



International Journal for Innovative Engineering and Management Research

A Peer Reviewed Open Access International Journal

www.ijiemr.org

COPY RIGHT



ELSEVIER
SSRN

2022IJIEMR. Personal use of this material is permitted. Permission from IJIEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper, all copy right is authenticated to Paper Authors

IJIEMR Transactions, online available on 06th Dec 2022. Link

[:http://www.ijiemr.org/downloads.php?vol=Volume-11&issue=ISSUE-12](http://www.ijiemr.org/downloads.php?vol=Volume-11&issue=ISSUE-12)

DOI: 10.48047/IJIEMR/V11/ISSUE 12/53

Title **FLIPKART REVIEWS SENTIMENT ANALYSIS USING SENTIMENT INTENSITY ANALYSIS USING NATURAL LANGUAGE PROCESSING**

Volume 11, Issue 12, Pages: 410-416

Paper Authors

Dr. PEDAPUDI SAMBASIVA RAO , POTHURAJU SWATHI , AMARTHALURI SURESH BABU



USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per **UGC Guidelines** We Are Providing A Electronic Bar Code

FLIPKART REVIEWS SENTIMENT ANALYSIS USING SENTIMENT INTENSITY ANALYSIS USING NATURAL LANGUAGE PROCESSING

Dr. PEDAPUDI SAMBASIVA RAO¹, POTHURAJU SWATHI², AMARTHALURI SURESH BABU³

Associate professor¹, Assistant Professor², Assistant Professor³

CSE Department, Sri Mittapalli College of Engineering, Guntur, Andhra Pradesh-522233

ABSTRACT— Now a days online shopping become a daily activity for humans. Before going to buy any product in e-commerce business organization like flipkart, amazon, etc. Customer checks the reliability of a product. Reviews is the one of the important way to check reliability of a product. Customer will check reviews posted by other customers to buy a product. In reviews there is a positive and negative reviews as well as fake reviews. If a customer bought a product by seeing fake review, if the product is really good no problem otherwise a product loses its reliability. We are here to perform sentiment analysis on restaurant reviews to find number of correct and number of wrong predictions made by the classifier which is further helpful to classify reviews into real or fake. The classifiers used in our project is Support vector Machine (SVM), Random Forest and Naïve Bayes. The measured results of our experiments show that the SVM algorithm outperforms other algorithms, and that it reaches the highest accuracy not only in text classification but also in detecting fake reviews.

Keywords— Supervised Machine Learning Techniques, Support Vector Machine, Random Forest and Naïve Bayes.

1. INTRODUCTION

Online purchasing is rising bit by bit since each service or product is easily accessible. Sellers are obtaining more reaction to one's corporation. factors. Several people generally frustrated kinds of persons misdirect others by sharing false comments to encourage or damage the image of any specific

goods or services according to wish. Such people are known as perception spammers and the false reviews they give are considered as fake comments. Although customer reviews could be beneficial, naive confidence in such comments is unsafe for either the buyers or sellers. Many consumers read research before making any online purchase. Moreover, the comments

could be misleading for additional benefit or profit, so any buying decision relied on web comments should be taken carefully. Our work is mainly directed to SA at the document level, more specifically, on movie reviews dataset. Machine learning techniques and SA methods are expected to have a major positive effect, especially for the detection processes of fake reviews in restaurant reviews, e-commerce, social commerce environments, and other domains. In machine learning-based techniques, algorithms such as SVM, NB, and Random forest are applied for classification purposes. SVM is a type of learning algorithm that represents supervised machine learning approaches and it is an excellent successful prediction approach. The SVM is also a robust classification approach. The main goal of our study is to classify restaurant reviews as a real reviews or fake reviews using SA algorithms with supervised learning techniques.

2. RELATED WORK

Kolli Shivagangadhar intended at determining if a review is fake or real. The classifiers used in this work are Naïve Bayes Classifier, Logistic

Regression and SVM. Consumers are much more reliant on taking decisions to purchase products at either ecommerce websites or offline retail shops in the current context. As such comments are dangerous for an item's strength or weakness in profits, comments are exploited for true or false views. Distorted comments could also be called fake or dishonest comments or misleading reviews or untrue feedbacks. Phishing on misleading sentiment in recent digital world is becoming a risk to both consumers and businesses. It is a critical and challenging job to discern such fake comments. Such disappointed users also get hired to advertise such comments. As a consequence, by taking a look at every review, it's a huge challenge for a normal consumer to distinguish dishonest comments from honest one. In the paper published with title, it explains about the classification of reviews using sentimental analysis algorithms. The following section gives a system outline of the proposed system. The system architecture is the conceptual model that defines the structural and behavioural representation of a proposed system. The system architecture is an outcome

of the design process. The proposed system architecture aims at the design and development of an effective system for fake review detection.

3. METHODOLOGY

We Used Python Jupyter notebook software to write the working code. The below block diagram describes the process done in this paper. We implemented code in python using machine learning supervised techniques like Support Vector Machine, Naïve Bayes and Random Forest. The following are the step by step process to reach to desired goal

1. Data Preprocessing: In the first step of proposed system. The preprocessing is a sequence of operation that performs on scanned input datasets. The first step in data preprocessing is removing punctuations, numbers and converting each word to lower case. Then next step is stemming the words. We found using stopwords, classifier working accurately. So implemented stopwords as next step in data preprocessing.

2. Feature Selection: Feature selection is an approach which is used to identify a subset of features which are mostly related to the target model, and

the goal of feature selection is to increase the level of accuracy. In this study, we implemented five feature selection methods widely used for the classification task of SA with Stopwords methods. The results differ from one method to the other. For example, in our analysis of Restaurant Review datasets, we found that the use of SVM algorithm is proved to be more accurate in the classification task.

3. Sentiment Classification algorithms: In this step, we used three algorithms which are mainly used for classification as well as regression models. In this project, we used those algorithms for classification of reviews. The classification algorithms used in this project are as follows → Support Vector Machine(SVM) → Random Forest → Naïve Bayes(NB) Naïve Bayes(NB) The NB classifier is a basic probabilistic classifier based on applying Bayes' theorem. The NB calculates a set of probabilities by combinations of values in a given dataset. Also, the NB classifier has fast decision-making process. Support Vector Machine (SVM) SVM in machine learning is a supervised learning model with a related learning algorithm, which examines data and

identifies patterns, which is used for regression and classification analysis. Recently, many classification algorithms have been proposed, but SVM is still one of the most widely and most popular used classifiers. Random Forest Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.

4. Detection process After training, as a next step predict the output of our model. In this step we predict the output using test data. Then a confusion matrix is generated as a output. It contains four fields as follows

True Positive (TP): The predicted review matches the actual review. The actual review was positive and the model predicted a positive review. Generally, these reviews are real positive reviews.

True Negative (TN): The predicted value matches the actual review. The actual review was negative and the model predicted a negative review. Generally, these reviews are real negative reviews. False Positive (FP): The predicted review was falsely

predicted. The actual review was negative but the model predicted a positive review. This type of review called as fake positive reviews.

False Negative (FN) : The predicted review was falsely predicted. The actual review was positive but the model predicted a negative review. This type of review called as fake negative reviews.

4. COMPARISON OF RESULTS

In this step, we compared the different accuracy provided by the dataset of restaurant reviews with various classification algorithms and identified the most significant classification algorithm for detecting Fake positive and negative Reviews.

After completing all the steps in this project, next step is displaying results.

Output for classifier support vector machine

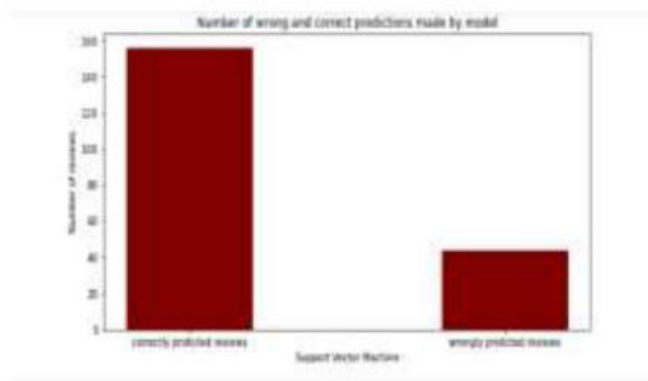


Fig: Output of classifier SVM

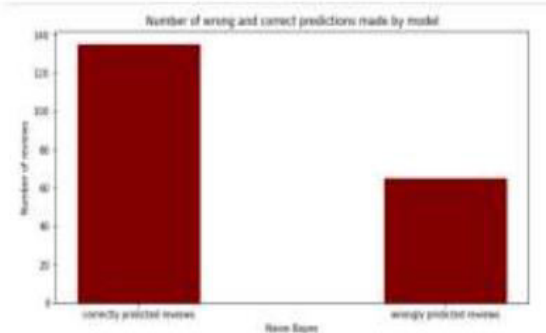


Fig: Output of classifier Naïve Bayes

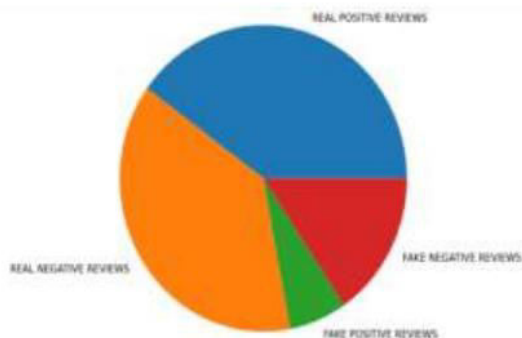


Fig: classification of reviews using SVM

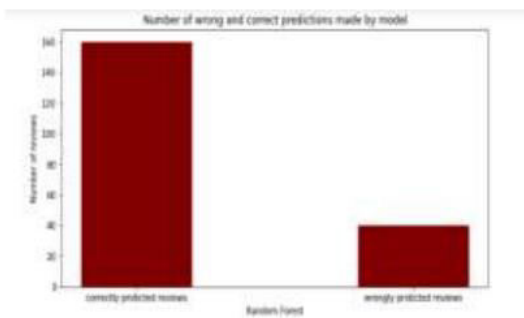


Fig:output for Random Forest

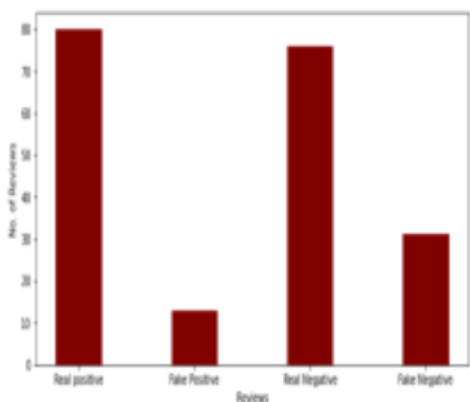


Fig:classification of reviews using SVM

5. CONCLUSION

In this paper, we proposed several methods to analyze a dataset of restaurant reviews. We also presented sentiment classification algorithms to apply a supervised learning of the movie reviews located in datasets. Our experimental approaches studied the accuracy of all sentiment classification algorithms, and how to determine which algorithm is more accurate. Furthermore, we were able to detect fake positive reviews and fake negative reviews through detection processes. Among three algorithms are used, SVM performs accurately.

REFERENCES

- [1]. B. Liu, "Sentiment analysis and opinion mining," Synthesis lectures on human language technologies, vol. 5, no. 1, 2012, pp. 1–167.
- [2]. W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis

algorithms and applications: A survey,”
Ain Shams Engineering Journal, vol. 5, no.
4, 2014, pp. 1093–1113.

[3]. B. Pang, L. Lee, and S. Vaithyanathan,
“Thumbs up? sentiment classification
using machine learning techniques,” in
Proceedings of EMNLP, 2002, pp. 79–86.

[4]. J. Malbon, “Taking fake online
consumer reviews seriously,” *Journal
of Consumer Policy*, vol. 36, no. 2, 2013,
pp. 139–157.

[5]. R. Xia, C. Zong, and S. Li,
“Ensemble of feature sets and
classification algorithms for sentiment
classification,” *Information Sciences*, vol.
181, no. 6, 2011, pp. 1138–1152.

[6]. SK Kotamraju, PG Arepalli, SS
Kanumalli (2021), Implementation
patterns of secured internet of things
environment using advanced blockchain
technologies, *Materials Today:
Proceedings*, 2021.

[7]. Gopi A.P., Patibandla R.S.M. (2021)
An Efficient Methodology for Avoiding
Threats in Smart Homes with Low Power
Consumption in IoT Environment Using
Blockchain Technology. In: Choudhury T.,
Khanna A., Toe T.T., Khurana M., Gia
Nhu N. (eds) *Blockchain Applications in
IoT Ecosystem*. EAI/Springer Innovations
in Communication and Computing.

Springer, Cham.
https://doi.org/10.1007/978-3-030-65691-1_16.

[8]. RSML Patibandla, SN Mohanty
(2021), Need of Improving the Emotional
Intelligence of Employees in an
Organization for Better
Outcomes, *Decision Making And Problem
Solving: A Practical guide*, 2021.

[9]. T. Barbu, “Svm-based human cell
detection technique using histograms of
oriented gradients,” *cell*, vol. 4, 2012, p.
11.

[10]. G. Esposito, LP-type methods for
Optimal Transductive Support
Vector Machines. Gennaro Esposito, PhD,
2014, vol. 3.

[11]. P. Kalaivani and K. L.
Shunmuganathan, “Sentiment
classification of movie reviews by
supervised machine learning approaches,”
*Indian Journal of Computer Science and
Engineering*, vol. 4, no. 4, pp. 285–292,
2013.

[12]. B. Pang and L. Lee, “A sentimental
education: Sentiment analysis
using subjectivity summarization based on
minimum cuts,” in *Proceedings of the 42nd
annual meeting on Association for
Computational Linguistics*. Association for

Computational Linguistics, 2004, p. 271.

[Online].

[13]. S. Hassan, M. Rafi, and M. S. Shaikh, “Comparing svm and naive bayesclassifiers for text categorization with wikitology as knowledge enrichment,” in Multitopic Conference (INMIC), 2011 IEEE, 2016, pp.84–89.