Performance Analysis of Supervised Techniques for Review Spam Detection

Mr. Ashok Badresiya
C.E. Department, Marwadi Education Foundation's Group of Institutions
Rajkot,Gujarat, India
ashok.badresiya2011@gmail.com
Prof. Saifee Vohra
C.E. Department, Marwadi Education Foundation's Group of Institutions
Rajkot,Gujarat, India
saifee.vohra@marwadieducation.edu.in
Prof. Jay Teraiya
C.E. Department, Marwadi Education Foundation's Group of Institutions
Rajkot,Gujarat, India
Jay.teraiya@marwadieducation.edu.in

--**ABSTRACT**---

Nowadays, millions of products and services are available online. Searching for the best products which targets the individuals' requirements would be difficult as the result of the existence of too many offers. One of the most useful approaches to choose a product or service is to use the reviews of the others who have already tried them. A reviewing system is a place where individuals write their reviews on their experienced products and services, and also benefit from others' reviews. Moreover, companies utilize reviewing systems to apply opinion mining techniques in order to improve their goods or services and to watch their competitors. However, the popularity of the reviewing systems ignites this motivation for some people to enter fake review to promote some products or defame competitors products. These review spam should get detected and eliminated in order to prevent misleading potential customers. Opinion mining techniques should use to locate and eliminate potential spam reviews. The objective of this paper is to discover the concept of Review spam detection in the field of opinion mining, and presents a performance analysis of its techniques in this field.

Keywords - *Machine learning, Opinion Mining, Review spam*

I. INTRODUCTION

With the increasing popularity of internet, online shopping and selling also increases nowadays. Hence many people buy product online through e-commerce websites. Many e-commerce website allows customers to review different product based on experience of product. So it can be helpful to other users to make decisions for buying products. Reviews are very common in e-commerce website and reviews are also very useful for customers and business organizations. Reviews can be useful to customers to take decision in purchase of products and reviews are useful for the organization to make quality improvement based on the reviews of specific product. It can be very useful to make business decisions. Clearly customer values feedback of others before making buying decision .This gives motivation to some users or organizations to posts spam reviews to promote or defame brand or some target products, services, organizations, individuals, and even ideas without disclosing their true intentions. These spammed opinion information is called Opinion Spam or Review Spam [1] [4].

Online customer reviews for both products and merchants have greatly affected others' decision making in purchase. Considering the easily accessibility of the reviews and the significant impacts to the retailers, there is an increasing incentive to manipulate the reviews, mostly profit driven. Websites containing customer reviews are becoming targets of opinion spam. -- Undeserving positive or negative reviews; reviews that reviewers never use the product, but is written with an agenda in mind. So, review spam detection is getting importance nowadays. Many researchers have been working on them today [1] [2] [3] [4].

There are mainly two kinds of methods for review spam detection: supervised methods and unsupervised method. Supervised methods can be implemented by building a classifier. This classifier is trained by examples which can be manually labeled. Machine learning starts with collecting training dataset. The next step is to train a classifier on the training data. Mostly used supervised methods are support vector machine (SVM), Naïve Bayes classifier, logistic regression, K-NN classifier, etc.

The objective of this paper is to discover the concept of review spam in the field of opinion mining and presents a comparative analysis of its techniques in this field. The paper is organized as follows: Section II provides the overview of the most commonly used superviesd techniques in review spam detection. Section III discusses the analysis and comparison of review spam detection techniques. Section IV concludes the manuscript.

II. SUPERVISED TECHNIQUE PERFORMANCE IN REVIEW SPAM DETECTION

In this section, we present mainly used supervised techniques for review spam detection and its performance. For performance evaluation, we have used dataset which contains total 1600 review from which 800 reviews are truthful (non-spam) and 800 reviews are deceptive (spam). M. Ott et al. [17] used this dataset for their research work and it is publically available. This is the gold standard dataset for review spam detection field. For performance analysis we have used Rapidminer data mining tool.

A. *Naïve Bayes Classifier*

Bayesian classifier is statistical classifier based on Bayes theorem. They can predict membership probabilities, such as the probability that a given tuple belongs to which specific class. A Naïve Bayes classifier accepts that the effect of an attribute value of a given class is independent of the value of the other attributes. This assumption is called class conditional independence. This classifier is very simple, quick, precise, and simple to implement. It is based on a basic assumption in real life and is only valid to multiply probabilities when the events are independent [5] [6].

We applied Naïve Bayes classifier on our dataset for review spam detection. This classifier gives around 64% accuracy for dataset [17] . The confusion matrix for this is shown in table I.

TABLE I
CONFUSION MATRIX FOR NAÏVE BAYES CLASSIFIER

Predicated Class		Actual Class	
		Spam	**Non spam**
	Spam	540	277
	Non spam	260	523

B. *Support Vector Machine (SVM) Classifier*

A support vector machine (SVM) is a set of related supervised learning methods used for classification and regression. In simple words, given a set of training examples, each marked as belonging to one of two categories, the SVM training algorithm builds a model that predicts whether a new example falls into which specific category. Intuitively, SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on [5] [7] [8].

More correctly, SVM constructs a hyper plane or a set of hyper planes in a high dimensional space, which can be used for classification, regression or other tasks. Intuitively, a good separation is achieved by the hyper plane that has the largest distance to the nearest training data points of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier [5] [7] [8].

Currently, SVM is widely used in object detection and recognition, content-based image retrieval, text recognition, biometrics, speech recognition, speaker identification, benchmarking time-series prediction tests. Using SVM in text classification is proposed by [9], and subsequently used in [10] [11].

SVM classifier gives around 83% accuracy for dataset [17] with this classifier. The confusion matrix for this is shown in table II.

TABLE II
CONFUSION MATRIX FOR SVM CLASSIFIER

Predicated Class		Actual Class	
		Spam	**Non spam**
	Spam	670	139
	Non spam	130	661

C. *K-Nearest Neighbour (K-NN) Classifier*

K-Nearest Neighbour (K-NN) algorithm is one of the supervised learning algorithms that have been used in many applications in the area of data mining, statistical pattern recognition and many others. It follows a method for classifying objects based on closest training examples in the feature space. An object is classified by a majority of its neighbours. K is always a positive integer. The neighbours are selected from a set of objects for which the correct classification is known [12].

K-NN works well even when there are some missing data. K-NN is good at specified which predictions have low confidence. It has some strong consistent results. As the amount of data approaches infinity, the algorithm is guaranteed to yield an error rate no worse than twice the Bayes error rate (the minimum achievable error rate given the distribution of the data) [13].

K-NN classifier gives around 69% accuracy for dataset [17] with this classifier. The confusion matrix for this is shown in table III.

TABLE III
CONFUSION MATRIX FOR K-NN CLASSIFIER

Predicated Class		Actual Class	
		Spam	**Non spam**
	Spam	600	303
	Non spam	200	497

D. *Logistic Regression Classifier*

Logistic regression classifier is very popular and widely used classification technique. This is simple, easy to implement, and provide good performance on a wide variety of problems. Logistic regression is a discriminative probabilistic classification model that operates over real-valued vector inputs. The dimensions of the input vectors being classified are called "features" and there is no restriction against them being correlated. Logistic regression is one of the best probabilistic classifiers, measured in both log loss and first-best classification accuracy across a number of tasks [14].

We have applied logistic regression classifier on [17] dataset for review spam detection. This gives around 82% accuracy for dataset [17] with this classifier. The confusion matrix for this is shown in table IV.

TABLE IV
CONFUSION MATRIX FOR LOGISTIC REGRESSION CLASSIFIER

Predicated Class		Actual Class	
		Spam	**Non spam**
	Spam	658	140
	Non spam	142	660

E. *Decision Tree Classifier*

Decision tree is a binary tree structure whose internal nodes correspond to input patterns and whose leaf nodes are categories of patterns. Classifier uses the concept of tree structure to classify the given data in to the different number of classes based on the training data, Structure is mainly divided into two parts nodes and branches, there are mainly two things in tree which plays very important role in classifying the data, one is Root node from which all the instances are going to be classified and goes to the leaf node based on their feature values, Leaf node contains the actual class label which is required to be determined. Every single node in decision tree represents a feature which will help in classification of an instance, and each branch in decision tree represents a value of a node [5] [15].

We have applied decision tree classifier which gives around 51% accuracy for dataset [17] with this classifier. The confusion matrix for this is shown in table V.

TABLE V
CONFUSION MATRIX FOR DECISION TREE CLASSIFIER

Predicated Class		Actual Class	
		Spam	**Non spam**
	Spam	777	761
	Non spam	23	39

III. COMPARATIVE PERFORMANCE ANALYSIS

This section presents the comparative performance analysis of the supervised techniques presented in above sections. We have compared the performance in term of accuracy, recall, and precision measures. The comparison is shown in below table VI.

Supervised machine learning techniques have shown relatively better performance than the other unsupervised methods. Supervised methods demand large amounts of labeled training data that are very expensive whereas acquisition of unlabelled data is easy.

Most of the researchers found that Support Vector Machines (SVM) has high accuracy than other classification algorithms. The main limitation of supervised learning is that it generally requires large expert-labelled training dataset to be created from scratch, specifically for the application at hand, and may fail when training data are insufficient.

TABLE VI
COMPARATIVE PERFORMANCE ANALYSIS OF SUPERVISED TECHNIQUES

Techniques	Measures		
	Accuracy	**Recall**	**Precision**
Naïve Bayes	66.44%	66.10%	67.50%
SVM	83.19%	83.75%	82.82%
K-NN	68.56%	75.00%	66.45%
Logistic regression	82.39%	82.25%	82.46%
Decision tree	51.00%	97.12%	50.52%

IV. CONCLUSIONS

This paper shows that support vector machine (SVM) outperform than all the other supervised techniques for review spam detection. Logistic regression classifier also gives good accuracy in review spam detection. However, when the set of training data is small, a Naive Bayes classifier might be more appropriate since SVMs must use large set of data in order to build a high-quality classifier.

More future work is needed on further improving the performance of the Review spam detection. There is a huge need in the industry for such applications because every company wants to know how consumers feel about their products and services and those of their competitors. Different types of techniques should be combined in order to overcome their individual drawbacks and benefit from each others merits, and enhance the Review spam detection performance.

Here we have use the content of review only to detect spam review. Other future direction for improving accuracy of each method is to use of other features such as using ratings of reviews, number of helpful feedbacks, time of reviewed, etc.

REFERENCES

[1] M. HU and B. LIU, "Mining and summarizing customer reviews," Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 168-177, 2004.

[2] A.M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing, pp.339-346, 2005.

[3] F.K. Rahaghi, “Towards Review spam Detection,” Master’s Thesis, University of Calgary, APRIL 2013.

[4] Ahmad S. J. Abu Hammad, “An approach for Detecting Spam in Arabic Opinion Reviews,” Master’s Thesis, Islamic University of Gaza, January 2013.

[5] Han, J. and Kamber, M. (2006) Data Mining: Concepts and Techniques. 2nd Edition. Morgan Kaufmann Publishers, San Francisco, USA. (ISBN-55860-901-6).

[6] Y. Sun, M.S. Wong, A.K.C. and Y. Wang, ”Cost-Sensitive Boosting For Classification of Imbalanced Data. Pattern Recognition,” Vol. 40, No. 12, Pages 3358-3378, 2007.

[7] R. Duda, P. Hart, and D. stork, Pattern Classification. 2nd Edition, Wiley Interscience, 2001.

[8] E. Frank, and I. Witten, Data Mining: Practical Machine Learning Tools and Techniques. 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

[9] T. Joachims, Text Categorization with support Vector Machines: Learning with Many Relevant Features. In Proceedings of ECML-98, 10th European Conference on Machine Learning, 1998.

[10] S. Dumais, D. Heckerman, J. Platt, and M. Sahami, ” Inductive Learning Algorithms and Representations for Text Categorization,” In Proceedings of ACM-CIKM98, Pages 148-155, 1998.

[11] M. Haruno, and H. Taira, “Feature Selection in SVM Text Categorization,” In Proceedings of the 16th National Conference on Artifical Intelligence, Pages 480-486, 1999.

[12] M. Bramer, Principles of Data Mining, Springer-Verlag, London, 2007. (ISBN 184628-765-0).

[13] O. David, and M. Francesco, "Research Challenge on Opinion Mining and Sentiment Analysis,” The CROSSROAD Roadmap on ICT for Governance and Policy Modeling, 2010.

[14] L. Liu, and Y. Wang, “a Method for Sorting Out the Spam from Chinese Product Reviews,” In Proceeding of the Conference on Consumer Electronics, Communications and Networks.(CECNet), 2012.

[15] A. S. Galathiya, “Improved Decision Tree Inducation Algorithm with Feature Selection, Cross Validation, Model Complexity and Reduced Error Pruning,” International Journal of Computer Science and Information Technologies, Vol. 3(2), PP. 3427-3431.

[16] Bing Liu, Sentiment analysis and subjectivity, Handbook of Natural Language Processing, second edition, 2010.

[17] Myle Ott, Yejin Choi, Claire, and Jeffrey Hancock. Finding Deceptive Opinion Spam by Any Stretch of the Imagination. ACL 2011:309-319.