

ABSTRAK

Pertumbuhan yang cepat dari internet, dalam hal ini komunikasi lewat *electronic mail (email)* menjadi salah satu bentuk komunikasi yang paling cepat dan ekonomis. Sebuah pesan *email* yang dikirim kepada sejumlah besar orang tanpa persetujuan dari orang tersebut, biasa disebut sebagai *unsolicited commercial email (UCE)*, *spam email*, *junk mail*, *bulk mail* atau *email* sampah.

Masalah *email* sampah (*spam* atau *junk email*) merupakan salah satu masalah yang dihadapi pada dunia internet. Untuk menyeleksi *email* yang datang secara manual akan membutuhkan waktu yang sangat banyak. Serta akan memakan kapasitas penyimpanan *email* yang akan memenuhi tempat penyimpanan *email-email* tersebut.

Pada tugas akhir ini telah dibuat aplikasi *spam filtering* dengan menggunakan metode pembeda markov yang akan melakukan filterisasi terhadap *email*. Pembuatan fitur dengan menggunakan *Sparse Binary Polynomial Hash (SBPH)* dengan skema pembobotan *Exponential Superincreasing Model (ESM)*. Metode pembeda markov mengklasifikasikan *email* menjadi *email spam* dan *legitimate* secara otomatis serta mengurangi kesalahan klasifikasi *email legitimate* menjadi *email spam*.

Kata kunci : *email*, *email spam*, *email legitimate*, fitur, filter *spam*, pembeda markov, SBPH, ESM.

ABSTRACT

The overwhelming growth of internet, in this case communication using electronic mail (email), become one of the fastest and economical type of communication. An email message that sent to a large amount of people without any permission from that person, that email can be called as an unsolicited commercial email (UCE), spam email, junk mail or bulk mail.

Spam Email problem represent one of the problem faced by the internet world. To filter the incoming email manually, it will need a large amount of time to read those email one by one. And it also will easily fill out the storage to save those email.