**Fast email spam filtering methods**

ADITYA BHUTE1,NAGRAJ AAJURE2,SHUBHAM DHANORKAR3,AND

PROF.KAPIL WAGH4

1(Information Technology,Pimpri Chichwad Trust’s Nutan Maharashtra Institute of Engineering and Technology,India)

2(Information Technology,Pimpri Chichwad Trust’s Nutan Maharashtra Institute of Engineering and Technology,India)

3(Information Technology,Pimpri Chichwad Trust’s Nutan Maharashtra Institute of Engineering and Technology,India)

4(Information Technology,Pimpri Chinchwad Trust’s Nutan Maharashtra Institute of Engineering and Technology,India)

**ABSTRACT:** The paper elaborates on how text analysis influences classification—a key part of the spam-filtering process. The authors propose a multistage meta-algorithm for checking classifier performance. As a result, the algorithm allows for the fast selection of the best-performing classifiers as well as for the analysis of higher-dimensionality data. The last aspect is especially important when analyzing large datasets. The approach of cross-validation between different datasets for supervised learning is applied in the meta-algorithm. Three machine-learning methods allowing a user to classify e-mails as desirable (ham) or potentially harmful (spam) messages were compared in the paper to illustrate the operation of the meta-algorithm. The used methods are simple, but as the results showed, they are powerful enough. We use the following classifiers: k-nearest neighbours(k-NNs), support vector machines (SVM), and the naïve Bayes classifier (NB). The conducted research gave us the conclusion that multinomial naïve Bayes classifier can be an excellent weapon in the fight against the constantly increasing amount of spam messages. It was also confirmed that the proposed solution gives very accurate results.

**Keywords:** classifiers; e-mail; sms; ham; machine learning; spam