PENERAPAN NAÏVE BAYES PADA *DETECTION MALWARE* DENGAN DISKRITISASI VARIABEL

TESIS

INDA ANGGRAINI

ENTERPRISE IT INFRASTUCTURE

172420048

PROGRAM STUDI MAGISTER INFORMATIKA-S2

PROGRAM PASCASARJANA

UNIVERSITAS BINA DARMA

PALEMBANG

2020

PENERAPAN NAÏVE BAYES PADA DETECTION *MALWARE* DENGAN DISKRITISASI VARIABEL

Tesis ini diajukan sebagai salah satu syarat
Untuk memperoleh gelar

MAGISTER KOMPUTER

INDA ANGGRAINI

ENTERPRISE IT INFRASTUCTURE

172420048

PROGRAM STUDI MAGISTER INFORMATIKA-S2

PROGRAM PASCASARJANA

UNIVERSITAS BINA DARMA

PALEMBANG

2020

Halaman Pengesahan Penguji Tesis

Judul Tesis : PENERAPAN NAÏVE BAYES PADA *DETECTION MALWARE* DENGAN DISKRITISASI VARIABEL

Oleh INDA ANGGRAINI NIM 172420048 Tesis ini telah disetujui dan disahkan oleh Tim Penguji Program Studi Teknik Informatika – S2 Konsentrasi Enterprise IT Infrastructure program Pascasarjana Universitas Bina Darma Palembang pada 30 Januari 2020 dan telah dinyatakan LULUS.

Palembang, 30 Januari 2020
Mengetahui
Program Pascasarjana
Universitas Bina Darma
Direktur,

Dr.Ir.Hj. Hasmawaty AR, M.M.,M.T

Tim Penguji
Ketua I,

Dr. Firdaus, MT

Penguji II,

Yesi Novaria Kunang, S.T., M.Kom

Penguji III,

Dr. Widva Cholil. S.Kom..M.IT

Penguji IV,

Dr. Edi Surya Negara, M.Kom

iii

Halaman Pengesahan Pembimbing Tesis

Judul Tesis : PENERAPAN NAÏVE BAYES PADA DETECTION MALWARE DENGAN DISKRITISASI VARIABEL

Oleh INDA ANGGRAINI NIM 172420048 Tesis ini telah disetujui dan disahkan oleh Tim Penguji Program Studi Teknik Informatika – S2 Konsentrasi Enterprise IT Infrastructure program Pascasarjana Universitas Bina Darma Palembang pada tanggal 30 Januari 2020 dan telah dinyatakan LULUS.

Mengetahui,
Program Studi Teknik Informatika-S2
Universitas Bina Darma
Ketua,

Magister Teknik Informatika

Darius Antoni, S.Kom., M.M., Ph.D.

Tim Pembimbing
Pembimbing I,

Dr. Firdaus, M.T

Pembimbing II,

Yesi Novaria Kunang, S.T., M.Kom

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah :

Nama : INDA ANGGRAINI

NIM : 172420048

Dengan ini menyatakan bahwa :

1. Karya tulis saya tesis ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik Magister di Universitas Bina Darma;
2. Karya tulis ini murni gagasan, rumusan dan penelitian saya sendiri dengan arahan tim pembimbing;
3. Dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasi orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkan ke dalam daftar pustaka;
4. Karena yakin dengan keaslian karya tulis ini, saya menyatakan bersedia tesis yang saya hasilkan diunggah ke internet;
5. Surat pernyataan ini saya tulis dengan sungguh-sungguh dan apabila terdapat penyimpangan atau ketidaksamaan dalam pernyataan ini, maka saya bersedia menerima sanksi dengan aturan yang berlaku di perguruan tinggi ini.

Demikianlah surat pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya.

Palembang. 30 Januari 2020

Hormat pernyataan,

INDA ANGGRAINI

NIM : 172420048

ABSTRAK

Malicious software (Malware) adalah *software* jahat yang dirancang khusus untuk melakukan aktifitas berbahaya atau perusak perangkat lunak pada komputer seperti virus, Trojan, dan lain-lain yang disebar melalui jaringan internet. Banyaknya aktifitas penyebaran *Malware* yang terjadi melalui jaringan *internet* membuat banyak pengguna menjadi resah salah satu bentuk dari serangan tersebut yaitu dengan melakukan penyisipan file-file berbahaya atau malicious ke komputer. Contohnya seperti penyisispan *skrip web shell* yang di sisipkan ke komputer penyedia layanan *internet*. Penelitian ini bertujuan untuk melakukan analisa terhadap serangan *Malware* dengan menggunakan Algoritma Naïve Bayes Clasiffier dengan diskritisasi variabel. *Discretization* (pendiskritan) atribut merupakan teknik untuk merubah sebuah fungsi atau nilai kontinu kedalam bentuk diskrit. Teknik ini dilakukan sebagai penyesuaian terhadap kemungkinan kemunculan nilai kontinu dalam fitur *dataset* yang sangat kecil. Pendiskritisasian variabel dilakukan pada dataset yang bertipe kontinu, sehingga nilai probabilitas menunjukan kemungkinan nilai yang sama keluar pada suatu kelas. Dengan menggunakan algoritma naive bayes ini diharapkan dapat membantu mempermudah pengguna dalam menemukan metode yang tepat untuk mendeteksi serangan dari *Malware*.

Kata kunci : *Malicious software,* Algoritma Naïve Bayes, Diskritisasi Variabel.

ABSTRACT

Malicious software (Malware) is rogue software specifically designed to carry out malicious or destructive software activities on computers such as viruses, Trojans, and others that are spread through the internet network. The number of activities that spread Malware that occurs through the internet network makes many users uneasy one form of the attack is to insert malicious or malicious files into the computer. For example, such as web shell scripting script that is inserted into the internet service provider computer. This study aims to analyze Malware attacks using the Naïve Bayes Clasiffier Algorithm with the discretization of 3-interval and 5-interval Min-Max variables for continuous attributes. Discretization (discretion) attribute is a technique for changing a function or continuous value into a discrete form. This technique is done as an adjustment to the possibility of the emergence of continuous values in a very small dataset feature. Discretization of variables is done in a dataset of type continuous, so that the probability value indicates the possibility of the same value coming out of a class. Using the Naive Bayes algorithm is expected to help facilitate users in finding the right method for detecting attacks from Malware. The experimental results show that the application of Naïve Bayes in the classification of data that has not gone through the discretion stage produces an accuracy of 69.72% with prediction of Malware 63.53 % while the data that has passed the discretization stage is able to provide accuracy of up to 79.97 % with 81.29 % Malware prediction. The use of the Naïve Bayes by binning method in this study has an increased detection ability compared to the classification process without using the binning process (discretization). The discretion process can make the Naïve Bayes algorithm more accurate in detecting Malware

Keywords: Malicious software, Naïve Bayes algorithm, variable diskritization.

MOTO DAN PERSEMBAHAN

Motto

- *Sukses adalah ketika kesiapan dan kesempatan bertemu dalam satu waktu*
- *Kegagalan adalah guru terbaik dalam menggapai keberhasilan*
- *Teman sejati adalah dia yang berada paling depan ketika kita gagal bukan ketika kita meraih kesuksesan*

Persembahan

Seiring doa dan puji syukur kepada Allah SWT. Kupersembahkan tesis ini Kepada :

- *Bapak, Ibu, Kakak dan Keluargaku tercinta*
- *Bapak Firdaus dan Ibu Yesi Novaria Kunang terima kasih atas bimbingan dan bantuannya selama menyelesaikan Tugas Akhir*
- *Terima kasih untuk rekan kerja, sahabat, dan teman seperjuangan Angkatan (MTI.17.A) terkhusus untuk Dina Dalillah dan Yoga Elang Johar yang memberikan dukungan serta masukkan*
- *Almamaterku yang kubanggakan Universitas Bina Darma Palembang*

KATA PENGANTAR

Alhamdulillah, atas segala nikmat yang diberikan oleh Allah SWT yang selalu memberikan berkah, rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan tesis yang berjudul Penerapan Naïve Bayes Pada *Detection Malware* Dengan Diskritisasi Variabel.

Tesis ini disusun sebagai salah satu syarat untuk untuk memperoleh gelar Magister Komputer pada Universitas Bina Darma Palembang. Dalam penulisan tesis ini penulis telah berusaha semaksimal mungkin memberikan dan menyajikan yang terbaik. Tetapi penulis juga menyadari bahwa tesis ini masih jauh dari sempurna, hal ini dikarenakan terbatasnya pengetahuan yang dimiliki penulis. Oleh karena itu, penulis mengharapkan saran dan kritik yang bersifat membangun untuk kesempurnaan tesis ini.

Pada kesempatan ini, tidak lupa penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan bimbingan, pengarahan, nasihat, dan pemikiran dalam menyelesaikan tesis ini, terutama kepada:

1. Prof. Dr. Sunda Ariana, M.Pd., M.M. selaku Rektor Universitas Bina Darma Palembang.
2. Dr. Ir. H. Hasmawaty AR, M.M., M.T. Selaku Direktur Pascasarjana Universitas Bina Darma Palembang.
3. Darius Antoni, S.Kom., M.M., Ph.D. Selaku Ketua Program Studi Magister Informatika Universitas Bina Darma Palembang

4. Dr. Firdaus, M.T selaku Pembimbing I yang telah memberikan bimbingan tesis ini.
5. Yesi Novaria Kunang, S.T., M.Kom selaku pembimbing II yang telah memberikan bimbingan dan arahan dalam penulisan tesis ini.
6. Pihak Sekretariat Pascasarjana Universitas Bina Darma Palembang yang telah memberikan bimbingan pelayanan dengan baik.
7. Berbagai pihak yang telah memberikan bantuan dan dorongan serta berbagai pengalaman pada proses penyusunan proposal tesis ini.

Palembang, Januari 2020

Penulis,

Inda Anggraini

NIM : 172420048

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN