International Journal of Research Publication and Reviews

SPAM REVIEW DETECTION USING NAÏVE BAYES CLASSIFIER

Tejas Pandit[1], Hemanshu Borikar[1], Rutvik Pathak[1], Shlok Karandikar[1], Prof. Dr Geetha Chillarge[2]

Department of - Computer Engineering, MMCOE Pune, Maharashtra, India – 411052

ABSTRACT

Online marketing is becoming increasingly prevalent as the internet's popularity grows. This is due to the fact that many items and services are widely accessible over the internet. As a result, consumer and organisational assessments of these many products and services are quite significant.

Unfortunately, scammers used to fabricate phoney reviews in order to make money or get publicity. Customers and businesses are unable to form accurate judgments regarding items due to the phoney reviews produced by scammers. As a result, to avoid deceiving potential consumers, fraudulent reviews or review spam must be recognised and eradicated. In this research, we look at how to identify review spam using supervised and semi-supervised learning techniques.

Keywords: Product Review, Spam Review, Online Marketing, Fraudsters, Machine Learning, Supervised algorithm.

1. INTRODUCTION

Machine Learning is an advanced highly intelligent system that targets to build a system that can improve via experience with the help of statistical and probabilistic data. In simple words, it is a remarkably very useful tool which helps us in predicting Spam Reviews. It also lets researchers, doctors, professors and even students get important information from raw data. This also helps to build-up personalized experiences, and further helps to make automated intelligent machines/systems. In the field of Artificial-Intelligence and Machine-Learning, the following algorithms are commonly used: SVM, decision tree, logistic regression, Naive Bayes. And all of these have been proven to predict and categorize the upcoming events.

In this paper, the main aim is to come up with a systematic literature review, analytical review, and abstract of the ML techniques that are used for predictions of Spam Reviews. This way, it'll also give rise to the challenges and new limitations of using ML techniques in this field. Apart from these things, many good chances and gaps in this area for future research can also be discussed. This paper therefore came up with an abstract summary and potential research method that could guide researchers, doctors, professors and even students to obtain knowledge about the methods and use of ML in spam review detection.

2. LITERATURE SURVEY

Sr. No	Title	Publication	Year	Advantages	Disadvantages
1	Fake review Detection using Machine Learning Techniques	IEEE Shilpa yadav, et al.	2021	A CountVectorizer, Ngram model, and TfidfVectorizer were used as machine learning algorithms. The random forest algorithm outperforms other algorithms.	Worked on a tiny real-world yelp dataset
2	Supervised Machine Learning for Fake Reviews Detection	ResearchGate Ahmed Mohamed, et al.	2021	Study of the relevance of reviews and how they affect practically all aspects of web-based data	118 Sample Spaces aren't enough.
3	An ensemble approach to detect review spam using hybrid ML technique	ResearchGate Istiaq Ahsan, et al.	2016	This study provides a strategy for detecting Review spam that combines the two machine learning techniques (Active learning and Supervised learning).	Accuracy of NB classifier 5-fold model is low as compared to its precision value.

4	Learning to Identify Review Spam	IEEE Fangtao Li, et al	2018	This research investigates the task of detecting review spam in a product review mining system in depth.	The Deep learning approaches like recursive neural networks can help enhance categorization accuracy.
5	Mining and Summarizing Customer Reviews	IEEE Minqing Hu, et al.	2016	Analysing different techniques for mining spam reviews	The recall and precision accuracy for sentence extraction and sentence orientation prediction is comparatively low
6	Product Review Analysis and Spam Review Detection	IEEE Shashank Chavhan, et al.	2017	Analysis of how to make a dictionary which has sentiment words and weight assigned to them according to its property	Proper class assignment is not guaranteed

3. PROPOSED SYSTEM

In this paper, we are proposing a system which will predict spam review and also factors responsible for it. So basically, here, we find the reviewer's excessive rating ratio to be an intriguing feature. Because a fake reviewer will always offer either a 1 or a 5-star rating to persuade others of their point of view, we estimated the extreme rate (1star or 5stars) ratio for each reviewer and used it as one of the review's features. The extreme rating ratio for all unique reviewers (1 or 5). Divide the total number of reviews by the number of extreme evaluations given by the reviser. We generated this value for each unique reviewer and entered it into their view, which was then reviewed by the relevant reviewers.

There has never been a good way to tell the difference between phoney and genuine products. To address such issues, machine learning technologies can be useful. The project's main purpose is users' assistance in determining if a product is genuine or counterfeit based on its reviews. We suggested a web-based tool for the detection of counterfeit products that uses machine learning technologies. The proposed approach ensures that bogus products are detected in everyday life. Customer or user web-based application, Manufacturer or firm web application, and Database are the three key components of the proposed system. The first application, which is the Manufacturers or Company Side Application, requires us to first register. We have some alternatives when logging into the application after registration. One alternative is to add a product and let the maker fill in the data. Another option is to show the order, which allows them to examine the customer's order information and then accept or reject the order. The manufacturer can also see whether or not the goods has been delivered. A second application is the Customer application, which requires us to first register in the app before logging in with our id and password. Customers can see product details such as name, total amount, price of a product, and manufacturer details in this application.

This System is represented by 5-different phases, all of these phases working with their own dependency System which is $S = (Q, \Sigma, \delta, q0, F)$ where,

- Q is a set of finite states.

- Σ here defines as a finite set of symbols called as alphabet.

- Δ is a transition function, where $\delta : (Q \times \Sigma) \rightarrow Q$

- q0 is the starting point for processing any input (q0 ε Q).

- F is a collection of Q's final state/states. (F Q).

The similarity weight of fitness function of specific rules is generated by all t(n) policies returning 1 from training patterns.

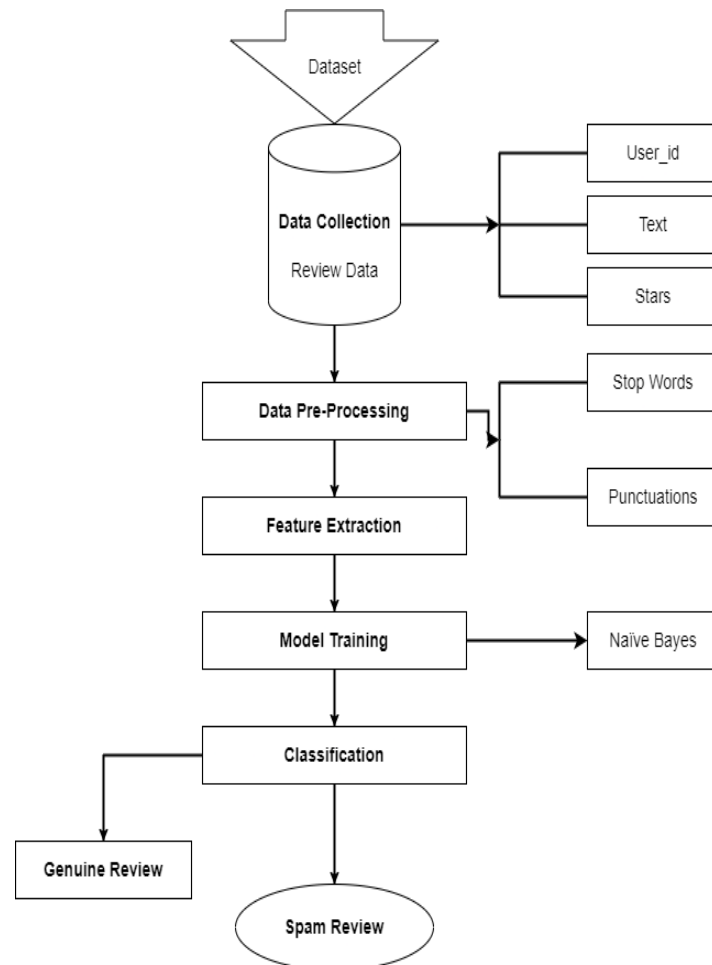Q = {VSet[i=0. . ... n]} is the initial set of produced qualities from various reviews

Σ = {saveinDB, data conversion}

Δ = { CorrectlyClassifiedInstnaces*100 / SumofF(x)}

q0 = {Sensor functions first event generation Σ i=0 }

F = {Report generated according to class [a,b,c,...., n]}

The system architecture is the model that conceptually defines the views, structure, and behavior of the system. System architecture in other words is the representation and description of how the system works and communicates with other system components in general. The diagrammatic representation of the system architecture is called the system architecture diagram. This diagram gives us the abstract view of the components and their relationship with the system that makes the system work.

Fig: System Architecture Diagram

Procedure to classify a review -

On developer end -

1. Obtain the csv file of the data.

2. Extract only the required features from it. In this case, we are only extracting the parameters "deceptive" and "text". The deceptive parameter contains the deceptiveness of the particular review. A review can either be truthful or it can be deceptive. The text parameter contains the actual review.
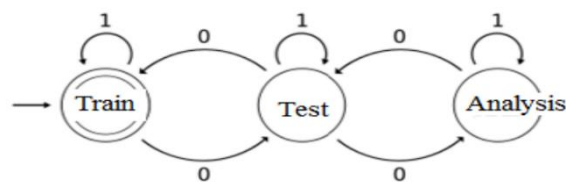
3. Assign 0 to categorical variable deceptive and 1 to categorical variable truthful.

4. Print the dataframe and separate the input and the output features. While the text portion forms the input, the deceptive parameter constitutes the output portion of the dataframe.

5. Import respective Multinomial Naïve Bayes classifier. Also import Gaussian Naïve Bayes classifier.

6. Split the dataset that we earlier obtained into two parts, one each for training and testing. Take 70% of the dataset for training and 30% for testing.

7. The text parts need to be converted into numerical feature. Here, we are using Count Vectorizer to facilitate this.

8. The data needs to be fitted for modelling. This step is known as Data Fitting.

9. The data is processed and we get 2 types of accuracies.

10. The Training accuracy is the percentage of the number of reviews where our model predicted correctly, the deceptiveness of a review when compared to the total number of reviews in our training dataset.

11. While the Testing accuracy is the percentage of the number of reviews where the prediction of our model matched the actual deceptiveness of a particular review, for reviews from the testing dataset.

On user end -

1. Goto the (locally hosted) url of our project in a web browser like Chrome.

2. The user should see a login page. Enter the username and the password in the placeholders for the same. Here, we have set the username as "admin" and the password as "admin".

3. Once logged in, the user will see a section for him to input his particular review which he wants to test the deceptiveness of. The user should enter the review here.

4. The user needs to press the review button located below this box.

5. The output of the screen displays whether the review is spam, or not.

6. If the user instead wants to test the accuracy of a bulk dataset of reviews, they should upload its .csv or .parquet file. The user will then see a tabular representation mentioning the accuracy of individual reviews. Here, we are displaying any 10 reviews as an example.

7. If the user wants to know the correctness of our model, he should navigate to the "Accuracy" pane in the menu.

8. The user will be able to get the Training and Testing accuracy of our model in this section.

9. The user can continue testing reviews in either ways mentioned as long as they want to.
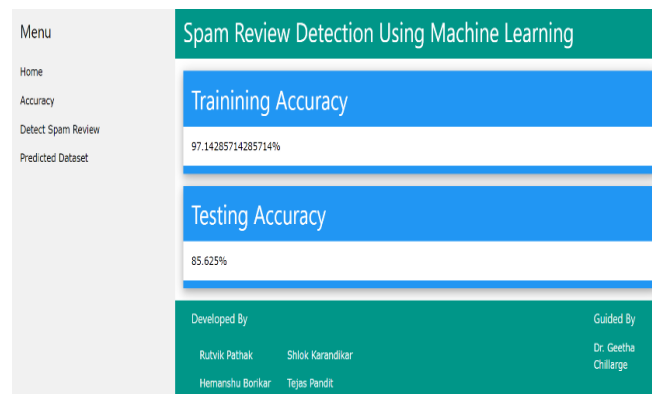
4.   RESULT

We have created a web application bases on our research which uses Naïve-Bays algorithm in ML which uses the input dataset that we have provided and tarins the model to predict whether given user input review is a spam or not.
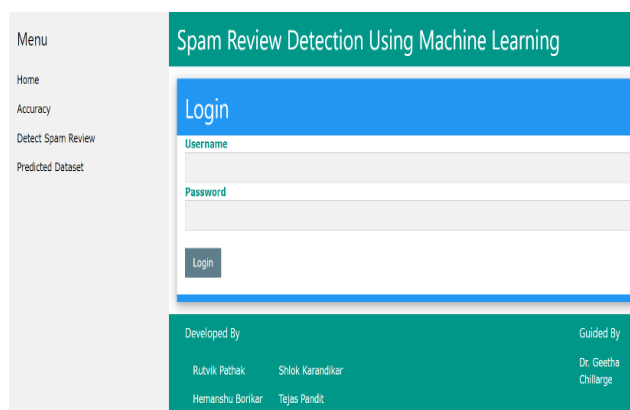
This is the login page where the admin can login to do all the operations available on dashboard.

Outputs :

Here we use training data that we provide to train the model.

We can take user review as an input and determine what the possibility of that review being spam is.

We can check any single review, as we check the review, we find that the possibility of being spam here is high. This is determined based on our training dataset.

We can detect a bulk dataset using this model as well. Here we uploaded a CSV file containing total of 1600 reviews which are unlabelled. We use our model to detect whether these reviews are spam or not. And we export a CSV file which has all the reviews classified and labelled as Spam or not.

Here we show some examples from among the CSV file which is labelled as spam or not spam.

5. CONCLUSION

Four common machine learning classification approaches for detecting false product reviews have been examined by this system.

Non-filtered reviews receive ratings such as informative, cool, and humorous, which means that once the reviews are filtered by product, their view is hidden and cannot be generated by others. When employing NLP, the experiment results indicated a very high prediction score. Many features, such as user trust based on user friendship and user profile (join date, photo, etc.), are still unavailable due to dataset limitations. Because an imbalanced dataset produces bad results in our experiment, it must be addressed.

During the experiment, we discovered that SVM took the longest to train the model and that Gaussian Naive Bayes produced the lowest average score. In our opinion, we cannot declare that all of the reviews filtered by the YELP recommendation system are phoney, because there are still other criteria

that can drive machine learning to make incorrect predictions. Other possibly reliable and selective review systems include the verified buyer method, which has been employed in various crowd source webs.

REFERENCES

[1] Hu Minqing, and Liu Bing. "Mining and summarizing customer reviews", Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press , 2004:168-177

[2] Kim S-M and HOVY E," Determining the sentiment of opinions", Proceedings of the 20th International Conference on Computational Linguistics. Philadelphia: Association for Computational Linguistics,2004:1367

[3] Ohana B and Tierney B, " Sentiment classification of reviews using SentiWordNet" 9th. IT & T Conference. 2009

[4] Li N and Wu D D, " Using text mining and sentiment analysis for online forums hotspot detection and forecast" Decision Support Systems, 2010, 48(2): 354-368

[5] Mudinas A, Zhang Dand Levene M, "Combining lexicon and learning based approaches for concept-level sentiment analysis", Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining. ACM, 2012: 5.

[6] Jindal Nitin, LIU Bing. "Review spam detection", Proceedings of the 16th international conference on World Wide Web. Canada. New York: ACM Press ,2007: 1189- 1190.

[7] Jindal Nitin and Liu Bing, 'Opinion spam and analysis", Proceedings of the 2008 International Conference on Web Search and Data Mining. New York: ACM Press , 2008:219-230.

[8] XieSihong, WANG Guan, LIN Shuyang, et al. " Review spam detection via temporal pattern discovery" , Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery And Data Mining. New York: ACM Press , 2012:823-831.

[9] Lim Ee-Peng, Nguyen Viet-An, Jindal Nitin, et al. Detecting product review spammers using rating behaviors [C]// Proceedings of the 19th ACM international conference on Information and knowledge management. New York: ACM Press , 2010:939-948.

[10] Jindal Nitin, Liu Bing, Lim Ee-peng, et al. " Finding unusual review patterns using unexpected rules", Proceedings of the 19th ACM International Conference on Information and Knowledge Management. New York: ACM Press -2010 ; pp1549-1552

[11] Wang Guan, XieSihong, Liu Bing,et al. " Identify online store review spammers via social review graph" , ACM Transactions on Intelligent Systems and Technology, 2011,3(4):61.1-61.21.

[12] MyleOtt, Yejin Choi, Claire Cardie, et al. Hancock. "Finding deceptive opinion spam by any stretch of the imagination" Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, 2011: 309-319

[13] Zhang, Rong, et al. "Exploiting shopping and reviewing behavior to re-score online evaluations" Proceedings of the 21st international conference companion on World Wide Web. ACM, 2012.

[14] QingxiPeng and,MingZhong- "Detecting Spam Review through Sentiment Analysis", State Key Lab of Software Engineering, Wuhan University, Wuhan, China

[15] E. D. Wahyuni and A. Djunaidy, "Fake review detection from a product review using modified method of iterative computation framework." 2016.

[16] D. Michie, D. J. Spiegelhalter, C. Taylor et al., "Machine learning," Neural and Statistical Classification, vol. 13, 1994.

[17] T. O. Ayodele, "Types of machine learning algorithms," in New advances in machine learning. InTech, 2010.

[18] F. Sebastiani, "Machine learning in automated text categorization," ACM computing surveys (CSUR), vol. 34, no. 1, pp. 1–47, 2002.

[19] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features." 1998.

[20] T. R. Patil and S. S. Sherekar, "Performance analysis of naive bayes and j48 classification algorithm for data classification," pp. 256–261, 2013.

[21] M.-L. Zhang and Z.-H. Zhou, "Ml-knn: A lazy learning approach to multi-label learning," Pattern recognition, vol. 40, no. 7, pp. 2038–2048, 2007.

[22] N. Suguna and K. Thanushkodi, "An improved k-nearest neighbor classification using genetic algorithm," International Journal of Computer Science Issues, vol. 7, no. 2, pp. 18–21, 2010.

[23] M. A. Friedl and C. E. Brodley, "Decision tree classification of land cover from remotely sensed data," Remote sensing of environment, vol. 61, no. 3, pp. 399–409, 1997.

[24] A. Liaw, M. Wiener et al., "Classification and regression by randomforest," R news, vol. 2, no. 3, pp. 18–22, 2002.

[25] D. G. Kleinbaum, K. Dietz, M. Gail, M. Klein, and M. Klein, Logistic regression. Springer, 2002.