IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⵂ Impact Factor 7.105 ⵂ Vol. 9, Issue 6, June 2022

DOI: 10.17148/IARJSET.2022.9689

Fake Review Detection on Yelp Dataset

Suraj Venkatesh Shetty[1], Vidya S[2]

[1]Dept. of MCA, Bangalore Institute of Technology (BIT), Bengaluru, India

[2]Dept. of MCA, Assistant Professor, Bangalore Institute of Technology (BIT), Bengaluru, India.

Abstract: Customer reviews are crucially significant in today's modern era. It is preferable for a consumer to interpret good or service or store reviews before having to decide that which and where to purchase. Consumers may well be misled into purchasing low-quality items given the prevalence of spam feedback/reviews, while satisfactory businesses may very well be defamed by fraudulent feedback. Unlike, say, advertisements, online consumer reviews contain experiences from actual people. It thus has a tremendous influence on the level of customers as well as indirectly on firms. Concerningly, these monetary inducements have produced a market for spammers to generate evaluations in order to falsely boost or criticize firms, practices known as opinion spam. To solve this issue, we discover that the usual reviewers' arrival pattern is consistent and generally indifferent to their ranking patterns. In contrary, spam operations are typically brief and associated with the ranking either favorably or unfavorably. Hence, we advise using abnormally correlated temporal variations to spot such threats. In order to view and exploit such associations, we identify and create multivariate data set based on aggregate data.

INTRODUCTION

Individuals and businesses are progressively utilizing reviews online to assist them make purchase decisions as well as how to operate a business. For firms and individuals, high ratings can result in massive financial gains and fame. Regrettably, this presents pretenders with tremendous reasons to exploit the system by posting false ratings either to praise or degrade specific target items or firms. These groups are regarded as opinion spammers & their operations are recognized as opinion spamming. The problem of spam or fraudulent ratings has risen in recent years, and lots of high incidents have been reported in public. Even consumer sites have accumulated a huge amount of automated fraud detection suggestions. Furthermore, there have already been media inquiries where fake reviewers have publicly stated accepting funds to make artificial reviews [1].

There are already several user reviews published for a vast range of goods and services thanks to the swift growth of internet retailing. A significant pool of potential visitors relies on them to analyze the caliber of products or services prior making the purchase. As an outcome, based on a desire of gain or competition, businesses and sellers develop motives and practices to alter reviews, deliberately publishing falsified feedback to purposely deceive potential consumers and manipulate their risky purchasing decisions. An individual (known as an individual spammer) or a gang (known as a spammer group) may also be sponsored by makers in order to publish enhanced positive perceptions on their items or damaging negative reviews on that of their counterparts in order to enhance customer satisfaction and brand [2].

This paper presents an overview of our attempts to determine a learning machine technique to predict whether reviews on the help dataset are factually true or not. We precisely assessed and compared different classification techniques in machine learning such as Support Vector Machine, Logistic Regression, and Naive Bayes to assess which one would lead to better results. Yelp recognizes that such a possible danger will result in erroneous details for their users. To tackle this question, Yelp has already formed a provision for entrepreneurs. Quite apart from that, Yelp has also put in place a preferred software program that plan to immediately extract all worrisome feedback.

LITERATURE SURVEY

From 2007, the process of sensing fake reviews has been explored, with the assessment of review spamming. The researchers assessed the specific instance of online sources in this work, concluding that manually classifying fake reviews may be tricky, as fake reviewers may meticulously create their feedback in order to make them quite credible for other users. As a side effect, they recommended using duplicate records or well almost as spam to develop a model that able to detect fraudsters.

Spam review sensing is a part of the larger problem of deception detection, in which both non-verbal and verbal components can be used. Fake review detection research has primarily focused on text - based and behavioral factors, while other strategies have regarded social or temporal factors. Nonverbal attributes of review activity, such as the number of reviews or the duration and gadget on which the review was submitted, are illustrations of behavioral

International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified ⁒ Impact Factor 7.105 ⁒ Vol. 9, Issue 6, June 2022
DOI: 10.17148/IARJSET.2022.9689

features. They were used to enhance the classifier model, generating impressive outcomes. Support Vector Machine was the most widely used classification methodology, preceded by Naive Bayes, Decision Tree, Random Forest, and Logistic Regression. Other methodologies were used in addition to supervised methods because information gathering for studies is a complex job [3].

EXISISTING SYSTEM

The reviews are evaluated according to a range of criteria, such as their authenticity, recent activity, and sheer quality. Presently, Yelp recommends reviews in exceeding of 75% of the time. Unfortunately, there isn't a method or system that is hundred percent protected. Since machine learning can help raise detection rates, it might be beneficial in the fight against misleading reviews. In particular, data can be used to train machine learning classification systems to distinguish legit reviews from fake reviews.

PROPOSED SYSTEM

We calculated the extreme rate (1star or 5stars) ratio for each reviewer and used it as a criterion for assessing the reviews we presented, presuming that fraudulent reviewers typically use either 1- or 5-star rating to persuade users. By dividing the total no. of of opinions each reviewer wrote by the total no. of reviews written by all reviewers, it was possible to calculate the proportion of those reviews that earned an extreme rating of 1 or 5. This score was determined for each reviewer separately and used to influence each reviewer's perspective.

METHODOLOGY

For finding the fake reviews machine learning classifier techniques are used. They include support vector machine, Logistic Regression, Naïve Bayes algorithm.

A. Support Vector Machine(SVM)

The SVM algorithm's priority is to determine the best line or decision boundary that can split n-dimensional space into classes, permitting us to swiftly categorize new data points in the long term. A hyper - plane is the name given to this optimal decision boundary.

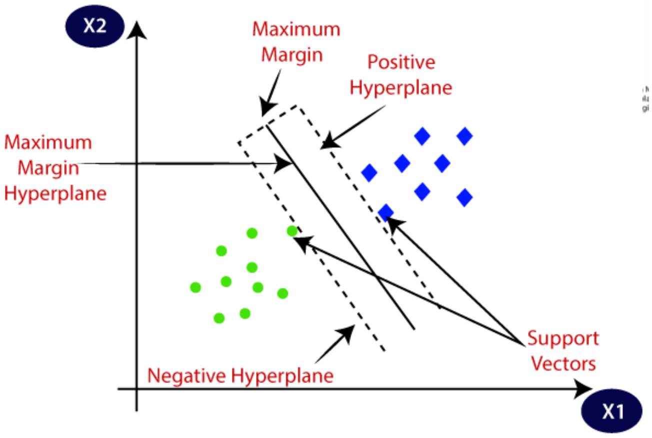


Fig. 1. SVM

B. Logistic Regression

Apart from how they are implemented, logistic regression and linear regression are very identical. Challenges concerning regression are solved using linear regression, while classification problems are solved using logistic regression.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 6, June 2022

DOI: 10.17148/IARJSET.2022.9689

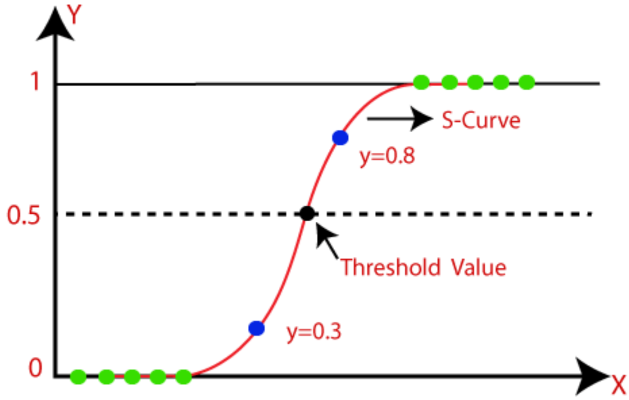


Fig. 2. Logistic Regression

C. Naïve Bayes

A class of supervised ML classification methodologies built on the Bayes theorem are known as naïve bayes classifier. Although it is a straightforward classification method, it is fully functioning.

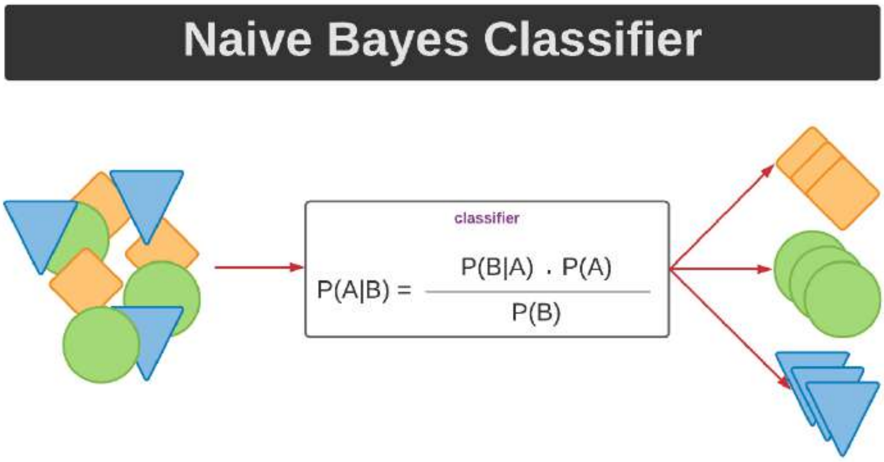


Fig. 3. Naïve bayes Classifier

ARCHITECTURE DESIGN

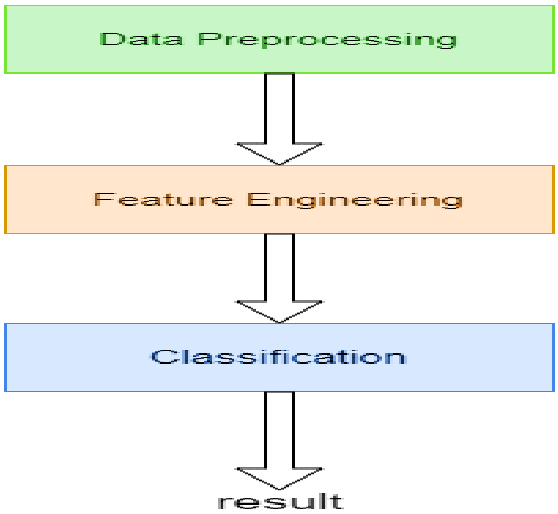


Fig. 4. Architecture Design

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⋇ Impact Factor 7.105 ⋇ Vol. 9, Issue 6, June 2022

DOI: 10.17148/IARJSET.2022.9689

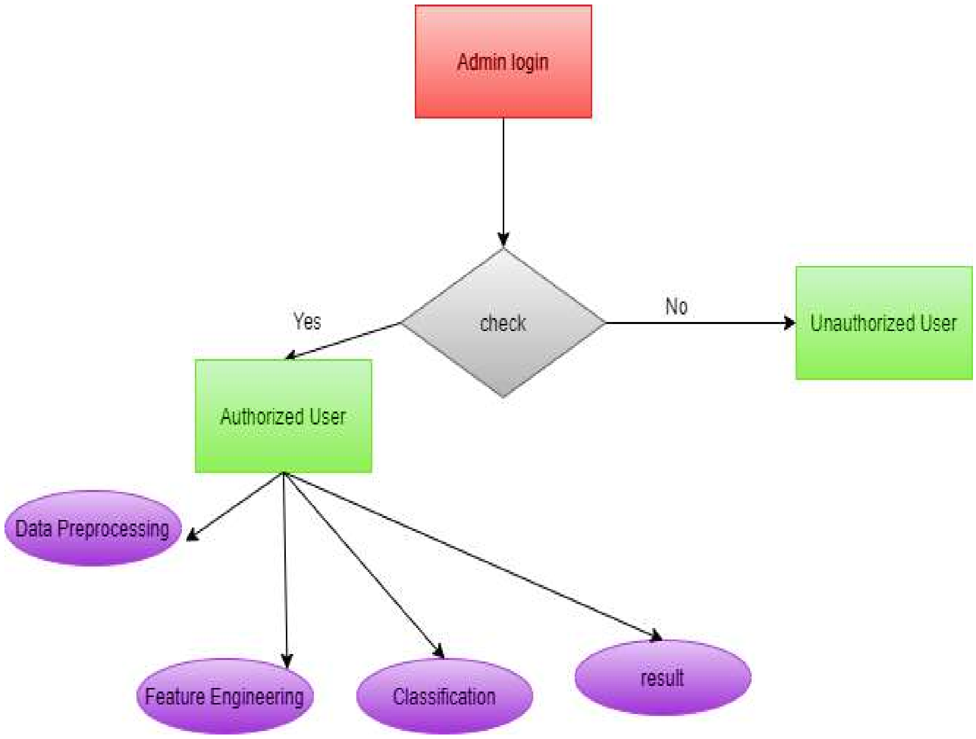


Fig. 5. Sequence of flow

I. Processing of Data

The set of data has a heavily slanted variety of filtered and unfiltered assessments, with a 1:6 proportion. In order to tackle this issue, we consider taking a dual strategy. The first tactic, known as over sampling, entails repeating data from a small group in order to boost the statistical impact of that group. Throughout this case, we're making 3 duplicates of the processed reviews to boost the number of duplicates, resulting in a ratio of around 1:2. The second strategy is under-sampling, which requires taking certain unfiltered assessments from the training data set. With the invalid reviews deleted, the ratio fell to somewhere around 1:3. The results indicate that the oversampling tactic performs better than the under-sampling strategy.

II. Feature Engineering

Before moving on to the feature engineering phase, we conduct some basic statistical analysis on the set of data. Figure 4 depicts how filtered reviews are much more likely to be highly positive or entirely nasty. We also encountered that filtered reviews are often shorter than non-filtered reviews; this is not clearly evident, but it is helpful data to have.

RESULTS

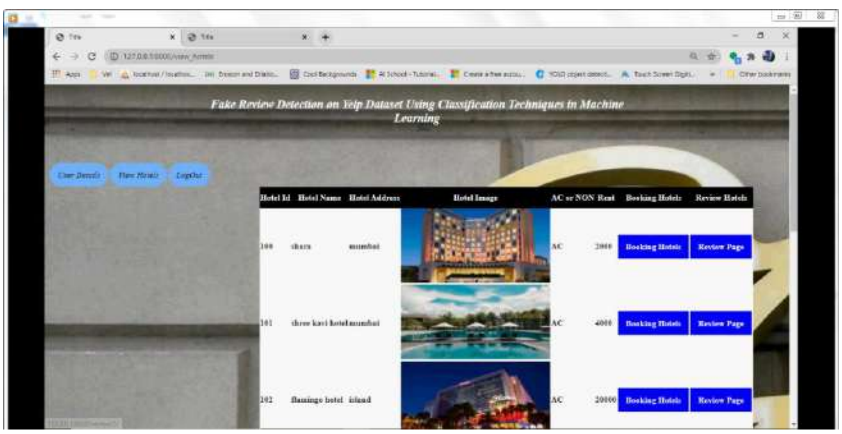


Fig .6. Hotel Reviews

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 6, June 2022

DOI: 10.17148/IARJSET.2022.9689

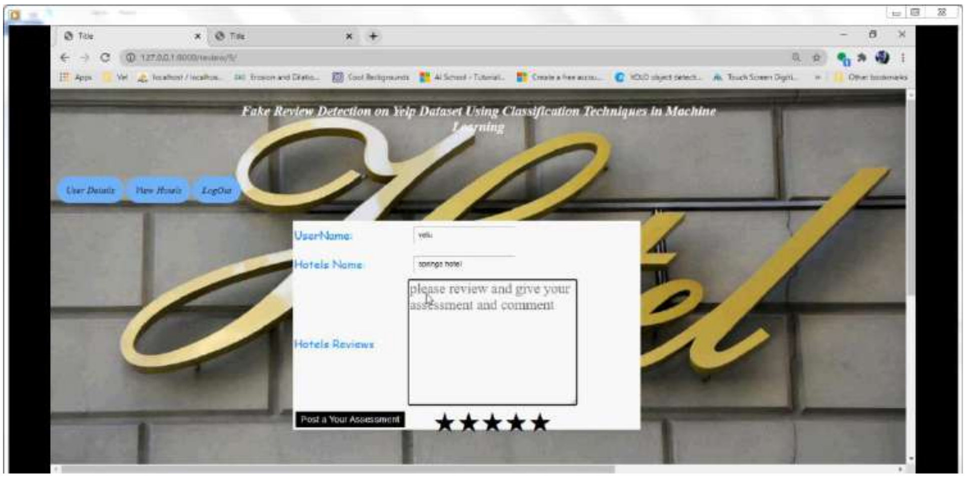


Fig .7. Posting Review

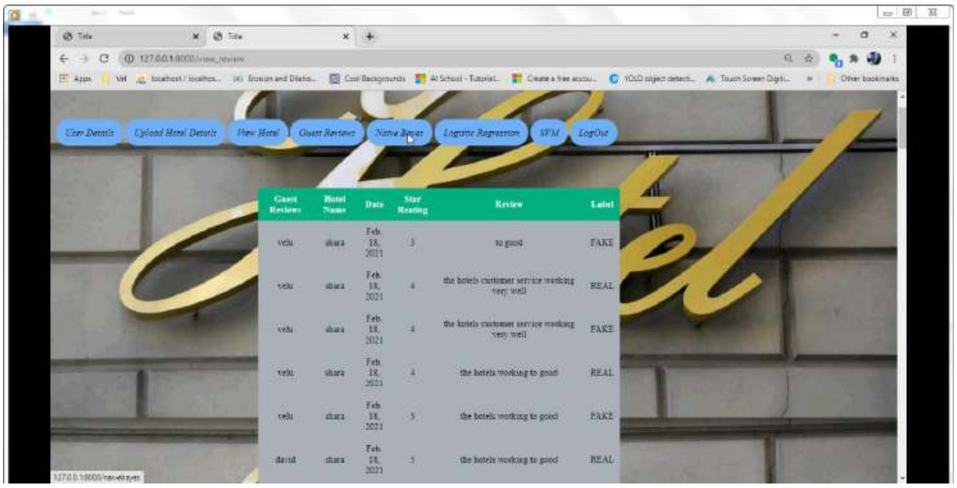


Fig .8. Review Result

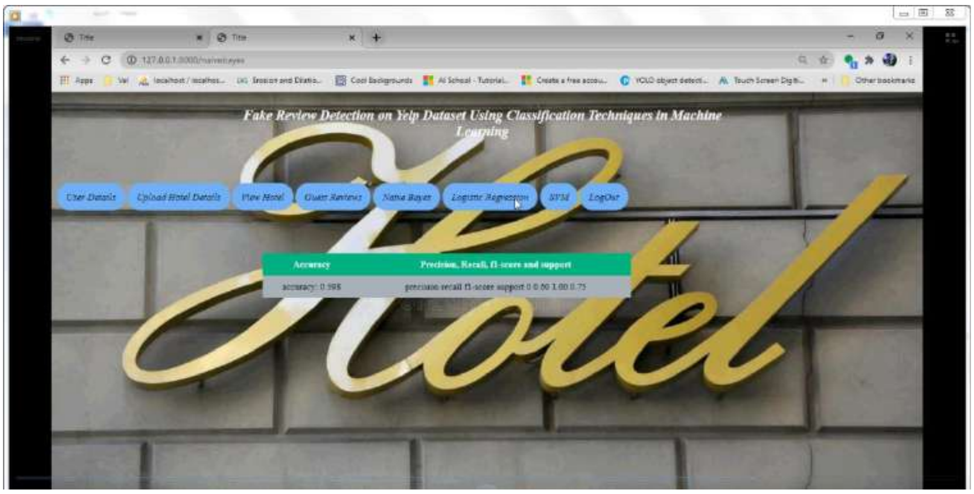


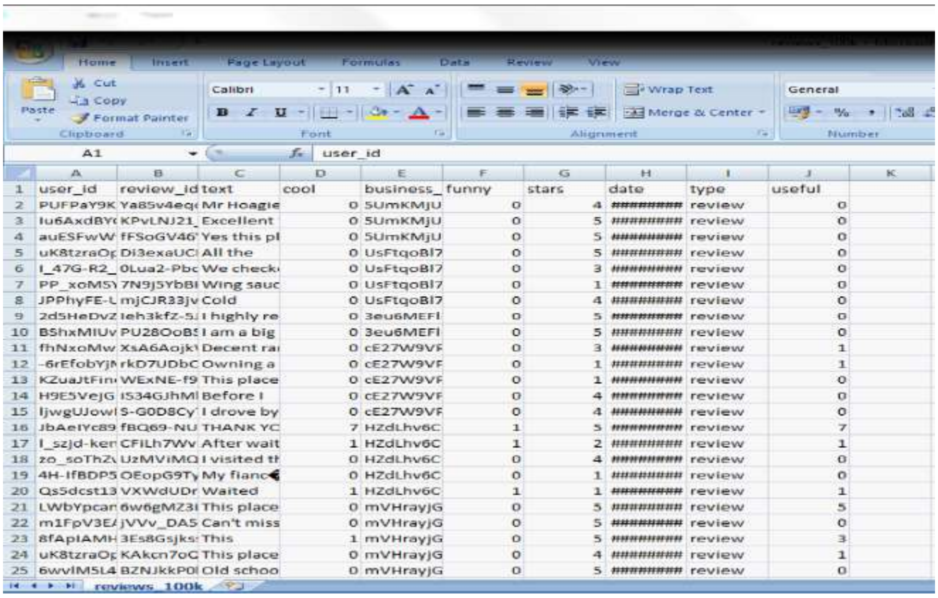
Fig .9. Precision of reviews

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 6, June 2022

DOI: 10.17148/IARJSET.2022.9689



	A	B	C	D	E	F	G	H	I	J
1	user_id	review_id	text	cool	business_	funny	stars	date	type	useful
2	PUFPaY9K	Ya85v4eq	Mr Hoagie	0	5UmKMjU	0	4	#########	review	0
3	Iu6AxdBY	KPvLNJ21_	Excellent	0	5UmKMjU	0	5	#########	review	0
4	auESFwW	fFSoGV46	Yes this pl	0	5UmKMjU	0	5	#########	review	0
5	uK8tzraOp	Di3exaUC	All the	0	UsFtqoBl7	0	5	#########	review	0
6	I_47G-R2_	0Lua2-Pbc	We check	0	UsFtqoBl7	0	3	#########	review	0
7	PP_xoMSY	7N9j5YbBI	Wing sauc	0	UsFtqoBl7	0	1	#########	review	0
8	JPPhyFE-L	mjCJR33jv	Cold	0	UsFtqoBl7	0	4	#########	review	0
9	2d5HeDvZ	Ieh3kfZ-5.	I highly re	0	3eu6MEFI	0	5	#########	review	0
10	BShxMIUv	PU28OoB	I am a big	0	3eu6MEFI	0	5	#########	review	0
11	fhNxoMw	XsA6Aojk	Decent ra	0	cE27W9VF	0	3	#########	review	1
12	-6rEfobYjN	rkD7UDbC	Owning a	0	cE27W9VF	0	1	#########	review	1
13	KZuaJtFin	WExNE-f9	This place	0	cE27W9VF	0	1	#########	review	0
14	H9E5VejG	i534GJhMI	Before I	0	cE27W9VF	0	4	#########	review	0
15	IjwgUJowI	S-G0D8Cy	I drove by	0	cE27W9VF	0	4	#########	review	0
16	JbAeIYc89	fBQ69-NU	THANK YC	7	HZdLhv6C	1	5	#########	review	7
17	I_szjd-ker	CFILh7Wv	After wait	1	HZdLhv6C	1	2	#########	review	1
18	zo_soThZ	UzMViMQ	I visited th	0	HZdLhv6C	0	4	#########	review	0
19	4H-IfBDP5	OEopG9Ty	My fianc	0	HZdLhv6C	0	1	#########	review	0
20	Qs5dcst13	VXWdUDr	Waited	1	HZdLhv6C	1	1	#########	review	1
21	LWbYpcar	6w6gMZ3	This place	0	mVHrayjG	0	5	#########	review	5
22	m1FpV3E	jVVv_DA5	Can't miss	0	mVHrayjG	0	5	#########	review	0
23	8fApIAMH	3Es8Gsjks	This	1	mVHrayjG	0	5	#########	review	3
24	uK8tzraOp	KAkcn7oC	This place	0	mVHrayjG	0	4	#########	review	1
25	6wvIM5L4	BZNJkkP0	Old schoo	0	mVHrayjG	0	5	#########	review	0

Fig .10. Data Anlysis

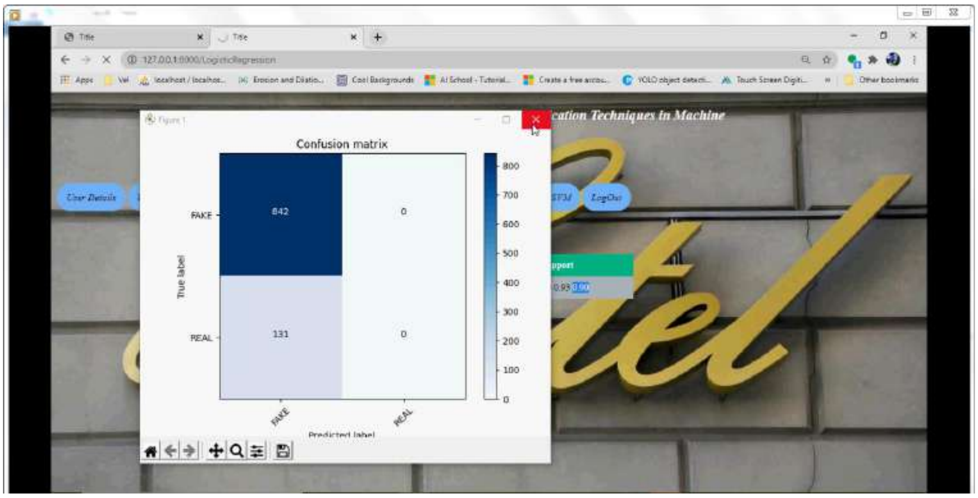


Fig .11. Confusion Matrix

CONCLUSION

The paper addresses four widely used ML classification methodologies for detecting spurious Reviews on yelp. Beneficial, nice, and hilarious ratings could only be obtained through un-doctored feedback, which indicates that once a review has been filtered, the ratings are lost forever. When there is disparity in the data we use, the results of our experiment suffer. We realized that SVM required the most time to train the model and that Gaussian Naive Bayes generated the lowest aggregate rating during the experiment.

We conclude that the evaluations removed from the YELP recommendation system were all fraudulent would be premature. Some web-based crowdsourcing projects have already used the authenticated purchaser approach, which is another presumably credible technique.

REFERENCES

1. M.J. Zhong, L.Tan, X.L. Qu, Identification of Opinion Spammers using Reviewer Reputation and Clustering Analysis
2. Shilpa Yadav, Fake Review Detection Using Machine Learning Techniques, JETIR.
3. Arjun Mukherjee, Vivek Venkataraman, Fake Review Detection: Classification and Analysis of Real and Pseudo

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⋇ Impact Factor 7.105 ⋇ Vol. 9, Issue 6, June 2022

DOI: 10.17148/IARJSET.2022.9689

Reviews.
4. Ott, M., Choi, Y., Cardie, C., Hancock, J.T.: Finding deceptive opinion spam by any stretch of the imagination. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Portland, Oregon, USA, pp. 309–319 (2011)
5. Heydari, A., Tavakoli, M., Salim, N.: Detection of fake opinions using time series. Expert Syst. Appl. **58**, 83–92 (2016)
6. Lim, E.-P., Nguyen, V.-A., Jindal, N., Liu, B., Lauw, H.W.: Detecting product review spammers using rating behaviors. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management, pp. 939–948. ACM (2010)
7. Xie, S., Wang, G., Lin, S., Yu, P.S.: Review spam detection via temporal pattern discovery. In: Proceedings of the 18th ACM International Conference on Knowledge Discovery and Data Mining, pp. 823–831. ACM (2012)
8. Ye, J., Kumar, S., Akoglu, L.: Temporal opinion spam detection by multivariate indicative signals. In: ICWSM, pp. 743–746 (2016)
9. Aghakhani, H., Machiry, A., Nilizadeh, S., Kruegel, C., Vigna, G.: Detecting deceptive reviews using generative adversarial networks. CoRR abs/1805.10364 (2018)
10. Abadi, M., et al.: Tensorflow: a system for large-scale machine learning.2016.
11. Castillo, C.,Donato,D.,Gionis,A.,Murdock,V.,Silvestri,F. Know your neighbors: web spam detection using the web topology. SIGIR 2007: 423-430.
12. Chawla, N.,Japkowicz,N. and Kolcz,A. Editorial: Special issue on learning from imbalanced data sets. SIGKDD Explorations Special Issue on Learning from Imbalanced Datasets, 6(1):1–6, 2004
13. Cheng, Z.,Gao,B.,Sun,C.,Jiang,Y.,Liu,T. Let web spammers expose themselves. WSDM 2011: 525-534
14. Chirita, P.A., Diederich, J., and Nejdl, W. MailRank: using ranking for spam detection. CIKM. 2005.
15. Chung, Y.,Toyoda,M.,Kitsuregawa,M. Detecting Hijacked Sites by Web Spammer Using Link-Based Algorithms. IEICE Transactions 93-D(6): 1414-1421 (2010)
16. Drummond, C. and Holte,R. C4.5, class imbalance, and cost sensitivity: Why under-sampling beats oversampling. In Proceedings of the ICML'03 Workshop on Learning from Imbalanced Data Sets, 2003.
17. Eric, Ask Yelp: Why can't I write reviews from my mobile? http://officialblog.yelp.com/2009/12/ask-yelp-why-cant-i-writereviews-from-my-mobile.html. Yelp Official Blog. December 2009.
18. Feng, S., Banerjee, R., and Choi, Y. Syntactic Stylometry for Deception Detection. ACL (short paper), 2012.
19. Feng, S., Xing, L., Gogar, A., and Choi, Y. Distributional Footprints of Deceptive Product Reviews. In ICWSM. 2012.
20. Holloway, D. Just Another Reason Why We Have a Review Filter. http://officialblog.yelp.com/2011/10/just-another-reason-why-wehave-a-review-filter.html. Yelp Official Blog. October 2011.
21. Jindal, N. and Liu, B. Opinion spam and analysis. WSDM. 2008.
22. Jindal, N., Liu, B. and Lim, E. P. Finding Unusual Review Patterns Using Unexpected Rules. CIKM.
23. Jin,X.,Lin,C.X.,Luo,J.,Han,J. SocialSpamGuard: A Data MiningBased Spam Detection System for Social Media Networks. PVLDB 4(12): 1458-1461 (2011)
24. Joachims, T. Making large-scale support vector machine learning practical. Advances in Kernel Methods. MIT Press. 1999.
25. Joachims, T. (1998) Text categorization with support vector machines: Learning with many relevant features.. ECML, 1998.
26. Identification of Opinion Spammers using
27. Reviewer Reputation and Clustering Analysis