KLASIFIKASI *E-MAIL SPAM* DENGAN MENGGUNAKAN WEKA (*WAIKATO ENVIRONTMENT FOR KNOWLEDGE ANALYSIS*) *EXPLORER*

TUGAS AKHIR

MUHAMMAD FIRZA ARYOGI

1152001027

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK DAN ILMU KOMPUTER

UNIVERSITAS BAKRIE

JAKARTA

2020

KLASIFIKASI *E-MAIL SPAM* DENGAN MENGGUNAKAN WEKA (*WAIKATO ENVIRONTMENT FOR KNOWLEDGE ANALYSIS*) *EXPLORER*

TUGAS AKHIR

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer

MUHAMMAD FIRZA ARYOGI

1152001027

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK DAN ILMU KOMPUTER

UNIVERSITAS BAKRIE

JAKARTA

2020

HALAMAN PERNYATAAN ORISINALITAS

Tugas Akhir adalah hasil karya saya sendiri,
dan semua sumber baik dikutip maupun
dirujuk telahsaya nyatakan dengan benar.

Nama : Muhammad Firza Aryogi

NIM : 1152001027

Tanda Tangan :

Tanggal : 10 Maret 2020

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama	: Muhammad Firza Aryogi
NIM	: 1152001027
Program Studi	: Informatika
Fakultas	: Teknik dan Ilmu Komputer
Judul Skripsi	: Klasifikasi E-mail Spam dengan Menggunakan WEKA Explorer

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian dari persyaratan untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Bakrie.

DEWAN PENGUJI

Pembimbing	: Prof. Dr. Hoga Saragih, ST., MT.	( )
Penguji I	: Reyful Rey Fatri, M.Sc.	( )
Penguji II	: Sigit Wijayanto, M.Sc.	( )

Ditetapkan di : Jakarta

Tanggal : 10 Maret 2020

KATA PENGANTAR

Alhamdulillah, puji dan syukur penulis panjatkan kehadirat Allah SWT karena berkat segala rahmat karunia dan hidayah-Nya penulis diberikan kesehatan untuk menyelesaikan Tugas Akhir yang berjudul "Klasifikasi *E-mail Spam* dengan Menggunakan WEKA (*Waikato Environtment for Knowledge Analysis*) *Explorer*" ini. Shalawat serta salam tidak lupa penulis panjatkan kepada Nabi besar Muhammad SAW.

Tugas Akhir ini dibuat guna memenuhi syarat lulus untuk mendapatkan gelar Sarjana. Selama proses penyusunan Tugas Akhir ini penulis mendapatkan banyak hambatan. Namun hal ini dapat teratasi berkat motivasi, dukungan, serta saran dari berbagai pihak. Untuk itu penulis ingin mengucapkan terimakasih kepada:

- Allah SWT atas limpahan berkah, nikmat sehat, keselamatan dan kelancaran setiap waktu hingga saat ini.
- Kedua orangtua penulis dan keluarga yang senantiasa selalu memberikan do'a, motivasi, serta semangat.
- Bapak Prof. Dr. Hoga Saragih, ST. MT. Ketua Program Studi Teknik Informatika Universitas Bakrie.
- Seluruh tim dosen Program Studi Teknik Informatika Universitas Bakrie yang telah memberikan pembelaran, saran dan arahan.
- Dina Audina, Rafi Afdan, Angga Aulia, Fathur Hadyan Syah, Imam Nugroho, Nandhika Dwi, Bhakti Nurhasan, Ramadhani Qodriansyah
- Ahmad Novel Gadran, Alifian Azmi, Wahyu Widodo, Aziz Sentosa, Alhamsya, Salmaa Badriatu, Prima Dona serta teman-teman Informatika 2015 lainnya yang tidak bisa disebutkan satu persatu yang telah memberikan dukungan, semangat dan kebersamaan dalam suka maupun duka dalam empat tahun perkuliahan.
- Kakak-kakak *senior* dan adik-adik angkatan 2016 dan 2017 Informatika Universitas Bakrie.

- Seluruh pihak Universitas Bakrie baik terlibat langsung maupun tidak yang telah memberikan pengalaman, motivasi, dukungan, dan fasilitas yang sangat membantu selama masa perkuliahan dan penyusunan Tugas Akhir ini.

Penulis sangat bersyukur dan mengucapkan banyak terimakasih untuk pihak0pihak tersebut atas bantuannya dalam bentuk apapun sehingga penulis dapat menyelesaikan Tugas Akhir ini. Semoga Tugas Akhir ini dapat bermanfaat baik bagi semua kalangan pendidikan, khususnya bidang Informatika.

Jakarta, 14 Maret 2020

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai *civitas* akademik Universitas Bakrie, saya yang bertanda tangan di bawah ini :

Nama : Muhammad Firza Aryogi

NIM : 1152002017

Program Studi : Informatika

Fakultas : Teknik dan Ilmu Komputer

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bakrie **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:**

KLASIFIKASI *E-MAIL SPAM* DENGAN MENGGUNAKAN WEKA (*WAIKATO ENVIRONTMENT FOR KNOWLEDGE ANALYSIS*) *EXPLORER*

Dengan Hak Bebas Royalti Noneksklusif ini Universitas Bakrie berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencamtumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta untuk kepentingan akademis.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada Tanggal : 10 Maret 2020

Yang menyatakan

Muhammad Firza Aryogi

KLASIFIKASI *E-MAIL SPAM* DENGAN MENGGUNAKAN WEKA (*WAIKATO ENVIRONTMENT FOR KNOWLEDGE ANALYSIS*) *EXPLORER*

Muhammad Firza Aryogi

ABSTRAK

E-mail adalah salah satu alat komunikasi antar individu maupun perusahaan (*professional*) yang sangat dibutuhkan pada saat sekarang ini. Hal ini disebabkan *e-mail* merupakan alat komunikasi yang dapat mengirim pesan berbentuk digital dengan efisien, cepat dan murah dimana dapat memberikan keuntungan yaitu memangkas waktu dan biaya. Tujuan *e-mail* yaitu sebagai media pertukaran pesan atau informasi digital antar manusia baik individu maupun kelompok. Konten pada *e-mail* berbentuk informasi antar individu dan informasi perusahaan (bisnis). Namun, terdapat informasi lain seperti mengenai promo suatu produk barang maupun jasa, informasi berupa iklan, dan konten yang tidak diinginkan bagi penerima. Hal tersebut merupakan *spam* yang dapat mengganggu kenyaman para pengguna *e-mail*. Permasalahan tersebut dapat dicegah atau diminimalisir dengan mengklasifikasi data *e-mail* kedalam kategori *spam* dan non-*spam*. Metode klasifikasi yang dipakai untuk menyaring *e-mail* yaitu *naive bayes* dan J48 (*decision tree*) dengan melakukan *training* pada dataset dan melakukan *test* untuk menentukan prediksi kelas pada data *e-mail* tersebut.

Kata Kunci : *E-mail, Spam Detection, Spam Mail Filtration, Naive Bayes, Decision Tree, Klasifikasi*

E-MAIL SPAM CLASSIFICATION USING WEKA (WAIKATO ENVIRONTMENT FOR KNOWLEDGE ANALYSIS) EXPLORER

Muhammad Firza Aryogi

ABSTRACT

Today, e-mail is a crucial and necessary tool either for individuals (private) or for professionals (company). E-mail is an effective tool for communication as it saves a lot of time and cost. The purpose of e-mail as a media for exchanging digital message or information between human either indivual or group. Contents of e-mails is information between individual or information between company (business). But, there is another information such as information regarding promotion of a goods and services product, advertisements, and an unsolicited messages which is unwanted by the recipient. That is a spam that can interfere e-mail user's convenience. This problem could be prevented or could be minimized by classifying the data of e-mails into spam or non-spam category. Classifier method that can be used for spam e-mail filtration are naive bayes and J48 (decision tree) by performing a training on the dataset, and do the test after it to determine the prediction of the class from the data.

Keywords: E-mail, Spam Detection, Spam Mail Filtration, Naive Bayes, Decision Tree, Classification.

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR RUMUS

DAFTAR SINGKATAN

SMD	*Spam Mail Detection*
KDD	*Knowledge Discovery from Databases*
ID3	*Iterative Dichotomiser 3*
IDS	*Intrusion Detection System*
SVM	*Support Vector Machine*
CHAID	*Chi-squared Automatic Interaction Detection*
CART	*Classification and Regression Tree*
ADABOOST	*Adapive Boost*
TF-ISF	*Term Frequency – Inverse Sentence Frequency*
RDBMS	*Relational Database Management System*
CLI	*Command-Line Interface*
CSV	*Comma-Separated Values*