

## ABSTRAK

Penyebaran *e-mail* yang tidak diinginkan, yang disebut sebagai *spam*, pada lalu-lintas internet bisa berdampak pada rendahnya efisiensi dan menurunnya produktifitas pekerjaan pengguna *e-mail*. Beberapa *mail server* seperti *yahoo mail* dan *google mail* yang sudah memiliki mesin *spam filtering* pun tidak luput dari *spam* sehingga memungkinkan untuk *spam* masuk ke dalam *inbox* pengguna *e-mail*. *Naïve Bayes Classifier* adalah salah satu cara untuk mengatasi permasalahan di atas. *Naive Bayes Classifier* atau *Bayesian Filter* merupakan metode terbaru yang digunakan untuk mengklasifikasikan sekumpulan dokumen. Algoritma ini memanfaatkan metode probabilitas dan statistik yaitu memprediksi probabilitas dimasa depan berdasarkan pengalaman dimasa sebelumnya, metode statistik Bayesian inilah digunakan pada teknologi anti *spam filter*. Kelebihan *Bayesian Filtering* adalah tingkat akurasi yang tinggi dan *error rate* yang minimum.

Hasil dari implementasi program ini adalah mengklasifikasikan *e-mail* menjadi *spam e-mail* dan ham (bukan *spam e-mail*) pada *mail client* secara otomatis serta mengurangi kesalahan klasifikasi ham menjadi *spam e-mail*. Diharapkan sistem dapat mengklasifikasikan dengan tingkat keakuratan yang tinggi sehingga *spam e-mail* dapat teridentifikasi dengan baik.

Kata Kunci : *e-mail*, *e-mail filtering*, *spam*, *Naive Bayes Classifier*

## **ABSTRACT**

The spread of e-mails that undesirable, also known as “spam” on the internet traffic cause to reduced efficiency and productivity of email users. Some of mail servers like yahoo mail and google mail which already has a spam filtering engine was not free from allowing spam to the inbox of e-mail users. Naïve Bayes Classifier method is the one of way to solve of the problem already described. Naive bayes classifier or Bayesian filter is the latest method is used to classify a set of documents. These algorithms utilize the methods of probability and statistics that predict the probability of the future based on the experience of earlier decades, this bayesian statistical methods used in anti-spam technology filters. The advantages of bayesian filtering is a high degree of accuracy and minimum error rate.

The results of the implementation of this program is to classify spam e-mails into e-mails and ham (non-spam e-mails) on the mail server automatically and reduce misclassification of ham into a spam e-mails. It is expected that the system can classify with a high degree of accuracy so that the spam e-mails can be identified correctly.

Keywords : E-mail, E-mail filtering, Spam, Naive Bayes Classifier