

OPTIMIZING SPAM FILTERING WITH MACHINE LEARNING

Specify The Bussiness Problem

Over recent years, as the popularity of mobile phone devices has increased, **Short Message Service (SMS)** has grown into a multi-billion dollar industry. At the same time, reduction in the cost of messaging services has resulted in growth in unsolicited commercial advertisements (**spams**) being sent to mobile phones. Due to Spam SMS, Mobile service providers suffer from some sort of financial problems as well as it reduces calling time for users. Unfortunately, if the user accesses such Spam SMS they may face the problem of virus or malware. When SMS arrives at mobile it will disturb mobile user privacy and concentration. It may lead to frustration for the user. So Spam SMS is one of the major issues in the wireless communication world and it grows day by day. To avoid such Spam SMS people use white and black list of numbers. But this technique is not adequate to completely avoid Spam SMS. To tackle this problem it is needful to use a smarter technique which correctly identifies Spam SMS. Natural language processing technique is useful for Spam SMS identification. It analyses text content and finds patterns which are used to identify Spam and Non-Spam SMS.

Many SMS Spam messages detection techniques are available these days to block spam messages and filtering spam messages. They have tried on them various messages portrayal techniques and machine learning calculations, as far as viability. The outcomes show that Bayesian separating methods can be successfully utilized to group SMS spam.

They have built two datasets one in English and another in Spanish. Their analysis shows that Bayesian filtering techniques that were earlier used in detecting email spam can also be used to block SMS Spam.

In the era of information technology, information sharing has become very easy and fast. Many platforms are available for users to share information anywhere across the world. Among all information sharing mediums, email is the simplest, cheapest, and the most rapid method of information sharing worldwide. But, due to their simplicity, emails are vulnerable to different **kinds of attacks**, and the most common and dangerous one is spam