and reviews are posted by the pre-ordered consumers.

Fake Review Monitoring: By the Continuous monitoring of the system, the system will categorize the reviews by their IP Addresses of the consumers from the existing rating and reviews. If the consumer has posted more reviews for the same product from the same IP, then the reviews has been declared as the Fake reviews.

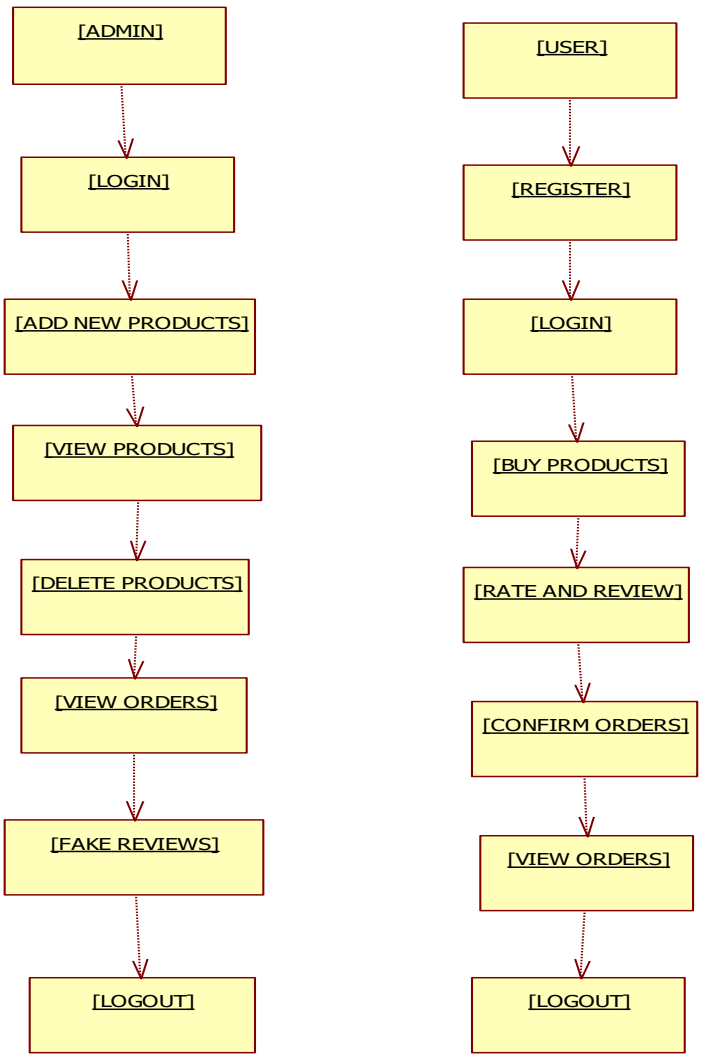


Figure. 1 Flowchart

V. USE CASE DIAGRAM

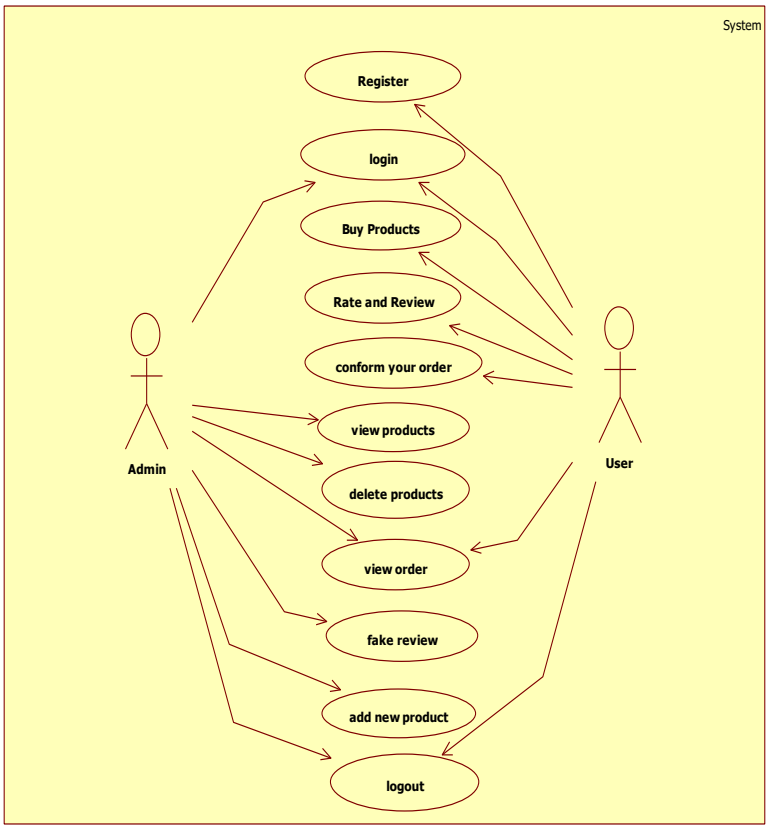
A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2024

DOI: 10.17148/IJARCCE.2024.13590



VI. RESULT

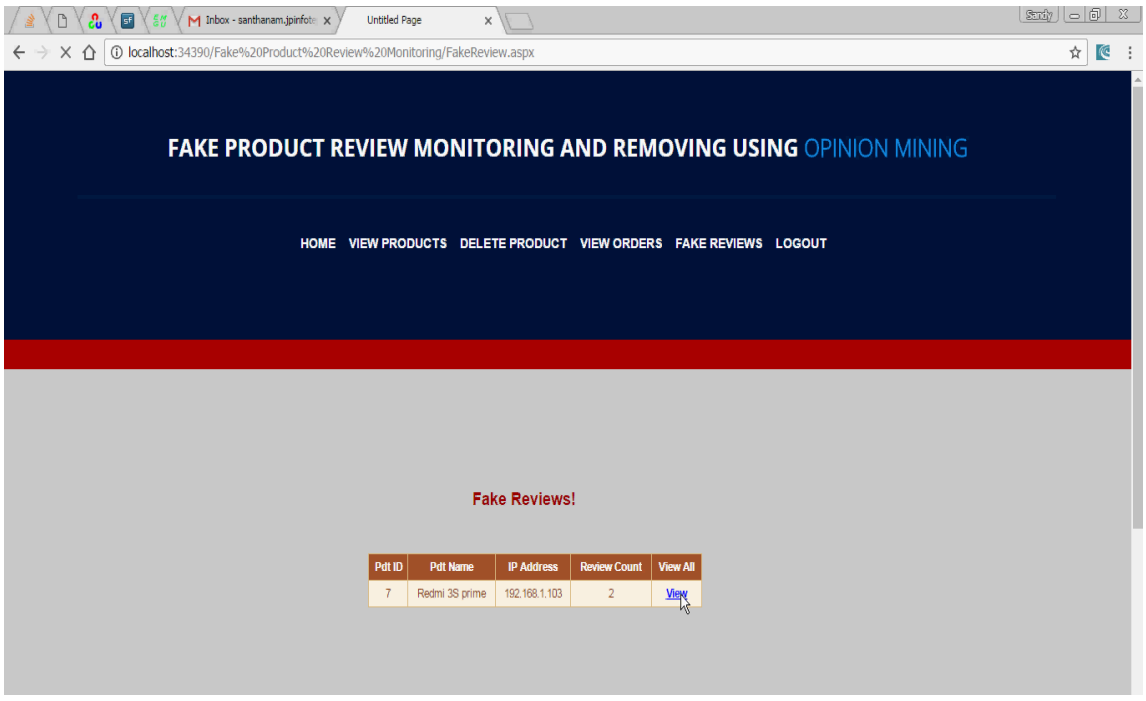


Figure. Fake review detection

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2024

DOI: 10.17148/IJARCCE.2024.13590

Above figure shows the number of fake reviews given by the users.And only admin can detect the fake reviews.

Admin will remove fake reviews.Following figure will show how the fake reviews will me removed.

Figure.Removing Fake reviews.

VII. CONCLUSION

Many schemes have demonstrated very good performance in protecting reputation systems, In this project we have implemented he fake product review monitoring. In the future, we will explore effective ways in incorporating review content into our early reviewer prediction model. Also, we have not studied the communication channel and social network structure in diffusion of innovations partly due to the difficulty in obtaining the relevant information from our review data. We will look into other sources of data such as Flipster in which social networks can be extracted and carry out more insightful analysis. Currently, we focus on the analysis and prediction of early reviewers, while there remains an important issue to address, i.e., how to improve product marketing with the identified early reviewers. We will investigate this task with real e-commerce cases in collaboration with e-commerce companies in the future.

REFERENCES

[1] J. McAuley and A. Yang, “Addressing complex and subjective product-related queries with customer reviews ,” in WWW, 2016, pp. 625–635.
[2] J. J. McAuley, C. Targett, Q. Shi, and A. van den Hengel, “Image based recommendations on styles as substitutes,” in SIGIR, 2015, pp. 43–52.
[3] X. Rong and Q. Mei, “Diffusion of innovations revisited: from social network to innovation network,” in CIKM, 2013, pp. 499–508.
[4] I. Mele, F. Bonchi, and A. Gionis, “The early-adopter graph and its application to web-page recommendation,” in CIKM, 2012, pp. 1682–1686.
[5] A. S. E, “Studies of independence and conformity: I. a Minority of one against a unanimous majority” Psychological monographs: General and applied, vol. 70(9), p. 1, 1956.
[6] V. G. D. W. Shih-Lun Tseng, Shuya Lu, “The effect of herding behaviour on online review voting participation,” in AMCIS, 2017.
[7] S. M. Mudambi and D. Schuff, “What makes a helpful online review? a study of customer reviews on amazon.com,” in MIS Quarterly, 2010, pp. 185–200.
[8] J. J. McAuley, R. Pandey, and J. Leskovec, “Inferring networks of substitutable and complementary products.” in KDD, 2015, pp. 785–794.
[9] E. Gilbert and K. Karahalios, “Understanding Deja reviewers.” In CSCW, 2010, pp. 225–228.
[10] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, “Detecting product review spammers using rating behaviours,” in CIKM, 2010, pp. 939–948.