

Spam Filtering Methods and machine Learning Algorithm - A Survey

Abha Tewari
Student, ME
VESIT

Smita Jangale
Associate Professor
VESIT

ABSTRACT

Social networking websites are used by millions of people around the world. People express their views, opinions and share current topics. Millions of data generated every day.

It's a good platform to connect with the people. Now a day's spammers used this platform to advertise spam content on the social networking websites.

The proposed system used to classify tweets into different groups as spam and non spam tweets .The system will use 120 character tweets for analysis purpose. Various active and verified twitter accounts would be chosen to extract the tweets. Each tweet is to be classified into 2 category-spam and non-spam. These classified tweets then are used to train the various machine learning techniques. Words of each tweet considered as features and a feature vector was created using bag-of-words approach in order to create the instances. The data will be trained using SVM (Support Vector Machine).

General Terms

Classification, Spam Filtering methods

Keywords

Spam tweets, svm, kernel functions, SVM LIB

1. INTRODUCTION

There are different ways an algorithm can model a problem based on its interaction with the experience or environment or whatever we want to call the input data.

There are two general approaches to spam filtering: knowledge engineering (KE) and machine learning (ML). In the former case, a set of rules is created according to which tweets are categorized as spam or legitimate. A typical rule of this kind could look like "if the tweet contains the text BUY NOW, then the tweet is spam". A set of such rules should be created either by the user of the filter, or by some other authority (e.g.

The software company that provides a particular rule-based spam-filtering tool). [2]

The paper is organized as follows. The section2 presents a machine learning algorithm, different types of machine learning algorithm Section 3 explains spam filtering methods, Section 4 elaborates Classification of tweets, Section 5 explains support vector machine (SVM), kernel functions. Here will analyze one small dataset. Section 6 offers conclusions and directions for future work.

2. MACHINE LEARNING ALGORITHM

"Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed."

The data set is trained in such a way that the predefined output always come whenever applying machine learning algorithm.

2.1 Different types of Machine Learning Algorithm

Two most important learning styles are supervised and unsupervised learning [4]:

Supervised Learning: Labeled and classified data are used to train the model. The training process predicts the output; if it not matches the actual output then process is repeated again until the actual output comes. Sometimes it is also called intelligent learning model.

Working: This algorithm consists of a target variable which is to be predicted from a given set of predictors (independent variables). Using these set of variables, it generate a function that map inputs to desired outputs. The training process continues until the desired output is not obtained.

Examples of Supervised Learning: Regression, Decision Tree, Random Forest, KNN, Logistic Regression, back propagation neural network etc.

Unsupervised Learning: The training process is done on unlabelled /not classified data. The model is to be trained for the unknown output. It requires deep understanding of mathematical concepts.

Working: In this algorithm, do not have any target or outcome variable to predict / estimate. It is used for clustering population in different groups, which is widely used for segmenting customers in different groups for specific intervention. Example problems are clustering, dimensionality reduction and association rule learning. Example algorithms include: the Apriori algorithm and k-Means.

2.2 Applications of Machine Learning Algorithm[5]

Machine learning algorithms are being used in lots of places in interesting ways. Some applications are

2.2.1 Face detection

The face detection Feature in mobile cameras is an example of what machine learning can do. Cameras can automatically snap a photo when someone smiles more accurately now than ever before because of advances in machine learning algorithms.

2.2.2 Face recognition

This is where a computer program can identify an individual from a photo. You can find this feature on Facebook for automatically tagging people in photos where they appear. Advances in machine learning mean more accurate auto-face tagging soft wares.

2.2.3 Image classification:

A good example is the application of deep learning to improve image classification or image categorization in apps such as Google photos. Google photos would not be possible without advances in deep learning.

2.2.4 Speech recognition

Another good example is Google now. Improvements in speech recognition systems has been made possible by, you guessed right, machine learning specifically deep learning.

2.2.5 Google

Google defines itself as a machine learning company now. It is also a leader in this area because machine learning is a very important component to its core advertising and search businesses. It applies machine learning to improve search results and search suggestions.

2.2.6 Anti-virus

Machine learning is used in Anti-virus software's to improve detection of malicious software on computer devices.

2.2.7 Anti-spam:

machine learning is also used to train better anti-spam software systems.

2.2.8 Genetics:

Classical data mining or clustering algorithms in machine learning such as agglomerative clustering algorithms are used in genetics to help find genes associated with a particular disease.

2.2.9 Signal denoising:

Machine learning algorithms such as the K-SVD which is just a generalization of k-means clustering are used to find a dictionary of vectors that can be sparsely linearly combined to approximate any given input signal. Thus such a technique is used in video compression and denoising.

2.2.10 Weather forecast:

Machine learning is applied in weather forecasting software to improve the quality of the forecast.

Machine learning makes it relatively easier to develop sophisticated software systems without much effort on the human side. Instead of spending years handcrafting features or fine tuning a system with a lot of parameters, machine learning does that quicker. It also only requires training data to learn better features or parameters needed to improve a given system.

3. SPAM FILTERING METHODS

Different techniques have been used by researchers to find out the spam content. Table 1 is showing the summary of the different methods with their advantages and disadvantages. [3]

Table 1 Different spam filtering methods

<i>List-Based Filters</i>		<i>Advantages</i>	<i>Disadvantages</i>
	Blacklist	Attempts to stop unwanted email by blocking messages from a preset list of senders.	Misidentify legitimate senders as spammers.
	Real-Time Black hole List	(i) Identically to a traditional blacklist. (ii)Real-time black hole lists are maintained by third parties	(i)Can also generate false positives if spammers happen to use a legitimate IP address as a conduit for junk mail. (ii)Less control over what addresses are on — or not on — the list.
	White list	White list lets specify which senders to allow mail from; these addresses are placed on a trusted-users list.	Very strict filter that only uses a white list would mean that anyone who was not approved would automatically be blocked.
	Grey list	The receiving mail server initially rejects messages from unknown users and sends a failure message to the originating server. If the mail server	They may delay mail delivery, which could be inconvenient when you are expecting time-sensitive messages.

		attempts to send the message a second time — a step most legitimate servers will take — the grey list assumes the message is not spam and lets it proceed to the recipient's inbox. At this point, the grey list filter will add the recipient's email or IP address to a list of allowed senders.	
Content-Based Filters	Word-Based Filters	Configured to block messages containing more common words	IT staff will need to make time to routinely update the filter's list of blocked words.
	Heuristic Filters	Rather than blocking messages that contain a suspicious word, heuristic filters take multiple terms found in an email into consideration.	Configured to be aggressive may generate false positives if a legitimate contact happens to send an email containing a certain combination of words.
	Bayesian Filters	Employ the laws of mathematical probability to determine which messages are legitimate and which are spam.	However, since this method does require a training period before it starts working well, you will need to exercise patience and will probably have to

			manually delete a few junk messages, at least at first.
Other Filtering Methods	Challenge/Response System	Filters that use a challenge/response system block undesirable emails by forcing the sender to perform a task before their message can be delivered.	Challenge/response filters might block email newsletters you subscribe to, as these messages are typically sent by automated programs
	Collaborative Filters	Collaborative content filtering takes a community-based approach to fighting spam by collecting input from the millions of email users around the globe	If a group of spammers mobilize in large numbers and pretend to be legitimate users of the system, they could skew results by falsely labeling spam emails as legitimate messages.
	DNS Lookup Systems	Attempts to verify that the domain name in the email address of the sender — the part after the at symbol (@) — exists.	Reverse DNS lookups have been known to produce false positives — legitimate messages marked as spam

4. CLASSIFICATION of TWEETS

There are five main types of Tweets that is posted by people [6]:

4.1 Links

Posting links to good info or maybe posting a quote you like. The aim of these is to get them shared by others.

4.2 Images

To post a photo or GIF in a Tweet on Twitter

4.3 Quotes

The famous quotes and saying about life, love, friendship, philosophy and more.

4.4 Retweets

Sharing good tweets from others

4.5 Questions or comments of your own

4.6 Spam Tweets

Unfortunately Twitter is not immune to spam. In fact, spam happens on a regular basis with people often using the “@...” function to directly sell a product or service. This is poor practice.

5. MACHINE LEARNING ALGORITHM-SVM

The Support Vector Machine is one of the most modern techniques used in mail classification. In abstract view, it is a kernel machine with the strong mathematical base. It is a technique of pattern recognition and data analysis. The training sample is a set of vectors of n attributes. At the end of training phase, we can say that, we are in hyperspace having dimensions equal to the number attributes. In process of spam filtering, SVM Builds hyperspace with two classes, namely, spam and ham. These two classes are separated by a hyper plane. Every mail instance is treated as a single point with n dimensions in hyperspace. The distance between the hyper planes and points of each class, is kept maximum, for good separation. [1]

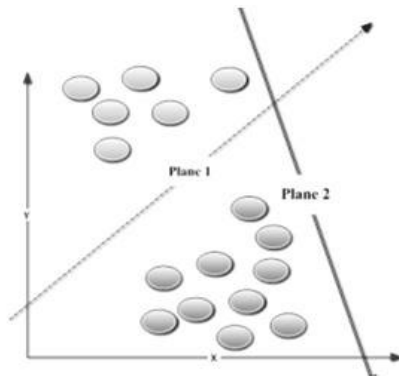


Fig.1: Hyper plane that separate the two classes

Plane1 is good classifier and Plane2 doesn't classify all instances. It may also happen that, the plane can't find good separator hyper plane (Plane 1) as in fig.1. In such case, hyperspace is called as non-linearly separable. To obtain linear separation in the non-linearly separable hyperspace, it is extended to more dimensions.

5.1 Kernel Functions

It gives the equivalent of mapping a nonlinear separable input space, to a higher dimensional feature space that is linearly separable. [1]

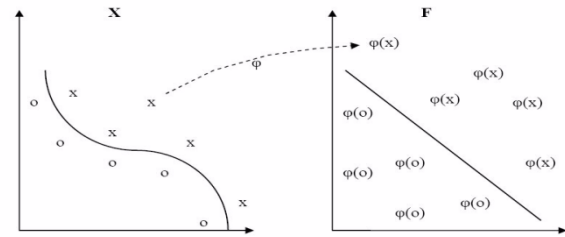


Fig 2: Mapping of kernel functions

The four types of kernel functions frequently used with SVM:

1. Linear
2. Polynomial
3. RBF
4. Sigmoid

There is no specific technique to detect which kernel is best for a particular problem.

5.2 LIBSVM: A Library for Support Vector Machines

Case study: Here analyzed svm using one small dataset. In R tool. The selection of kernel functions in the svm function parameters gives different output.

```
data <- seq(1,10)
classes <- c('b','b','b','b','a','a','a','a','b','b')
mysvm = svm (data, classes, type='C', kernel='radial',
gamma=0.1, cost=10)
pred = predict (mysvm, data)
print(table(pred, classes))
Output: classes
pred a b
a 4 0
b 0 6
mysvm = svm (data, classes, type='C',
kernel='polynomial/linear/sigmoid', gamma=0.1, cost=10)
Output: classes
pred a b
a 0 0
b 4 6
```

So the choice of kernel function plays a very important role in SVM machine learning algorithm.

6. CONCLUSION

In this paper we explained about machine learning algorithm, different spam filtering methods and implementation of SVM in the case study in the section 5.2 done on R. The project envisions two directions towards which the work can evolve. Future work would be planning to classify tweets based on spam vs. ham (legitimate tweets).The work would explore the SVM with different parameters and comparing SVM with

other classifier algorithm.. Finally, aim at investigating other kinds of machine learning algorithm for spam filtering

7. ACKNOWLEDGMENT

It gives immense pleasure to express my deep and sincere gratitude to my project guide Prof. Mrs.Smita Janjale for her kind help and valuable advice during the study and for her guidance and suggestions.

8. REFERENCES

- [1] “Dynamic Feature Selection for Spam Filtering Using Support Vector Machine” by Md. Rafiqul Islam, Wanlei Zhou and Morshed U. Choudhury
- [2] “Machine Learning Techniques in Spam Filtering” Konstantin Tretyakov,kt@ut.ee Institute of Computer Science, University of Tartu Data Mining Problem oriented Seminar, MTAT.03.177,
- [3] <http://www.techsoup.org/learningcenter/internet/page6028.cfm>
- [4] <http://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/> “A Tour of Machine Learning Algorithms by **Jason Brownlee** on November 25, 2013 in **Machine Learning Algorithms**
- [5] <https://www.quora.com/What-are-some-interesting-possible-applications-of-machine-learning>
- [6] “A Performance Evaluation of Machine Learning-Based Streaming Spam Tweets Detection” Chao Chen, Jun Zhang, *Member, IEEE*, Yi Xie, Yang Senior Member, IEEE, @2015
- [7] <https://blog.bufferapp.com/the-5-types-of-tweets-to-keep-your-buffer-full-and-your-followers-engaged>
- [8] “Spam Filtering Techniques and MapReduce with SVM: A Study” by Amol G. Kakade, Prashant K. Kharat, Anil Kumar Gupta, Tarun Batra ©2014 IEEE