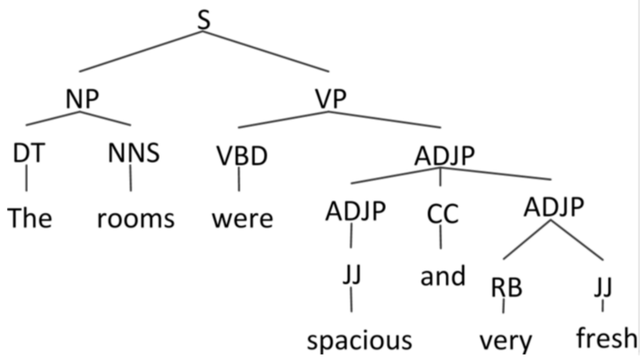
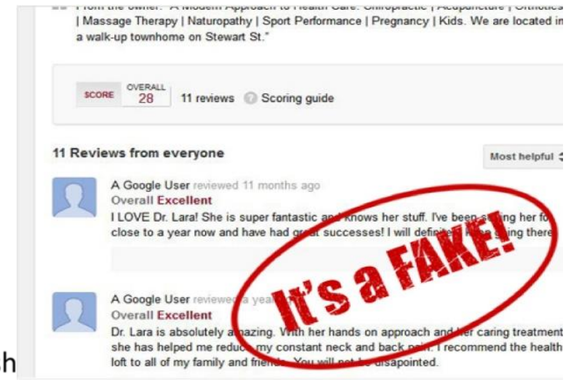
Sentiment Analysis for Spam Detection





A Project Report

On

Sentiment Analysis for Spam Detection

Guided By

Yashvardhan Sharma

[Assistant Professor, Department of Computer Science and Information Systems]

Submitted by

Astha Gupta

(2012C6PS736P)

DEPARTMENT OF COMPUTER SCIENCE AND INFORMATION SYSTEMS

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

May – 2016

Acknowledgement

I would like to take this opportunity to express my sincere thanks of gratitude to Dr. Yashvardhan Sharma for accepting me under his guidance for the project and helping me thought the project.

Rupal Bhargava for her ideas and involvements in the project.

Professor Rahul Banerjee, Head of Department and all faculty members at Department of Computer Science and Information Systems, BITS Pilani, who though various courses and project made me equipped to do this project with required skills set for its successful conductance.

My friends Anushka and Suyash for helping me throughout the project.

I am grateful to my friends and family for their constant support, love and encouragement.

Abstract

Online review have increasingly become the most popular and easy way to get information about various products, brands, stores and other commodity through online reviews posted by users over different websites. Customer choice rely heavily on these reviews and some people are using it to influence people by writing spurious reviews, these people are called spammers. Email spam which is a very well researched domain is similar to online spam detection but there are stubble difference that makes it difficult to correctly classify reviews. Most of the previous work majorly concentrated on features such as review centric, reviewer centric and product centric using rating as one of the important features. Problem with such approach is that they do not consider the hints dropped by the spammer in reviews majorly due to lack of his true views or experience with the product, which are many times reflected in sentiment of the review. Therefore major focus of this project is to use sentiment analysis to better classify reviews.

Table of Contents

1. Introduction

With help of internet one can now express themselves and interact with other, by various means such as emails, chats, reviews, comments and posts on different websites. Word-of-mouth effect is a phenomenon that is based on the influence caused by people's opinion and attitude towards a particular subject and its effect in shaping process of making decisions [1]. With these increased interaction and easily accessible information, which are most of the time very helpful, concerns regarding credibility of information floating on internet comes into picture. "User-generated contents" are contribution made by individual user on Web [2]. Importance and impact of user-generated contents have been fairly recognised and felt, but many times their power is used by organizations to influence potential users for different purpose such as increasing sales, participation in an event, increasing reach or band value.

In this project focus is on review spams. Review are a feedback given about a product that they have used and want to share information about it and their experience with potential customers with the motive of helping other customers to make informed decision, since given large variety of every product in market user may not have experience with a particular product of a particular brand. Review spams are spurious reviews written by spammers that are not written to serve purpose as stated above, but with an intention of making profit by influencing potential customers.

In this project we are trying to solve two problems in the field of spam detection. First use of Sentiment Analysis to detect Spams in online product reviews and second creating a framework for Data Annotation. Also side by side work on Survey paper for spam detection is being conducted.

1.1 Sentiment Analysis

Sentiment analysis is also known as opinion mining. Natural language processing (NLP), computational linguistics and text analysis are three fields that is widely used under this topic identify. The aim is to extract subjective information in given data source. Studies in spam detection have majorly focused on feedback score and user ratings as indicator to detect the spam reviews. Less work is done towards incorporating sentiment analysis to improve prediction. Detecting Spam Review through Sentiment Analysis [3], claims to be the first paper that used these scores for spam detection. Paper [4, 5] deal with features that can be extracted through sentiment analysis for spam and spammer detection respectively. Here in this project we use sentiment analysis to learn classification with help of other features suggested by Bing Liu et al. [6].

1.2 Data Acquisition and Annotation

Due to difficulty of marking a review as spam or non-spam, there is very less annotated data available for the domain. Most common tool used in spam detection problem is machine learning [2] and for any machine learning attempt accurate and reliable data set is very important. However it is easy to collect huge amount of unlabelled data, labelling the data is difficult and costly and therefore very less free annotated data is available on web. Difficulty in labelling arises due to the domain, for even a skilled expert it is very difficult to correctly identify spam and non-spam and hence many techniques have been developed to tackle the issue.

2. Literature Survey

2.1 Sentimental Analysis

Sentimental Analysis incorporated different kinds of techniques and many times hybrid of different techniques. Below is an attempt to classify Sentimental Analysis based on its attributes and different techniques used to deal with the attributes [7].

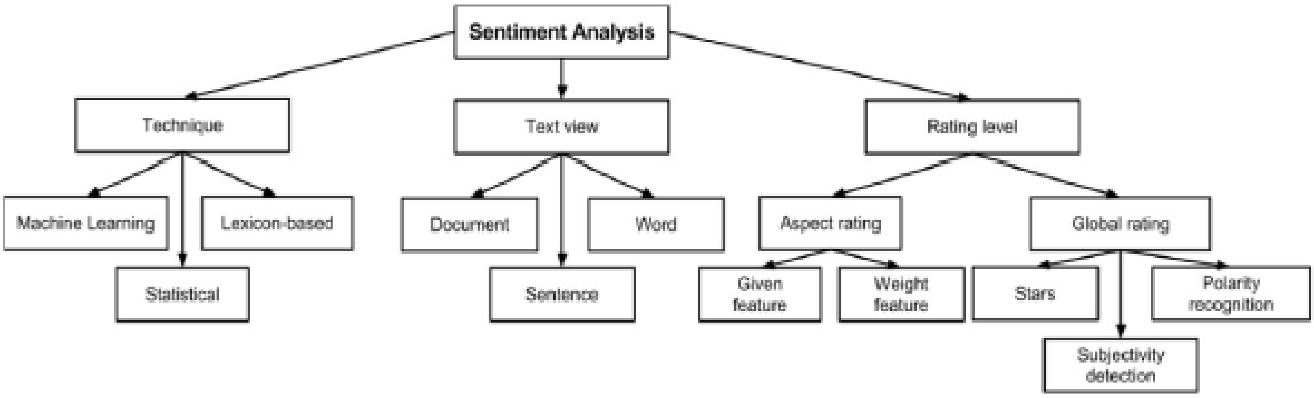


Figure 1: Classification of Sentimental Analysis

Sentimental Analysis can be best understood by knowing different applications [8]. There are different domains under sentimental analysis as stated.

- Document Sentiment Classification
- Sentence Subjectivity and Sentiment Classification
- Opinion Lexicon Expansion
- Aspect-Based Sentiment Analysis
- Mining Comparative Opinions
- Opinion Spam Detection
- Utility of Reviews
- Others

A lot of work has been done under these domains, most of the domains can be thought as a combination of techniques under various attributes as shown by Figure 1.

2.2 Opinion Spams

Opinion spams can be thought in three different ways. Detecting review spams, detecting spammer and detecting spammer group. All the three are very essential approaches in order to give credible information to consumer online. In paper by Heydari et al. [9] different ways in which these three approaches have been implemented and different papers that talk about these approaches have been discussed in detail.

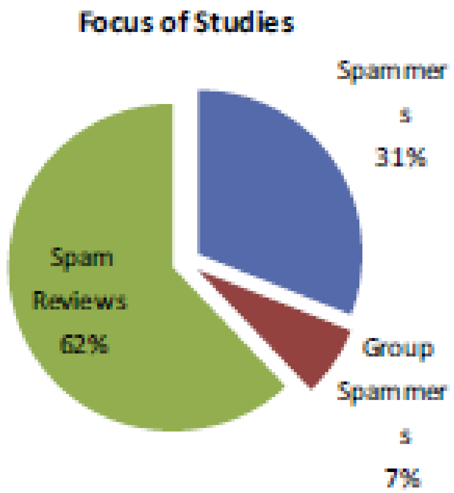


Figure 2: Proportion of focus of study

Opinion spam detection can be divided in three different classes depending the kind of approach it follows as mentioned in paper by Arjun Mukherjee [17].

I. Linguistic Signals: Leveraging linguistic signals present in the data. It can be further divided into three categories.

   a. N-gram language models: Sequence of *n* items in text or speech data is called N-gram. This approach is used to form models for detecting spams using sequence of words that occur in text.

   b. Psycholinguistics: Psychology of language (linguistic), it is study of factors that help humans to understand and comprehend language . Linguistic Inquiry and Word Count (LIWC) is one of the major feture used in the field under this category.

   c. Stylometry: Study of linguistic style, use of shallow and deep parsing has been used under stylometry [18].

II. Behavioural Signals: Recurring pattern that can be detected due to mass production of spams.

   a. Rating, Reviewing, & Collusion Behaviours: Analysis with respect to rating or reviewing has been used to identify spam in many papers.

b. Distributional and Time-Series Analysis: Checking for pattern based on time or any other measure.

c. Graph Based Methods: Graphically identifying related items, such as graphical analysis for group spam detection has been explored

d. Behavioural Features: Analysis of behavioural patterns have been suggested by Mukherjee et al. [19, 20].

III. Machine Learning & Statistical Modelling: Machine Learning is the most common method used in this domain. A survey paper by Crawford et al. [10] discussed different approaches in machine learning, features used and results produced.

a. Supervised: Most of the paper that use Machine Learning for classification use supervised learning.

b. Unsupervised Methods: Methods that do not require a classified training data. Used in paper by Mukherjee et al. [21]

c. Semi-Supervised Learning: Semi-Supervised framework has been proposed by Li et al. [4]

d. Latent Variable Models: Wentao Tian et al. [22] discusses latent variable.

3. Methodology

Methodologies for different part of the project is listed in different sections. Focus of this project has been mainly on sentimental analysis and survey for spam detection.

3.1 Semantic Analysis

As stated in introduction, not much work has been done in use of Sentimental Analysis for spam detection, one of the significant work is by Peng et al. [3] where sentimental analysis is used to generate scores regarding sentiment of a review and its comparison with average sentiment and rating are used in time series analysis to generate rules for spam detection. The three scores proposed by this paper are used as added features in different machine learning algorithm with features suggested by Liu et al. [6]. Different algorithms are learned such as Naïve Bayes, Ada Boost and Random Forest etc.

3.2 Data Acquisition and Annotation

Data sets in the project and their acquisition are stated below.

a. Deceptive Opinion Spam Corpus v1.4 – used in papers by M. Ott et al. [11, 12]. Please refer to appendix 10.1 to get link for downloading corpus. This data set is used for Sentimental Analysis

b. Amazon Product Review Data - Link given in appendix 10.2, this dataset is used in paper by Bing Liu et al. [13]. In the project data set is used for annotation.

For Data Annotation methodology followed is making an exhaustive list of points that needs to be considered while annotating a data set. First we need some pre-processing step as suggested by Lim et al. [14], secondly reviews that can be directly be marked as spam or non-spam should be remove them from the data set and then look at the various aspects of a review to manually annotate reviews. Since human accuracy is very less as suggested in paper by Somayeh et al. and Ren et al. [15, 16], we ask 3 people to review same set of reviews and then take a vote in case of a conflict.

3.3 Survey for Spam Detection

Methodology used in Survey of spam detection is first to analyse different surveys, already present in this domain. A corpus of 83 papers is created through three different methods. First is to take basic papers such as by Liu et al. and papers that pop us most in google search of spam detection in opinion mining, then back-search for references listed in these papers. Remove papers that are not related to online opinion spams of product reviews. Second method is to search for papers in famous journals and publisher’s websites, using keywords such their synonyms. Third find names of famous people in that domain, by looking at frequent names and then search for papers written by these authors.

4. Proposed Technique

4.1 Data Acquisition and Annotation

A semi supervised technique is proposed for data annotation. There are four step for data annotation.

1. Pre-processing

    a. Remove anonymous user
    b. Remove duplicated products: products that are essentially same but with minor variation, choose on representative out of the variants
    c. Remove inactive users and unpopular products
    d. Resolution of same brand names: such as “HP” and “Hewlett Packard”

2. Definite Spams

    Definite spams should be removed. These are rules which can be used to directly mark reviews as spams

    a. Remove reviews with all capital letters and mark them as spams
    b. Remove duplicates and near duplicates
    c. Remove reviews with links in the review
    d. Remove review if it is reviewed weeks/months before said product is actually released
    e. Remove if any kind of discount code given or review tell you where to go to buy the product – “can buy at “ and its variant

3. Likely Spams

    This is a list of points that can help us create set of likely spams and then based on different points in the same list we can mark spam or non-spams. Such as sets can be created based on criterions stated below.

1. Extreme Rating : Very High or Very Low
    a. No Information
    b. Low Information
2. Number Of Reviews:
    a. 0-2 reviews
    b. 3-5 reviews
    c. 5 and above
3. Time based Categorization: Use of box plot to see the time difference between first review and other reviews.
    a. Looking at the first few reviews – cookie cutter
    b. Looking at the burst of reviews – spamming behaviour, if many review between 48 hours then they are all likely to be spam

4. Categorization of the length
    a. Only one sentence – less than 20 words – high chances of fake, since no Information provided
    b. 5-10 lines of review
    c. More than 10 – lengthy and therefore crafted well by spammers, user will not spend much time on writing lengthy reviews

An exhaustive list of points that should be kept in mind while making spam is provided in appendix 8.3 with all the sources used to compile the list. Annotator needs to keep these points in mind and identify if a spam belongs to any of the stated category.

4. Semi Supervised Learning

   Create some amount of dataset for learning using the above strategy, rest data can be annotate using a semi supervised framework proposed by Li et al. [4] using three different models to learn the same data but on different features and taking a vote on their prediction.

4.2 Sentimental Analysis

For each review sentiment scores need to be calculated as given by Peng et al. [3]. For calculating these scores SentiWordNet 3.0 is used, based on different POS tag and word combination a score is provided to each sentimental word. Scores are between -1 to 1 and multi word as well as inversions are handled.

There are three scores proposed by Peng.

Sentiment Score (SS): Denoted by $o(d)$ represents sentiment polarity of a reviews.

$$o(d) \in [-1, 1]$$

Sentiment Ratio, SR: Denoted by $r(d)$ is a ratio of SS to all sentiment of all sentences,

$$r(d) \in [0,1] .$$

Difference of Sentiment Polarity, DSP: Denoted by $f(d)$ represent inconsistency in the rating and actual sentiment score

Since the Hotel Review Data has equal no of deceptive and genuine reviews with no ratings Sentiment Ratio and Difference of Sentiment Polarity were not calculated. Though once data set for Amazon Review is ready, these scores can also be used.

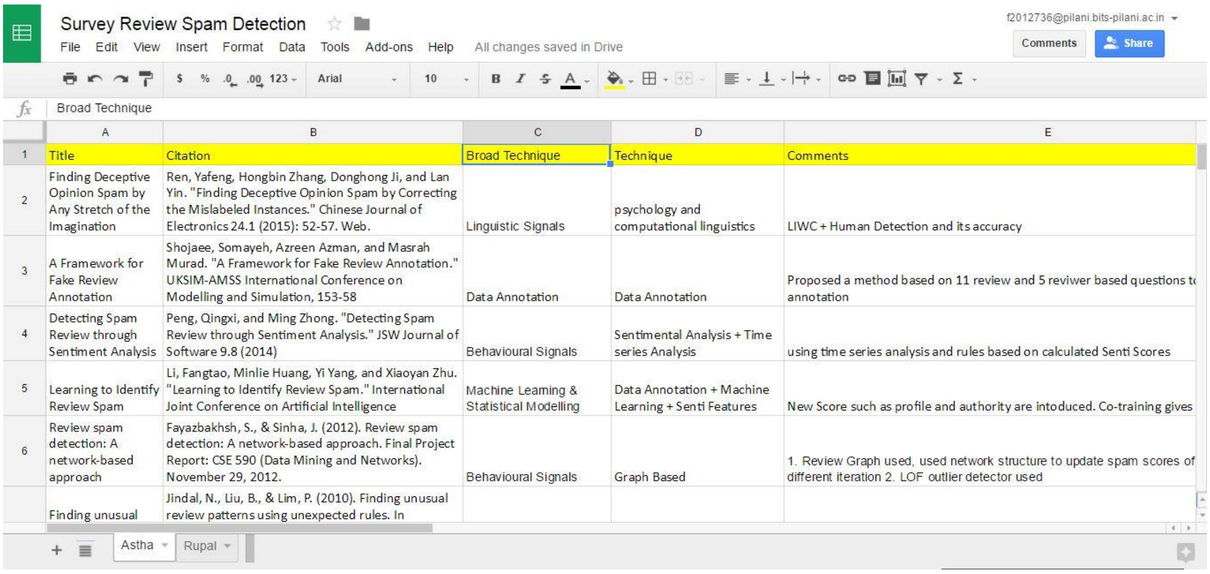
4.3 Survey for Spam Detection

Proposed technique is to create a excel sheet with Title of research paper, Citation, Broad Technique, Detail of Technique and Comments. Also side by side keeping a track of different new ideas proposed by each paper and features that a paper proposes.

Categorization of techniques and detailed categorization as mentioned above is done on the basis of theory provided by Arjun Mukherjee [17]. Different kind of analysis such as

1. Different broad topics
2. Different methods used
3. Listing features for Review centric, Reviewer centric and product centric proposed in different papers
4. A list of general trends and major focus in the area

These are some of the analysis to be done.



	Title	Citation	Broad Technique	Technique	Comments
1	Title	Citation	Broad Technique	Technique	Comments
2	Finding Deceptive Opinion Spam by Any Stretch of the Imagination	Ren, Yafeng, Hongbin Zhang, Donghong Ji, and Lan Yin. "Finding Deceptive Opinion Spam by Correcting the Mislabeled Instances." Chinese Journal of Electronics 24.1 (2015): 52-57. Web.	Linguistic Signals	psychology and computational linguistics	LIWC + Human Detection and its accuracy
3	A Framework for Fake Review Annotation	Shojaee, Somayeh, Azreen Azman, and Masrah Murad. "A Framework for Fake Review Annotation." UKSIM-AMSS International Conference on Modelling and Simulation, 153-58	Data Annotation	Data Annotation	Proposed a method based on 11 review and 5 reviwer based questions t annotation
4	Detecting Spam Review through Sentiment Analysis	Peng, Qingxi, and Ming Zhong. "Detecting Spam Review through Sentiment Analysis." JSW Journal of Software 9.8 (2014)	Behavioural Signals	Sentimental Analysis + Time series Analysis	using time series analysis and rules based on calculated Senti Scores
5	Learning to Identify Review Spam	Li, Fangtao, Minlie Huang, Yi Yang, and Xiaoyan Zhu. "Learning to Identify Review Spam." International Joint Conference on Artificial Intelligence	Machine Learning & Statistical Modelling	Data Annotation + Machine Learning + Senti Features	New Score such as profile and authority are intoduced. Co-training gives
6	Review spam detection: A network-based approach	Fayazbakhsh, S., & Sinha, J. (2012). Review spam detection: A network-based approach. Final Project Report: CSE 590 (Data Mining and Networks). November 29, 2012.	Behavioural Signals	Graph Based	1. Review Graph used, used network structure to update spam scores of different iteration 2. LOF outlier detector used
	Finding unusual	Jindal, N., Liu, B., & Lim, P. (2010). Finding unusual review patterns using unexpected rules. In			

Figure 3: Snapshot of excel sheet for survey

5. Discussion and Result

Hotel Reviews only contain text review, no other data or meta-data such as rating, time or user name is available and hence only few features could be implemented as of now. These include

1. Length
2. Percent Sentiment Word Positive
3. Percent Sentiment Word Negative
4. Percent Numerals
5. Percent Capitals
6. Percent All Capital Word

In the following results F, stands for learning on features 1,4,5 and 6. SP denotes feature 2 and SS represents sentiment Score.

Precision = TP/(TP+FP)

Therefore precision is how many were actually from positive class from the ones predicted.

Recall = TP/(TP+FN)

Recall is how many from the actual negative class were detected. Hence in our case making some negative reviews as positive (non-spam as spam) is still better than making positive reviews as negative (spam as non-spam). Therefore desire is to have good recall.

Table 1 contains numbers for different algorithms with only above 5 features, Table 2 contains results with these Six features plus sentiment score (SS) as given in section 4.2 of the report.

F+SP	Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM-SVC
TN	128	116	112	135	109
FP	72	84	88	65	91
TP	136	156	148	116	97
FN	64	44	52	84	103
Precision	0.654	0.650	0.627	0.641	0.516
Recall	0.680	0.780	0.740	0.580	0.485

Table 1: Results for Features without all the features and Sentiment Percentage

F+SS+SP	Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM
TN	127	125	129	134	107
FP	73	75	71	66	93
TP	147	144	135	115	98
FN	53	56	65	85	102
Precision	0.668	0.658	0.655	0.635	0.513
Recall	0.735	0.720	0.675	0.575	0.490

Table 2: Results for Features including all the features, Sentiment Percentage and Sentiment Score

Random forest seems to benefit from Sentiment Score especially in recall, though increment in Ada Boost and Naïve Bayes seem justified, due to benefit that we expected, but low recall is unexpected. Logistic Regression and SVM on other hand, precision for without SS is slightly better than with SS and recall is differing with just .005.

F	Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM-SVC
TN	125	120	90	136	110
FP	75	80	110	64	90
TP	142	141	149	114	98
FN	58	59	51	86	102
Precision	0.654	0.638	0.575	0.640	0.521
Recall	0.710	0.705	0.745	0.570	0.490

Table 3: Results for Features including all the features

F+SS	Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM-SVC
TN	137	124	107	136	108
FP	63	76	93	64	92
TP	137	143	141	108	98
FN	63	57	59	92	102
Precision	0.685	0.653	0.603	0.628	0.516
Recall	0.685	0.715	0.705	0.540	0.490

Table 4: Results for Features and sentiment scores

Table 5 and 6 clearly show that only Random forest benefits from SS the most, precision and recall both improves, percentage sentiment bring results down for Random Forest. Ada Boost and Naïve Bayes also do well with F+SS+SP in terms of precision. For Recall Ada Boost, Naïve Bayes seem to benefit from SP and SS brings the performance down. Logistic Regression and SVM show slight changes.

Precision	Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM
F+ SS+SP	0.668	0.658	0.655	0.635	0.513
F+SP	0.654	0.650	0.627	0.641	0.516
F	0.664	0.631	0.598	0.646	0.516
F+SS	0.653	0.657	0.606	0.633	0.516

Table 5: Precision for all the 4 sets

Recall	Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM
F+ SS+SP	0.735	0.720	0.675	0.575	0.490
F+SP	0.680	0.780	0.740	0.580	0.485
F	0.710	0.760	0.730	0.575	0.490
F+SS	0.715	0.755	0.700	0.560	0.485

Table 6: Recall for all the 4 sets

Figure 4 and 5 are graphical representation of precision and recall of the four different runs.

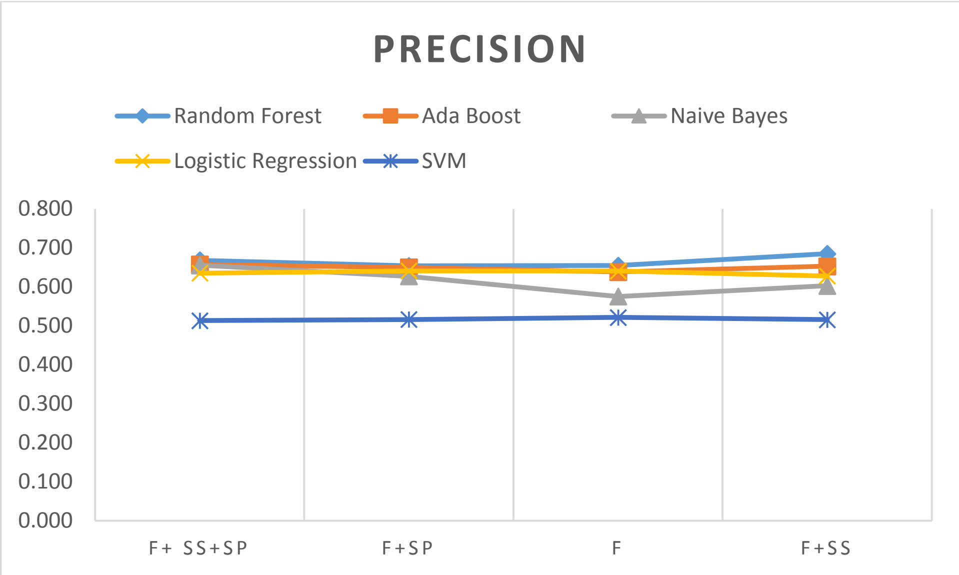


Figure 4: Line Graph for comparison of precision

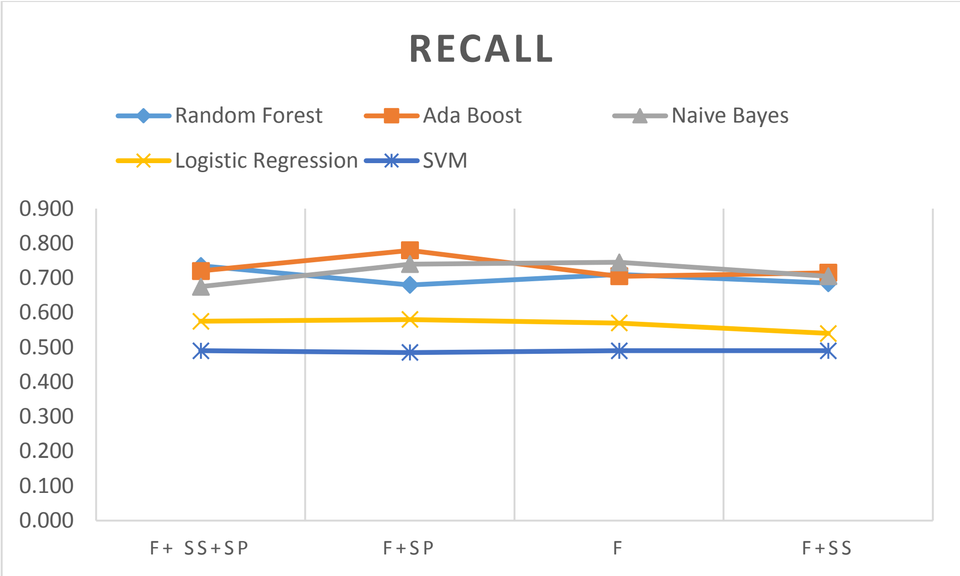


Figure 5: Line Graph for comparison of recall

6. Conclusion and Future Work

A lot of variation from the expected results in seen in the experiments. Sentiment Score helps to improve performance in precision and recall both for Random Forest. Precision is improved in Ada Boost and Naïve Bayes with help of SS (Sentiment Score) but recall suffers. Recall is improved by using SP (Sentiment Percentage). One of the hypothesis is that spams tend to have more sentiment words, since spammer does not have time to mention details of the product and hence in reviews more of sentiment is expressed.

These hypothesis needs to be checked and also concrete reasons needs to be found. Future work is to create an annotated data set, and use other Sentiment Score mentioned in the project and learn classification as an added feature. Not all features in presented by Liu in their original Paper have been implemented. Hence we need to test against them as well.

7. References

[1] Ma, Yingying, and Fengjun Li. "Detecting Review Spam: Challenges and Opportunities." Proceedings of the 8th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing(2012)

[2] Jindal, Nitin, and Bing Liu. "Analyzing and Detecting Review Spam." Seventh IEEE International Conference on Data Mining (ICDM 2007), 2007

[3] Peng, Qingxi, and Ming Zhong. "Detecting Spam Review through Sentiment Analysis." JSW Journal of Software 9.8 (2014)

[4] Li, Fangtao, Minlie Huang, Yi Yang, and Xiaoyan Zhu. "Learning to Identify Review Spam." International Joint Conference on Artificial Intelligence

[5] Hu, Xia, Jiliang Tang, Huiji Gao, and Huan Liu. "Social Spammer Detection with Sentiment Information." 2014 IEEE International Conference on Data Mining, 2014

[6] Jindal, Nitin, and Bing Liu. "Analyzing and Detecting Review Spam." Seventh IEEE International Conference on Data Mining (ICDM 2007),2007

[7] Collomb, Ana• s, and Crina Costea. "A Study and Comparison of Sentiment Analysis Methods for Reputation Evaluation."

[8] Liu, Bing, and Lei Zhang. "A Survey of Opinion Mining and Sentiment Analysis." Mining Text Data (2012): 415-63

[9] Heydari, Atefeh, Mohammad Ali Tavakoli, Naomie Salim, and Zahra Heydari. "Detection of Review Spam: A Survey." Expert Systems with Applications 42.7 (2015): 3634-642. Web.

[10] Crawford, Michael, Taghi M. Khoshgoftaar, Joseph D. Prusa, Aaron N. Richter, and Hamzah Al Najada. "Survey of Review Spam Detection Using Machine Learning Techniques." Journal of Big Data 2.1 (2015)

[11] M. Ott, Y. Choi, C. Cardie, and J.T. Hancock. 2011. 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.

[12] M. Ott, C. Cardie, and J.T. Hancock. 2013. Negative Deceptive Opinion Spam. In Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.

[13] Jindal and Liu, WWW-2007, WSDM-2008; Lim et al, CIKM-2010; Jindal, Liu and Lim, CIKM-2010; Mukherjee et al. WWW-2011; Mukherjee, Liu and Glance, WWW-2012

[14] Lim, Ee-Peng, Viet-An Nguyen, Nitin Jindal, Bing Liu, and Hady Wirawan Lauw. "Detecting Product Review Spammers Using Rating Behaviors." Proceedings of the 19th ACM International Conference on Information and Knowledge Management - CIKM '10 (2010): n. pag. Web.

[15] Shojaee, Somayeh, Azreen Azman, and Masrah Murad. "A Framework for Fake Review Annotation." UKSIM-AMSS International Conference on Modelling and Simulation, 153-58

[16] Ren, Yafeng, Hongbin Zhang, Donghong Ji, and Lan Yin. "Finding Deceptive Opinion Spam by Correcting the Mislabeled Instances." Chinese Journal of Electronics 24.1 (2015): 52-57. Web.

[17] Mukherjee, Arjun. "Detecting Deceptive Opinion Spam Using Linguistics, Behavioral and Statistical Modeling." Tutorials (2015): n. pag. Web.

[18] Feng, Song, Ritwik Banerjee, and Yejin Choi. "Syntactic stylometry for deception detection." Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2. Association for Computational Linguistics, 2012.

[19] Mukherjee, A., Kumar, A., Liu, B., Wang, J., Hsu, M., Castellanos, M. and Ghosh, R., 2013, August. Spotting opinion spammers using behavioral footprints. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 632-640). ACM.

[20] Mukherjee, Arjun. "Detecting Deceptive Opinion Spam Using Linguistics, Behavioral and Statistical Modeling." Tutorials (2015): n. pag. Web.

[21] Mukherjee, A., Liu, B., & Glance, N. (2012). Spotting fake reviewer groups in consumer reviews. In Proceedings of the 21st international conference on World Wide Web. ACM.

[22] WentaoTian, Y.B. and Lam, W., 2015. A unified model for unsupervised opinion spamming detection incorporating text generality.

8. Appendix

8.1 Hotel reviews Data : http://myleott.com/op_spam/
8.2 Amazon Product Review datahttp://liu.cs.uic.edu/download/data/
8.3 GENERIC READING FOR SPAM DETECTION POINTS

REVIEW CENTRIC

- Empty Adjectives Reviews: Reviews that have zero caveats and with empty adjectives with praise and no downsides.
- Brand Language Reviews: Using "brand approved" language such as a particular name version. Something that No normal person would be bothered to write but a marketer would.
    - Product Name in all caps
    - Same language as product description
    - Writing Entire product name multiple times
    - Giving mainly tally of features and no view from reviewers side (Real users tend to write about performance, reliability, durability and overall value)
- Copy Paste Description Reviews : Mostly using the same buzzwords that is given in product description or website of its products/services (or broken English given on webisite)
- The company reviews: Review talks about the company as a whole and not much about the product. Example "Company A wouldn't do that" or "Company B makes quality products". Normally people would say such statements by telling how the problem was dealt with or explain what can be done to fix it.
- Very High and Low rating anomaly: Negative reviews can provide a very good insight in the potential problems that can arise. But if review contains something very lazy such as "This sux!" , it can be categorized as spam.
- Hyperventilating negative reviews: Can be sign of a company jacking their competitor or of mental. For Example, "DOES NOT WORK!!!!! Broke after 10 seconds, so I got an (name of competitor) instead, and It works like magic!"
- Link Reviews: Link in reviews can be marked as spam. It is mostly the case that the review was left just to post the link and also review itself is useless."
- cookie-cutter reviews : Many times first reviews are cookie-cutter to highlight a peculiar phrasing, by doing a search we can find if similar cookie-cutters were used
- Lazy Reviews: Mentioning overall sentiment and just one or two things about the attribute. Example "Food at xyz is brilliant, they have the best [one of the mid ranged items] around!". Similar reviews for other places can be found by same reviewer.
- Disappointed and Switch Reviews: Starting on a disappointed note about something and then suddenly realizing the light and wanted to spread word about it. Example "I didn't like xyz product but my friend said I should try one, and it turns out is amazing".
- No Information :
    - Involving a third person, contradictory statement should be carefully looked at. Example "Guacamole like my grandmother used to make!, it was a little pricey, but worth due to the large serving sizes and great service!"
- Falsely Proven Reviews:

- Claim to be clinically proven but names of the scientists and scientific journals are fraudulent
    - Gushing Testimonials from customers who were at very first sceptical, but then they researched or found out from sources that the product indeed worked.
- Long-winded explanations why one product is way better than other when in reality they are identical
- Red flag words: Words such as “Treat” and “Recommend” or any variation should raise eyebrows, since they tend to be written by spammers to promote the product.

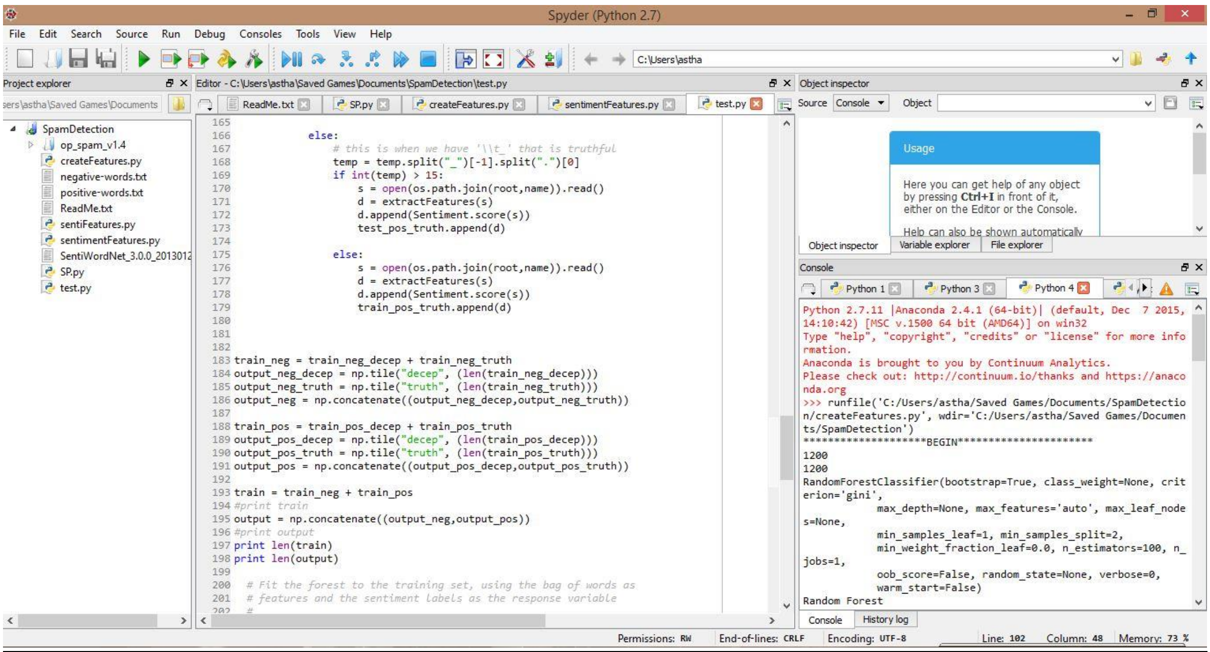
REVIEWER CENTRIC

- Names of reviewers are variations each other, i.e. varman1234, varman12345
- Polyamorous Reviewer: every product an unvarnished review
- Monogamous Reviewer: Reviews only for products by one brand or manufacturer.
- Username has more than 3 numbers: Especially if may other reviews are by accounts with more than 3 numbers in there username. Usual sign of automated program.
- Accounts that are created around the time domain name was registered is usually indication of spammer accounts. whois.net can be used to find out time of registration.
- Franchise Operation: a reviewer have reviewed all other locations of chain within a few days of time spam another.

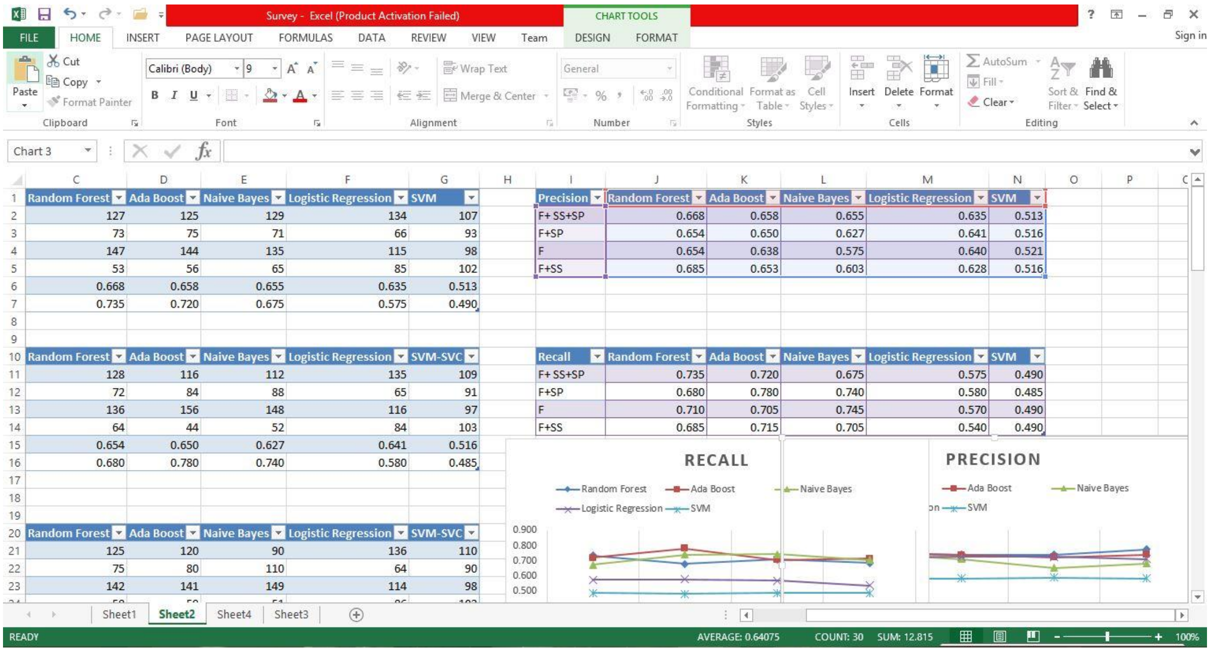
SOURCES

1. https://consumerist.com/2010/04/06/spot-fake-online-reviews/
2. https://consumerist.com/2010/04/14/how-you-spot-fake-online-reviews/
3. http://www.mybestbuddymedia.com/2015/06/9-quick-ways-to-detect-online.html

8.4 Screen Shot of runs



8.5 Screenshot of excel sheet


Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM
127	125	129	134	107
73	75	71	66	93
147	144	135	115	98
53	56	65	85	102
0.668	0.658	0.655	0.635	0.513
0.735	0.720	0.675	0.575	0.490

Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM-SVC
128	116	112	135	109
72	84	88	65	91
136	156	148	116	97
64	44	52	84	103
0.654	0.650	0.627	0.641	0.516
0.680	0.780	0.740	0.580	0.485

Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM-SVC
125	120	90	136	110
75	80	110	64	90
142	141	149	114	98

Precision	Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM
F+ SS+SP	0.668	0.658	0.655	0.635	0.513
F+SP	0.654	0.650	0.627	0.641	0.516
F	0.654	0.638	0.575	0.640	0.521
F+SS	0.685	0.653	0.603	0.628	0.516

Recall	Random Forest	Ada Boost	Naive Bayes	Logistic Regression	SVM
F+ SS+SP	0.735	0.720	0.675	0.575	0.490
F+SP	0.680	0.780	0.740	0.580	0.485
F	0.710	0.705	0.745	0.570	0.490
F+SS	0.685	0.715	0.705	0.540	0.490