OPTIMIZING SPAM FILTERING WITH MACHINE LEARNING

SOCIAL OR BUSSINESS IMPACT

The spam sms or emails are bad for various businesses in the first place. Spam is unsolicited or unwanted electronic messaging (usually in the form of email advertisements) that is blasted to lots of people. Receiving one out of a hundred isn't that bad, but imagine being deluged by countless useless emails every day. Beyond lost productivity, more dangerous spam comes in the form of phishing attempts and the dissemination of malware such as ransomware.

Three categories of spam filters

Since email is a primary communication medium for organizations, business email accounts send and receive emails in far higher volumes than personal ones. The higher capacity requirement led to the development of spam filters made especially for businesses. These filters are categorized as hardware spam filters, software spam filters, and cloud-based Hardware spam filters and the social or Business impact of spam emails are as follows :

Hardware spam filters

Hardware spam filters are dedicated appliances that are housed on-site and are placed in between the company's firewall and mail server to act as a gateway for all email traffic. Appliances vary by fixed capacity and must be chosen according to the number of active email users and domains a business has. Among the different spam filter types, hardware spam filters offer the least flexibility.

Software spam filters

These are programs installed on existing machines, virtually turning these into hardware spam filters in almost every way. However, if the machine is modifiable, then it can have far more flexibility when it comes to capacity.

Cloud-based spam filters

Instead of owning and managing your spam filtering solution yourself, you can have a cloud-based service provider take care of the filtering for you. It's essentially an outsourcing arrangement that lets you:

Real-time blacklists

These are continually updated lists of domains and IP addresses that are known to have been used to send spam. Emails from these points of origin are not delivered to their intended inboxes.

In the last two decades, spam detection and filtration gained the attention of a sizeable research community. The reason for a lot of research in this area is its costly and massive effect in many situations like consumer behaviour and fake reviews. The survey covers various machine learning techniques and models that the various researchers have proposed to detect and filter spam in emails and IoT platforms. The study categorized them as supervised, unsupervised, reinforcement learning, etc. The study compares these approaches and provides a summary of learned lessons from each category. This study concludes that most of the proposed email and IoT spam detection methods are based on supervised machine learning techniques. A labelled dataset for the supervised model training is a crucial and time-consuming task. Supervised learning algorithms SVM and Naive Bayes outperform other models in spam detection. The study provides comprehensive insights of these algorithms and some future research directions for email spam detection and filtering.

.