International Journal of Computer Engineering in Research Trends

Multidisciplinary, Open Access, Peer-Reviewed and fully refereed

Research Paper Volume-7, Issue-6,2020, Regular Edition E-ISSN: 2349-7084

Semantic & Behavioral Feature analysis for Detecting Fake Reviews using Machine Learning

[1]Anjali S Dilliwala, [2]Raghavendra G S

[1]Student, Dept. Of CSE, Shri Dharmasthala Manjunatheshwara College of Engineering, Dharwad, India.*
[2]Professor, Dept. Of CSE, Shri Dharmasthala Manjunatheshwara College of Engineering, Dharwad, India.

e-mail: *anjalidilliwala06@gmail.com*, *raghu_git123@rediffmail.com*

Available online at: http://www.ijcert.org

Received:27/06/2020 Revised:28/06/2020 Accepted:05/07/2020 Published: 14/07/2020

Abstract: Background: In this age of technology, online business is playing a vital role in the growth of the economy of the business. Hence, people need feedback on various products, technologies, businesses, etc. Since their opinions are the input for an individual to evaluate and adapt them. Therefore, the Review system is playing a vital role in decision making. So there arises a necessity to evaluate the reviews as well since the business units are trying to generate fake reviews to identify more clients for their products.

Methods/ Statistical Analysis: In this paper, we implement two machine learning algorithms SVM and Naïve Bayes algorithm and analyse the data and predict for the new set of data. We also compare the performance of both algorithms.

Findings: In this paper, we are trying to develop a Machine learning model which analyses the reviews on various factors and obtain the necessary features and classify the reviews as a fake or non-fake review. This helps in identifying fraudulent reviews and predicts the trustworthiness of the reviews in the future.

Improvements: The system can introduce and make available Machine learning techniques and identifying fake reviews at the earliest stage.

Keywords: Reviews, Machine learning, SVM, Naive Bayes, Review system.

1. Introduction

An increasing quantity of our lives is spent interacting online through social media platforms, additional and additional individuals tend to hunt out and consume news from social media instead of ancient news organizations. The explanations for this alteration in

Consumption behaviours are inherent in these social media platforms: It is commonly additional timely and fewer big-ticket to consume news on social media compared with ancient print media, like newspapers or television; and

It is simpler to more share, comment on, and discuss the news with friends or alternative readers on social media. As an example, 62% of U.S. adults get news on social media in 2016, whereas in 2012, solely 49% according to seeing news on social media.

With this, the trend of online shopping is also increasing, and a lot many people end up buying the essential commodity online. This option of buying commodities is more practiced because people do not have to go out of their home to buy anything, it doesn't take a lot of, and it's simpler with just a few clicks by the blessings of the Internet.

However, because it is easy and straightforward to shop online, customers believe in false opinions from fake reviews. If they don't pay proper attention to examining the product and related reviews, they end up getting cheated. Since one can't examine the quality in online stores, it all depends on the reviews about the product. Therefore, if there is a system that provides the consumer to find genuine reviews and respective ratings for a product, then the system of online shopping can be reliable. Fake review analysis depends on behavioral features, linguistic and textual features, and relational features. This is shown in Figure1.1.

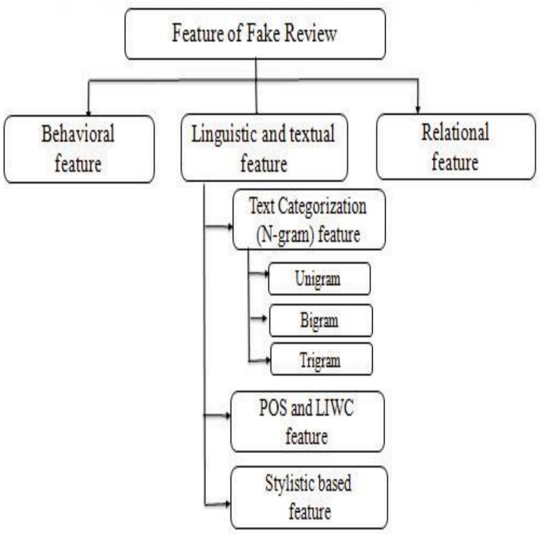

Figure 1.1 Types of Fake review Features.

Therefore, the Review system is playing a vital role in decision making. So there arises a necessity to evaluate the reviews as well since the business units are trying to generate fake reviews to identify more clients for their products. Hence, in this work, we are trying to develop a Machine learning model that analysis the reviews on various factors and obtains the necessary feature analysis and classifies the reviews as a fake or non-fake review.

The rest of this paper is organized as follows. Section II presents the background analysis. Section III shows the Objectives. Section IV explains the Proposed work; Section V explains the experiment results, Section VI presents the conclusion, and finally, Section VII future works.

2. Background

Jindal et al. initially researched the system of fake review detection.[2]. Among the various ways to identify fake reviews, the machine learning technique is one among them. In machine learning techniques, algorithms such as SVM is one of the robust classifier approaches, which also represents it as an excellent prediction method [1].

In [5], the authors used supervised Machine learning techniques for identifying fake reviews. They conjointly perform different pre-processing steps before the classification technique is applied, which embody stemming, removal of punctuation marks, and stop word removal. The linguistic feature is employed to spot faux reviews, contains POS and bag-of-words. Then to get the detailed result, different classification algorithms like call tree, random forest, support vector machine, Naive Bayes, and gradient boosted trees square measure applied.

Ott et al. [3], made the primary dataset of gold-standard deceptive opinion spam, using crowdsourcing through the Amazon Mechanical Turki. The authors found that though part-of-speech n-gram options provide a smart prediction on whether or not a private review is faux or not. The classifier indeed performed slightly higher once cognitive psychology options were added to the model.

In [4], Xie et al. propose a system that uses a net crawler to scrap the information on the web site, that successively focuses on police investigation spam and pretend reviews by victimization sentimental analysis. It removes the reviews that have the curse and vulgar words. Within the pre-processing, the information is reborn into the specified format. So the faux reviews square measure far away from the mixture of original and spam reviews victimization the projected model.

3. Objectives

Many analysis works had been created in ancient ways, and still, researchers try to boost the extent of sleuthing spam review a lot of exactly. During this paper, a system is planned supported 2 machine learning algorithms SVM and Naïve Bayes that helps in serving the aim of sleuthing pretend reviews during a higher economic approach. They conjointly compare between the results of each the algorithms in distinguishing the most effective among the 2.

Thus, by implementing the best available approach for detection of fake reviews using opinion mining (sentiment analysis) techniques provide necessary information to customers to know if each review is trustworthy and reliable.

4. Proposed Work

The proposed system is divided into 5 phases as named below,

1. Data Acquisition.
2. Text pre-processing.
3. Feature Vectorization.
4. Sentiment data analysis.
5. Machine learning prediction.
   a. Linear SVM
   b. Naïve Bayes

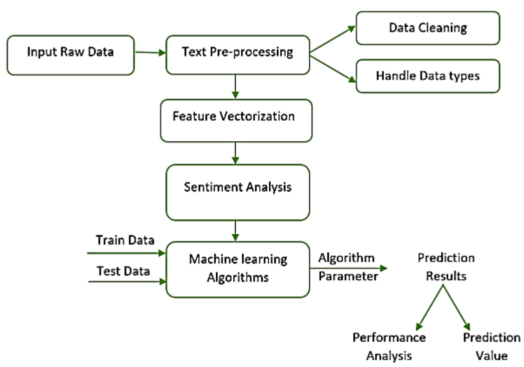

Figure 4.1 Proposed Systems

Step 1: **Data Acquisition**

Dataset is prepared by collecting the reviews of the web users on the product from the website. There are a handful of datasets that contain good quality genuine reviews and the deceiving ones. After previous inquiring works supported the references mentioned, we've got learned that one tagged and one unlabelled dataset was principally used. It was created by Ott et al. [3] and additionally mostly referred to as the Ott dataset. On the opposite hand, real-life reviews from Yelp dataset that are publicly accessible and may be used because of the unlabelled dataset [10]. A sample of the no inheritable set of information and their attribute are shown within the figure 4.2. below

	DOC_ID	LABEL	RATING	VERIFIED_PURCHASE	PRODUCT_CATEGORY	PRODUCT_ID	PRODUCT_TITLE	REVIEW_TITLE	REVIEW_TEXT
0	1	__label1__	4	N	PC	B00008NG7N	Targus PAUK10U Ultra Mini USB Keypad, Black	useful	When least you think so, this product will sav...
1	2	__label1__	4	Y	Wireless	B00LH0Y3NM	Note 3 Battery : Stalion Strength Replacement ...	New era for batteries	Lithium batteries are something new introduced...
2	3	__label1__	3	N	Baby	B000I5UZ1Q	Fisher-Price Papasan Cradle Swing, Starlight	doesn't swing very well.	I purchased this swing for my baby. She is 6 m...
3	4	__label1__	4	N	Office Products	B003822IRA	Casio MS-80B Standard Function Desktop Calculator	Great computing!	I was looking for an inexpensive desk calcolat...
4	5	__label1__	4	N	Beauty	B00PWSAXAM	Shine Whitening - Zero Peroxide Teeth Whitenin...	Only use twice a week	I only use it twice a week and the results are...

Figure 4.2 Types of Fake review Features.

We have a dataset with 80000 records, which have 9 attributes respectively collected from the major e-commerce sites like Amazon, Flipkart, eBay, etc.

Step 2: Text Pre-processing

Text pre-processing is a technique that is employed to refine or clean data by applying the little necessary and needed transformation. This section is disbursed before feeding the info to the algorithmic rule; therefore on get results on a reliable and valid set of information for our purpose. To be explaining it in straightforward terms, the info that is gathered from totally different sources isn't possible for the analysis and so must be treated consequently to realize higher results from the applied machine learning model. Few specific Machine Learning model wants data in an exceedingly specific format, as an example, Random Forest algorithmic rule doesn't support null values, so all null values have to be compelled to be consequently reorganized or changed from the obtained data in keeping with the need of the model. Also, since we have a tendency to create a comparison between the results of 2 algorithms, the info ought to be consequently managed that it's appropriate to be employed by each the algorithms utilized in the analysis.

Step 3: Feature Engineering

- We have the majority standard reviews, so we use oversampling to handle biased dataset.
- Use 80-20 train, test split
- Tokenize using Spacy and Keras to split the review into words
- Remove punctuations, convert to lowercase.
- Remove stop words using Spacy word-labels ("POS", "stop word", etc.)
- Vectorizer to convert word to index, TF-IDF to scale using inverse document frequency.
- Plot word frequency graph and word cloud.

Step 4: Sentiment Data analysis

Sentiment information analysis (EDA) is an indispensable advance that happens once part coming up with and getting data, and it needs to be done before any demonstrating. This can be because it's essential for Associate in nursing data man of science to actually nearly comprehend the thought of the data while not creating suspicions. The after-effects of data investigation will be useful in obtaining a handle on the structure of the data, the appropriation of the qualities, and the distance of extraordinary qualities and interrelationships within the informational index.

I. Label v/s Product Category

The below figure 4.1 represents the product category labels for both legitimate and fake review against the number of their review occurrences in each of the label.

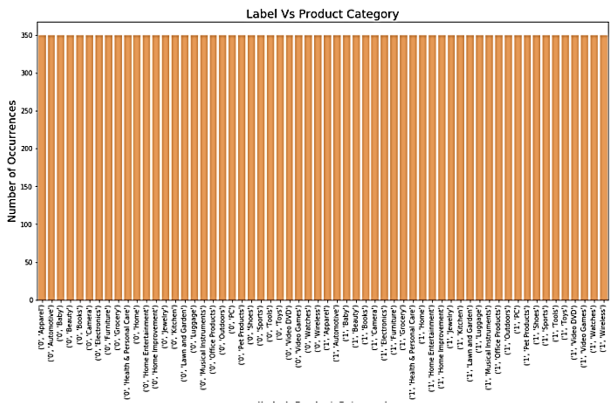

Figure 4.1 Label v/s Product Category

II. Label v/s Rating

The below Figure 4.2 shows the distribution of Ratings ranging from 1 to 5 for both fake and legitimate reviews.

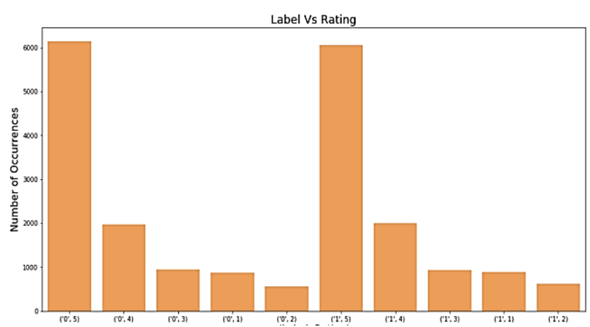

Figure 4.2 Label v/s Rating

III. Verified Purchase v/s Labels

The below figure 4.3 represents the graph of the products which are verified purchased for both the kinds of review against the count of their occurrences.

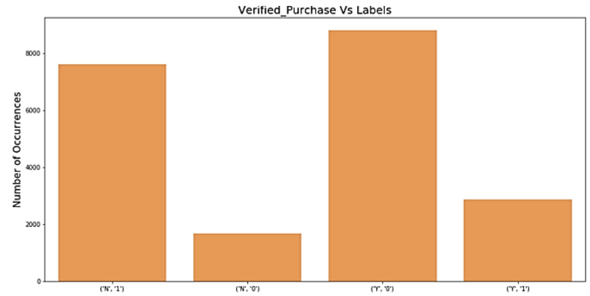

Figure 4.3 Verified Purchase v/s Rating

IV. Machine learning Prediction

Machine learning prediction has these following steps:

1. Divide data into 2 parts: training and testing data.
2. We are defining the algorithms, namely Linear SVC, and the Naïve Bayes algorithm.
3. Training and testing against both the algorithms.
4. Comparing the performance for both the algorithms.

Machine learning Algorithms used are:

1. Linear SVC algorithm
2. Naïve Bayes algorithm.

Linear SVM Algorithm

Support vector machines (SVMs) [7] learning formula are accustomed to building the prognosticative model. SVMs area unit one in every of the foremost standard classification algorithms, and have a chic method of reworking nonlinear knowledge so that one will use a linear formula to suit a linear model to the info (Cortes and Vapnik k 1995).

- SVMs take into consideration complex choice limits, regardless of whether the information has just a couple of highlights.
- They function admirably on low-dimensional and high-dimensional information (i.e., few and numerous highlights), however, don't scale high with the number of tests.
- Running an SVM on information with up to l0, 000examples may function admirably, yet working with datasets of size at least 100,000can wind up testing regarding runtime and memory use.
- SVMs require cautious pre-processing of the information and tuning of the parameters. In this way, these days the vast majority incline toward tree-based models.

Naive Bayes Algorithm

Naive Bayes could be a natural model that might be used for text classification. Naive Bayes, additional technically noted because the Posterior chance updates the previous belief of an incident given new data. The result's the chance of the category occurring given the new information. The classification model may handle binary and multiple classifications. When predicting a category, the model calculates the posterior chance for all categories and selects the most important posterior chance because the foretold category. This value is referred to as the Maximum A Posterior (MAP).

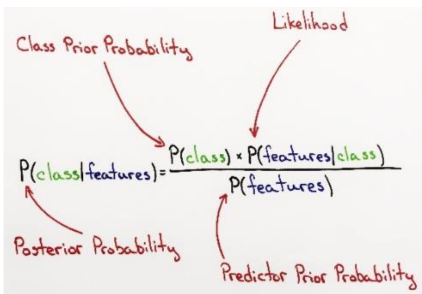

Figure 4.4 Bayes theorem

Posterior Probability:

- This is the updated belief given the new data, and the objective Probability of each class, derived from the Naive Bayes technique.

Class Prior Probability:

This is the Prior Belief; the Probability of the class before updating the belief.

Likelihood:

The likelihood is calculated by taking the product of all Normal Probability Density Functions (assume independence, ergo the "Naivete"). The Normal PDF is calculated using the Gaussian Distribution. Hence, the name Gauss Naive Bayes.

- We will use the Normal PDF to calculate the Normal Probability value for each feature given the class.
- The likelihood is the product of all Normal PDFs.

There's an important distinction to keep in mind between Likelihood and Probability.

- Average Probability is calculated for each feature given the class and is always between 0 and 1.
- The likelihood is the product of all Normal Probability values.
- The number of features is infinite and limited to our imagination.
- Since there will always be features that could be added, the product of all Normal Probabilities is not the Probability but the Likelihood.

Predictor Prior Probability:

- Predictor Prior Probability is another way of saying Marginal Probability.
- It is the Probability given the new data under all possible features for each class.
- It isn't necessary for the Naive Bayes Classifier to calculate this, because we're only looking for the prediction and not the exact Probability.
- The results do not change at all, however we do calculate it here to show that this is the case.

5. Evaluation

- There are two possible outcome classes:" 0" and "1". 0 = Legitimate, 1= Fake review.
- The original data of 1000 records is evaluated, in which 850 records were legitimate review and 150 fake reviews.
- Out of those 1000 records, Naive Bayes predicted 790 records as legitimate and 210 records as fake reviews.
- Similarly, the Linear SVM algorithm predicted 689 records as legitimate and 311 records as fake reviews.

In the evaluation, we want to understand, for several metrics, whether our method works well for the problem statement we are trying tackle. Figure 5.1 shows the graph of the accuracy of both the algorithms used.

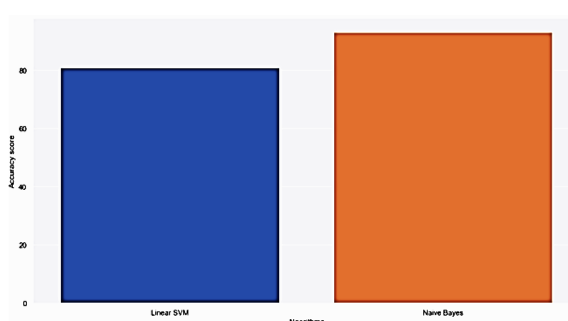

Figure 5.1 Accuracy Comparison

From the outputs obtained, it is a notable result obtained with the Naive Bayes algorithm giving a good result. Also, comparing the algorithms on various metrics, we obtain the following results, as seen in table 1.

Table 1: Metrics Comparison

Algorithms / **Metrics**	**Linear SVM**	**Naive Bayes**
Accuracy	80.93	93
Precision	0.819	0.84
Recall	0.8093	0.83
F-score	0.80	0.80

6. Conclusion

Due to fast development of the web, the scale of the reviews of the things/merchandise will increase. These immense amounts of data are generated on the Internet; there's no analysis of the quality of reviews that are written by the client. Anyone will write something that once and for all results in pretends reviews, or some firms are hiring individuals to post reviews. Several the pretend reviews that are advisedly unreal to look real, the capability to spot pretend online reviews are crucial. During this paper, we've got mentioned completely different pretend reviews detection techniques that are supported unattended, supervised additionally as semi-supervised methodologies. During this paper, we've got seen completely different options very well, like linguistic options, behavioural and relative options. We've got additionally compared entirely different techniques to spot pretend reviews. We've got additionally mentioned significant challenges of pretend review detection.

7. Future Scope

This paper majorly focuses on using supervised learning techniques to predict the trustworthiness of the reviews. Hence as future work, the plan is to work on unsupervised learning techniques where the data need not be training while reducing the complexity.

References

[1] Esposito, Gennaro. LP-type methods for Optimal Transductive SVMs. Vol. 3. Gennaro Esposito, PhD, 2014.

[2] N. Jindal and B. Liu, "Review spam detection", Proceedings of the 16th international conference on World Wide Web - WWW 07 (2007), ACM, pp. 1189–1190, 2007

[3] Myle Ott, Yejin Choi, Claire Cardie, and Jeffrey T. 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, HLT, 2011.

[4] S. Xie, G. Wang, S. Lin, and P. S. Yu, "Review spam detection via temporal pattern discovery," p. 823, 2012.

[5] W. Etaiwi,G.Naymat, "The impact of applying pre-processing steps on review spam detection", The 8th international conference on emerging ubiquitous system and pervasion networks, Elsevier, pp. 273-279, 2017.

[6] P.Rosso, D.Cabrera, M. Gomez, "Detecting positive and negative deceptive opinions using PU-learning", Elsevier, pp.1-11, 2014.

[7] S. Feng, L. Xing, A Gogar and Y. Choi, "Distributional footprints of deceptive product reviews". In Proceedings of the 6th International AAAI Conference on Weblogs and Social Media (ICWSM), pp. 98- 105, 2012.

[8] M. I. Ahsan, T. Nahian, A. A. Kafi, M. I. Hossain, and F. M. Shah, "Review spam detection using active learning," 2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2016.

[9] M. Singh, L. Kumar, and S. Sinha, "Model for Detecting Fake or Spam Reviews," Advances in Intelligent Systems and Computing ICT Based Innovations, pp. 213–217, Jan. 2017.

Authors Profile

Ms.Anjali S Dilliwala, Student, M.Tech, Dept. Of CSE, Shri Dharmasthala Manjunatheshwara College of Engineering, Dharwad, India.

Dr.Raghavendra G S, Associate Professor, Dept. Of CSE, Shri Dharmasthala Manjunatheshwara College of Engineering, Dharwad, India.

