ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

A SURVEY: SPAM FILTERING USING MACHINE LEARNING ALGORITHMS

Deepak Kumar Agarwal*
* M.Tech CS-IInd year ABES Engineering College Ghaziabad, India.

ABSTRACT

In this paper, we show an exhaustive survey of late improvements in the utilization of machine learning algorithms to Spam filtering, concentrating on both textual-and image-based approaches. Rather than considering Spam filtering as a standard classification issue, we highlight the significance of considering particular attributes of the issue, particularly idea float, in outlining new filters. Two especially vital viewpoints not generally perceived in the writing are talked about: the troubles in overhauling a classifier taking into account the sack of-words representation and a noteworthy contrast between two early credulous Bayes models. By and large, we presume that while vital headways have been made in the most recent years, a few perspectives stay to be investigated, particularly under more reasonable assessment settings.

KEYWORDS: Spam filtering, Support Vector Machine, Naive Bayes, Artificial Neural Network.

INTRODUCTION

As of late, the expanding utilization of e-mail has prompted the development and further heightening of issues brought on by spontaneous mass e-mail messages, regularly alluded to as Spam. Advancing from a minor aggravation to a noteworthy concern, given the high circling volume and hostile substance of some of these messages, Spam is starting to decrease the dependability of e-mail (Hoanca, 2006). Individual clients and organizations are influenced by Spam because of the network data transmission squandered accepting these messages and the time spent by clients recognizing Spam and ordinary (authentic or ham) messages. A business model depending on Spam showcasing is normally beneficial in light of the fact that the expenses for the sender are little, so that a substantial number of messages can be sent, expanding the profits, this forceful conduct being one of the characterizing attributes of Spammers (those that send Spam messages) (Martin-Herran, Rubel, and Zaccour, 2008). The efficient effects of Spam have driven some countries to embrace enactment (e.g., Carpinter and Hunt, 2006; Hoanca, 2006; Stern, 2008), despite the fact that it is constrained by the way that numerous such messages are sent from different countries (Talbot, 2008). In addition, challenges in following the real senders of these messages can likewise constrain the utilization of such laws. Notwithstanding enactment, a few creators have Proposed changes in conventions and operation models (talked about in Hoanca (2006)). Another methodology received is the utilization of Spam filters, which, in light of examination of the message substance and extra information, endeavor to recognize Spam messages. The move to be made once they are recognized generally relies on upon the setting in which the filter is connected. In the event that utilized by a solitary client, as a customer side filter, they are normally sent to an envelope which contains just Spam-named messages, making the recognizable proof of these messages less demanding. Conversely, if the filter works in a mail server, taking care of messages from a few clients, they might either be marked as Spam or erased. Another plausibility is a community setting, in which filters running in distinctive machines offer information on the messages got, to enhance their execution. On the other hand, the utilization of filters has made a developmental situation (Goodman, Cormack, and Heckerman, 2007; Hayes, 2007), in which Spammers utilize apparatuses (Stern, 2008) with different procedures particularly custom-made to minimize the quantity of messages distinguished. At first, Spam filters depended on client characterized tenets, composed in view of information of regularities effortlessly saw in such messages. Accordingly, Spammers then started utilizing content "confusion" (or darkening), by camouflaging certain terms that are exceptionally normal in Spam messages (e.g., by composing "f r 3" rather than "free"), on an endeavor to keep the right recognizable proof of these terms by Spam filters. These days, Spam filtering is typically handled by machine learning (ML) algorithms, went for segregating in the middle of true blue and Spam messages, giving an automated, adaptive methodology, which are the center of this audit. Rather than depending close by coded rules, which are inclined to the always showing signs of change nature of Spam messages, ML methodologies are equipped for extracting knowledge from an arrangement of messages supplied, and utilizing the got information as a

part of the classification of recently got messages. At last, albeit spontaneous substance is flow influencing e-mail, as well as internet searchers (Gyongyi and Garcia-Molina, 2005) and websites (Kolari, Java, Finin, Oates, and Joshi, 2006), this study concentrates singularly on managing e-mail Spam.

SUPPORT VECTOR MACHINES

support Vector Machines (SVM) (Scholkopf and Smola, 2002; Vapnik, 1998) were at first connected by Drucker, Wu, and Vapnik (1999), utilizing the BoW representation with binary, frequency or tf-idf features, chose by information gain, and two private corpora. It was confirmed that boosting with decision trees accomplished a marginally bring down false positive rate than SVM, however the recent being more vigorous to distinctive datasets and pre-processing methodology, and considerably more for effective for training (as a rule, somewhere around one and two requests of greatness). The best results were acquired utilizing a binary representation for SVM and frequency-based for boosting. At last, as boosting consequently chooses the quantity of features, and SVM are fit for managing an extensive number of features, no feature determination should be performed.

In a server-side filter, a vital trouble confronted are the distinctive attributes of Spam and genuine messages got by diverse clients. For this situation, the utilization of openly accessible datasets for training a solitary classifier for all clients seriously predispositions the classification issue. Concentrating on this situation, Bickel and Scheffer (2007) added to a system for learning a classifier utilizing openly accessible (marked) and client (unlabeled) messages. For the situation that n clients are as of now subscribed, the classifier for another client is gotten via training a linear SVM, where the misclassification expense of the messages depends on the evaluated predisposition rectification term of each message. The trial results were gotten utilizing a binary representation, the Enron corpus and Spam messages from different open and private sources. It was confirmed that the plan proposed diminished the danger in up to 40%, in examination with the client of a solitary classifier for all clients. Considering that numerous Spam and genuine messages are created by layout, Haider, Brefeld, and Scheffer (2007) considered the use of a SVM-based incremental regulated grouping algorithm for distinguishing, in light of the substance, such messages. Examinations were directed with messages from a private and the Enron corpora, including pamphlets. Utilizing a frequency-based linear SVM, in which the cluster information was fused through four extra features, it was confirmed that this extra information lessened the danger by up to 40%. Also, the incremental rendition proposed required an execution time more than six requests of extent lower than that of a cluster based. Kanaris et al. (2007) utilized character n-grams, when n is prespecified and a variable plan, to prepare a linear SVM utilizing a binary or frequency representation, utilizing the information gain for feature choice. Investigations depended on the LingSpam and SpamAssassin datasets, 10-fold cross-validation and 3-, 4-and 5-gram models. The n-gram models accomplished better or similar results than a word-based representation, with variable n being unrivaled in an expense delicate situation. Sculley and Wachman (2007a) connected linear SVMs in a web setting, utilizing a binary representation and no feature determination. The creators proposed the utilization of character n-gram models (n ¼ 3 and n ¼ 4), such that all the conceivable features are known from the earlier. In this detailing, the quantity of conceivable features has a tendency to be substantial, yet the creators pointed out (Sculley and Wachman, 2007b) that the quantity of exceptional 4-grams watched was on the request of millions in each of the TREC corpora. As far as possible the quantity of features utilized, just the initial 3000 characters of each message were utilized. Trials directed with the TREC2005 and 2006 corpora demonstrated that the online SVMs acquire extremely focused results, particularly when n ¼ 4, at an expense of a long execution time. The creators proposed a rearranged detailing, termed loose online SVMs, which permitted somewhere around 5-and 10-fold diminishments in the execution time, without impressive punishments in the classification execution. It was presumed that online SVMs perform extremely well, and that the improvements proposed make their expansive scale application practical. Sculley (2007) considered an online active learning system, where a filter arranges a message and can then demand its actual mark for overhauling the model, the goal being the boost of the classification execution while minimizing the quantity of names asked. Three hyperplane-based classifiers, to be specific perceptron, perceptron with margins and loose online SVM (Sculley& Wachman, 2007a), were concentrated on, alongside three systems for choosing to ask for the mark. Two of these strategies, b-Sampling (BS) and Logistic Margin Sampling (LMS), are probabilistic, asking for the mark with a sure likelihood as an element of the separation to the isolating hyperplane. In Fixed Margin Sampling (FMS), the mark is asked for if the separation is littler than an edge. Utilizing the same test setup as in Sculley and Wachman (2007a), perceptron performed the most exceedingly bad, while SVMs acquired the best results, with LMS and FMS beating BS. Then again, when the quantity of asked for names was huge, the three active learning techniques performed comparatively

ARTIFICIAL NEURAL NETWORK

Clark, Koprinska, and Poon (2003) proposed LINGER, which utilizes a multi-layer perceptron (Haykin, 1998) for e-mail order and Spam filtering. Messages are spoken to as BoW, with feature selection taking into account information gain or term-frequency change. Analyses were led utilizing the LingSpam and PU1 corpora, notwithstanding a private dataset, with 256 features and 10-fold cross-validation. Utilizing the information gain, LINGER acquired flawless results, and beat a naive Bayes classifier when utilizing term-frequency change, however with somewhat all the more false positives. Nonetheless, when the filter was prepared and tried with diverse datasets, it was checked that the outcomes were inadmissible. It was inferred that neural networks can be effectively connected to Spam filtering and e-mail arrangement, and that more investigations are expected to assess the versatility of filters. Luo and Zincir-Heywood (2005) utilized a Self-Organizing Map (SOM) (Kohonen, 2001) for grouping examination (SBSA), which considers the request of words in a given message. The tests utilized the LingSpam corpus with the information gain for feature determination, in an expense delicate setting. It was checked that SBSA can legitimately encode and utilize the arrangement of words in a message to group it, outflanking the naive Bayes classifier, particularly as far as false positives. Notwithstanding, it ought to be noticed that feature choice does not consider the request of words, which can contrarily influence the execution of the network. Wang, Jones, and Pan (2006) connected two online linear classifiers, Perceptron and Winnow, which depend on the determination of parameters of an isolating hyperplane. Other than being quick, adjusting the models with new information is straightforward, albeit no online assessment occurred in the work. Investigations were led utilizing the LingSpam and PU1 corpora, 10-fold cross-validation and considering the F1 measure for execution assessment. At first, it was confirmed that utilizing the information gain or report frequency for feature determination prompted comparative results, with chances proportion performing the most noticeably awful. Winnow and Perceptron got comparative results for an extensive variety of number of features, with the previous marginally beating the last, and a naive Bayes classifier performing the most noticeably awful, particularly for countless. By and large, the creators pointed out that both algorithms performed exceptionally well, being vigorous to the quantity of features, emphases and illustrations utilized for training. Winnow acquired marginally enhanced results in examination with the Perceptron, with both impressively outflanking the naive Bayes classifier.

Tzortzis and Likas (2007) considered the use of deep belief networks, neural networks with a few hidden-layers, for which an effective training algorithm was as of late created. Analyses were directed utilizing the LingSpam and SpamAssassin corpora, and a subset of the Enron corpus, with a frequency based representation of the 1000 or 1500 features chose in view of the information gain. It was confirmed that the deep belief networks with three hidden-layers got somewhat preferable results over a SVM utilizing a cosine kernel, with better exactness in each of the three corpora. Rather than breaking down the message substance, Wu (2009) defined a model considering the exchange for the conveyance of each message, as quite a bit of this information has a tendency to be faked or overlooked on account of Spam messages. This information is utilized to prepare a neural network, went for separating between conduct regular of Spam and true blue messages. From an arrangement of six header and four server log document handle, a sum of 26 literary features were determined. Trials depended on an individual corpus, isolated by time in five gatherings. The framework accomplished low false positive and negative rates, and with better results in correlation to utilizing the same building design for substance based classification utilizing 3000 features spoke to as tf-idf. At long last, in an examination with 16 content based routines, the proposed strategy accomplished the best results as far as both false positive and negative rates.

LOGISTIC REGRESSION

Goodman and Yih (2006) connected a logistic regression model (e.g., Hastie, Tibshirani, and Friedman, 2001), which is straightforward and can be effectively redesigned. It utilizes binary features, recognized taking into account whether they happen on the message-headers or body, without the use of feature choice. It was confirmed that, in investigations with the TREC2005 and Enron datasets, notwithstanding a private corpus, the outcomes got were focused or even better than a percentage of the best known filters for every corpus. Zhou (2007) outlined a versatile coding-based system, in which the structure of messages is put away in a Huffman tree. It utilizes one tree for every class, putting away the words and their number of events in messages, producing a feature vector encoding these terms in the tree. A logistic regression model is then prepared utilizing the features of haphazardly, especially later, examined messages. In trials utilizing the LingSpamcorpus, it accomplished the best general results in examination with a few algorithms, including SVM and boosting trees, in spite of the fact that the outcomes exhibited considered just the exactness. It was inferred that the proposed methodology is exceptionally productive, right around three requests of greatness quicker than SVM, powerful and stronger to idea drift and unbalanced information. Jorgensen, Zhou, and Inge (2008)

defined a numerous occasion learning methodology for managing great word assaults, portrayed by the incorporation in a Spam message of words that are run of the mill of honest to goodness messages. A sample is spoken to by an arrangement of occurrences, and is grouped relying upon the classification of every case, with four systems for producing the examples being proposed. A various occasion logistic regression (MILR) model is utilized to consolidate the classification of every example. Examinations were directed utilizing the TREC2006 corpus, sequentially separated into 11 sections, in a web setting, where the filters are prepared on one section and used to order the following. The four techniques were contrasted and a solitary occurrence logistic regression model, a SVM and naive Bayes, with 500 features chose in view of the information gain and messages spoke to utilizing tf-idf. It was checked that the attack recognizably corrupted the genuine positive rates of all the filters, which was lessened by up to half, with the exception of one of the MILR-based systems. At the point when the classifiers were permitted to prepare on camouflaged Spam messages, the impacts of the attack were lessened, with MIRL having the best execution. It was presumed that the created systems are compelling, notwithstanding when training with the hidden messages is impractical.

ARTIFICIAL IMMUNE SYSTEM

Oda and White (2003a) proposed an Artificial Immune System (see, e.g., de Castro and Timmis, 2002) for Spam filtering, where detectors, spoke to as customary expressions, are utilized for pattern-matching as a part of a message being broke down. It relegates a weight to every indicator, which is augmented (decremented) when it perceives an expression in a Spam (real) message, with the threshold whole of the weights of the coordinating locators being utilized to decide the classification of a message. The framework can be revised by either increasing or decrementing the weights of all the coordinating identifiers. Utilizing an individual corpus and 100 identifiers, created for the most part from SpamAssassin heuristics, genuine positive and negative rates of 90% and 99%, separately, were gotten. It was reasoned that the proposition accomplished satisfactory results considering the little number of identifiers utilized. Later on, the same creators (Oda and White, 2003b) thought about two different strategies for deciding the classification of a message, utilizing the SpamAssassin corpus and around 150 identifiers. The main strategy is like that beforehand utilized, aside from that the weights are increased just when examples in Spam messages are perceived, while the second system weights the quantity of expressions perceived in Spam and all the messages. It was inferred that the second system is more fitting, because of its more prominent general classification accuracy, regardless of somewhat higher false positive rates. Oda and White (2005) then performed extra trials utilizing the SpamAssassin corpus, utilizing likewise a Bayesian mix of finder weights. Notwithstanding heuristics usually utilized in Spam filters, a dictionary of words, and examples separated from an arrangement of messages were considered for creating the indicators. To decide the classification of a message, the second technique proposed in Oda and White (2003b) stayed as the best option, with the Bayes-roused methodology performing somewhat more terrible, in spite of the fact that the outcomes were investigated just as far as the general blunder rate. At long last, it was watched that utilizing the heuristics drove, without a doubt, to the best results, with the other two strategies performing correspondingly

COMPARATIVE STUDIES

Different learning ideal models, and challenges in looking at filters construct singularly with respect to execution figures, a few works have been committed to contrasting diverse filters under the same conditions. These works can give not just a comprehension of the best performing algorithms in specific cases, additionally light up different parts of the issue, for example, the significance of considering extra message features other than the body, for example, the headers. Schneider (2003) analyzed two naive Bayes models in view of diverse occasion models utilizing the LingSpam and PU1 corpora. The multivariate Bernoulli model considers just the event of terms, while the multinomial model utilizes their frequency. Given a message to be grouped, the previous considers, notwithstanding the event of the chose terms, the nonattendance of the remaining terms. In both models, feature determination was led in view of the information gain. Utilizing 10-fold cross-validation, the multivariate Bernoulli model accomplished lower false positives and higher false negatives, with the multinomial methodology having a more adjusted execution. In general, it was reasoned that the multinomial model is stronger to skewed class appropriations, accomplishing a higher general exactness. Androutsopoulos et al. (2004) thought about the execution of credulous Bayes, Flexible Bayes, SVMs and boosting on the four PU corpora utilizing 10-fold cross-validation. Messages were encoded utilizing a frequency-based representation, with the information gain for feature determination. On account of k ¼ 1, all algorithms had a comparable execution as far as the weighted exactness, with boosting and SVM being slower to prepare however speedier while classifying, while the Bayesian classifiers had the inverse conduct. On the other hand, credulous Bayes had an inadmissible execution for k ¼ 9. The creators then introduced Filtron, a learning-based filter, which was

utilized by one of the creators for seven months for classifying around 6700 messages, 75% being honest to goodness. Utilizing a SVM as learning algorithm, prepared on the PU3 corpus with k ¼ 1, genuine positive and negative rates of 89% and 99%, separately, were gotten. The filter was not upgraded amid the entire time, and it was checked that the vast majority of the inaccurately arranged Spam messages contained either next to zero content, were composed in dialects other than English or were encoded. With regards to the authentic messages mistakenly arranged, half were newsletters or automatic reactions. Ozgur, Gungor, and Gurgen (2004) explored the utilization of Spam filters in light of Artificial Neural Networks (ANN) and naive Bayes for Turkish messages, utilizing a module for morphological investigation. For both algorithms, a standard BoW representation was embraced, utilizing the information gain for feature determination. Examinations were led with an individual corpus of Turkish messages and six-fold cross-validation. In a first analysis, it was confirmed that the frequency-based representation, with 40 or 70 features, created the best results for the ANN, with a linear network having the best exchange off in the middle of adequacy and execution time. Then again, the binary representation created the best results for the naive Bayes classifier, unless the quantity of features was too expansive. It was reasoned that the routines constitute a fascinating way to deal with examining Turkish messages, with adequate right classification rates. Zhang et al. (2004) analyzed the execution of five Spam filters utilizing the SpamAssassin, LingSpam and PU1 corpora, notwithstanding the Chinese corpus ZH1 assembled by the creators. In the examinations, a binary representation, alongside 10-fold cross-validation, was utilized, with the information gain, record frequency and v2-test for feature choice and the TCR as execution measure. With respect to the feature determination strategy, it was confirmed that the information gain prompted the best results, trailed by the v2-test. Also SVM, AdaBoost and logistic regression model achieved the best general results, with naive Bayes and an apathetic learning methodology not being practical in an expense delicate situation with k ¼ 999. It was additionally watched that utilizing just the message headers prompted predominant or if nothing else comparable results as utilizing just the message body, while the mix of both accomplished the best execution. The creators reasoned that the BoW representation of messages is viable additionally to classify Chinese messages and that the great results acquired while considering the message headers highlight the significance of considering this information in a filter. Webb and Chitti (2005) assessed four Spam filters, to be specific (linear) SVM, regression-based boosting and two naive Bayes classifiers (SpamProbe and a standard variant utilizing a multivariate Bernoulli model), utilizing a vast scale corpus collected by the creators with more than 1 million messages. For feature determination, the information gain was utilized. The message headers were not utilized, as it was pointed out that utilizing them has a tendency to present a generous inclination on the classification, clarifying to some degree the upgrades normally watched when they are considered. It was checked that SpamProbe, SVM and boosting performed correspondingly, trailed by naive Bayes, demonstrating that the filters perform extremely well. An assault setting was then considered, in which disguised Spam messages containing parts of genuine messages were made, gone for corrupting the execution of the filters. At the point when the classifiers were prepared utilizing "typical" Spam and honest to goodness messages, they can't distinguish a significant number of the covered messages, however when some of these disguised messages were utilized for training, the precision of the classifiers was restored. Taking everything into account, this assault can impressively influence the execution of the filters, yet can be taken care of by utilizing some of these Spam messages for training or redesigning the classifiers. Metsis et al. (2006) thought about the execution of five credulous Bayes classifiers. For binary qualities, a multivariate Bernoulli and a multinomial model were considered. With a frequency-based representation, a multinomial model was utilized. In addition, two genuine esteemed models, multivariate Gauss and Flexible Bayes, were considered, which utilize a standardized frequency representation. Tests were led with messages from the Enron-Spam corpus, sequentially isolated into subsets containing 100 messages, and afterward utilized as a part of a web setting, where the classifiers were prepared on all the already seen messages and used to classify the following subset. It was confirmed that the best results were acquired when utilizing 3000 features, chose on the premise of the information gain, with Flexible Bayes and the binary multinomial model, while the multinomial frequency and multivariate Bernoulli models performed the most exceedingly awful. The creators presumed that, because of its less complex execution and smoother exchange off between genuine positive and negative rates, the multinomial binary model seems, by all accounts, to be the best variation. Yu and Ben Xu (2008) thought about the execution of four algorithms: neural networks, naive Bayes, SVM and Relevance Vector Machine (RVM) (Tipping, 2001). In examination with SVM, the recent does not require the change of parameters and results in a normally more meager arrangement, along these lines bringing about quicker testing. On the other hand, it includes the arrangement of a non-curved optimization issue, having a tendency to be slower for training. For the neural network, binary features were utilized as information, while, in the SVM, a sigmoid kernel was utilized, in spite of the fact that it ought to be noticed that, in this setting, the optimization issue may get to be non-curved, because of the way that this kernel is not positive unequivocal for certain parameter values (e.g. Scholkopf and Smola, 2002, p. 45). The test assessment was directed utilizing the SpamAssassin corpus and a private dataset, with features chose utilizing LINGER (Clark et al., 2003). It was confirmed that SVM and RVM have a comparative execution, with the decision between them relying

upon the requirement for quicker training (SVM) or testing (RVM), trailed by naive Bayes, and with the neural network performing the most exceedingly terrible, albeit just the accuracies were exhibit

CONCLUSION

In this paper, an extensive survey of late machine learning ways to deal with Spam filters was introduced. A quantitative examination of the utilization of feature determination algorithms and datasets was led. It was checked that the information gain is the most regularly utilized strategy for feature determination, in spite of the fact that it has been proposed that others (e.g., the term-frequency change, in Koprinska et al. (2007)) may prompt enhanced results when utilized with certain machine learning algorithms. Among the few openly accessible datasets, the LingSpam and SpamAssassin corpora stand as the most mainstream, while the late TREC corpora, which endeavor to replicate a reasonable, web, setting, are tolerably prevalent at present. Regarding assessment measures, the genuine positive and negative rates, which are given, separately, by the relative number of Spam and honest to goodness messages accurately ordered, are proposed as the favored files for assessing filters, particularly as ROC bends (Fawcett, 2006). Two imperative perspectives not generally perceived in the writing were examined. Albeit most algorithms speak to messages as sack of-words, it ought to be precisely utilized, as it forces an extreme inclination in the issue. This is because of that reality that redesigning a model to consider new terms, which were not at first accessible, can be a feeble point, as it as a rule requires re-constructing the classifier starting with no outside help.

REFERENCES

[1] Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues,methodological variations, and system approaches. Artificial IntelligenceCommunication, 7(1), 39–59.
[2] Abi-Haidar, A., & Rocha, L. M. (2008). Adaptive spam detection inspired by a crossregulationmodel of immune dynamics: A study of concept drift. Lecture Notes inComputer Science, 5132.
[3] Aha, D. W. (1997). Lazy learning. Artificial Intelligence Review, 11(1–5), 7–10.
[4] Androutsopoulos, I., Koutsias, J., Chandrinos, K. V., Paliouras, G., & Spyropoulos, C.(2000). An evaluation of naive Bayesian anti-spam filtering. In G. Potamias, V.Moustakis, & M. van Someren, M. (Eds.), Proc of the 11th Eur conf on mach learn.
[5] Androutsopoulos, I., Koutsias, J., Chandrinos, K. V., & Spyropoulos, C. D. (2000). Anexperimental comparison of naïve Bayesian and keyword-based anti-spamfiltering with personal e-mail messages. In Proc of the ann int ACM SIGIR conf onres and devel in inform retrieval.
[6] Androutsopoulos, I., Paliouras, G., & Michelakis, E. (2004). Learning to filterunsolicited commercial e-mail. Tech. rep. 2004/2, NCSR ‘‘Demokritos”.
[7] Aradhye, H., Myers, G., & Herson, J. (2005). Image analysis for efficientcategorization of image-based spam e-mail. In Proc int conf doc analysis andrecog (Vol. 2).
[8] Bezerra, G. B., Barra, T. V., Ferreira, H. M., Knidel, H., de Castro, L. N., & Zuben, F. J. V.(2006). An immunological filter for spam. Lecture Notes in Computer Science,4163, 446–458.
[9] Bickel, S., & Scheffer, T. (2007). Dirichlet-enhanced spam filtering based on biasedsamples. Advances in Neural Information Processing System, 19, 161–168.
[10] Biggio, B., Fumera, G., Pillai, I., & Roli, F. (2007). Image spam filtering using visualinformation. In Proc int conf on image analysis and proc.
[11] Biggio, B., Fumera, G., Pillai, I., & Roli, F. (2008). Improving image spam filteringusing image text features. In Proc of the fifth conf on email and anti-spam.
[12] Blanzieri, E., & Bryl, A. (2008). A survey of learning-based techniques of email spamfiltering. Tech. rep. DIT-06-056, University of Trento, Information Engineeringand Computer Science Department.
[13] Bratko, A., Filipic, B., Cormack, G. V., Lynam, T. R., & Zupan, B. (2006). Spam filteringusing statistical data compression models. Journal of Machine Learning Research,7, 2673–2698.
[14] Byun, B., Lee, C.-H., Webb, S., & Pu, C. (2007). A discriminative classifier learningapproach to image modeling and spam image identification. In Proc of the fourthconf on email and anti-spam.
[15] Camastra, F., & Verri, A. (2005). A novel kernel method for clustering. IEEETransactions on Pattern Analysis and Machine Intelligence, 27(5), 801–805.
[16] Carpinter, J., & Hunt, R. (2006). Tightening the net: A review of current and nextgeneration spam filtering tools. Computers and Security, 25(8), 566–578.Carreras, X., & Marquez, L. (2001). Boosting trees for anti-spam email filtering. InProc of the fourth int conf on recent adv in nat lang proc.
[17] Ciltik, A., & Gungor, T. (2008). Time-efficient spam e-mail filtering using n-grammodels. Pattern Recognition Letters, 29(1), 19–33.

[18] Clark, J., Koprinska, I., & Poon, J. (2003). A neural network based approach toautomated e-mail classification. In Proc of the IEEE/WIC int conf on web intell.
[19] Cormack, G. V. (2006). TREC 2006 spam track overview. In Proc of TREC 2006: The15th text retrieval conf.Cormack, G. V. (2007). TREC 2007 spam track overview. In Proc of TREC 2007: The16th text retrieval conf.
[20] Cormack, G. V., & Lynam, T. (2005). TREC 2005 spam track overview. In: Proc of TREC2005: The 14th text retrieval conf.
[21] Cormack, G. V., & Lynam, T. R. (2007). Online supervised spam filter evaluation. ACMTransactions on Information Systems, 25(3), 11.