



**KLASIFIKASI MALWARE BERDASARKAN FITUR API
CALL DAN ANDROID PERMISSIONS MENGGUNAKAN
RADIAL BASIS FUNCTION NETWORK**

SKRIPSI

Oleh

Bagas Aditya Wibisono

NIM. 1810511075

**PROGRAM STUDI INFORMATIKA, PROGRAM SARJANA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

2022



**KLASIFIKASI MALWARE BERDASARKAN FITUR API
CALL DAN ANDROID PERMISSIONS MENGGUNAKAN
RADIAL BASIS FUNCTION NETWORK**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana
Komputer**

Oleh

Bagas Aditya Wibisono

NIM. 1810511075

**PROGRAM STUDI INFORMATIKA, PROGRAM SARJANA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

2022

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Bagas Aditya Wibisono
NIM : 1810511075
Tanggal : 24 Juni 2022

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 24 Juni 2022

Yang Menyatakan,

A handwritten signature in black ink is written over a 1000 Rupiah stamp. The stamp features the Garuda Pancasila emblem and the text 'SEPULUH RIBU RUPIAH', '1000', and 'METRAI TEMPEL' with the serial number '9896EAJX804627822'.

(Bagas Aditya Wibisono)

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Bagas Aditya Wibisono
NIM : 1810511075
Fakultas : Ilmu Komputer
Program Studi : Informatika

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

KLASIFIKASI MALWARE BERDASARKAN FITUR API CALL DAN ANDROID PERMISSIONS MENGGUNAKAN RADIAL BASIS FUNTION NETWORK

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilih Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 24 Juni 2022

Yang Menyatakan,



(Bagas Aditya Wibisono)

LEMBAR PERSETUJUAN

Dengan ini dinyatakan bahwa Skripsi berikut :

Nama : Bagas Aditya Wibisono

NIM : 1810511075

Program Studi : S1 Informatika

Judul : Klasifikasi Malware Berdasarkan Fitur API Call dan Android Permissions Menggunakan Radial Basis Function Network.

Sebagai bagian dari persyaratan yang diperlukan untuk mengikuti Sidang Skripsi pada Program Studi S1 Informatika Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Menyetujui,
Ketua Program Studi



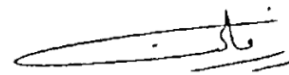
(Desti Sandya Prasvita, S.Kom., M.Kom.)

Menyetujui
Dosen Pembimbing I



(Dr. Didit Widiyanto S.Kom., M.Si.)

Menyetujui
Dosen Pembimbing II



(Noor Falih, S.Kom., M.T.)

Ditetapkan : Jakarta

Tanggal Persetujuan : 24 Juni 2022

LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa skripsi berikut :

Nama : Bagas Aditya Wibisono
NIM : 1810511075
Program Studi : S1 Informatika
Judul : Klasifikasi Malware Berdasarkan Fitur API Call dan Android Permissions Menggunakan Radial Basis Function Network

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Henki Bayu Seta, S.Kom., MTI.

Penguji I



Nurul Chamidah, S.Kom., M.Kom

Penguji II



Dr. Didit Widiyanto, S.Kom., M.Si.

Pembimbing I



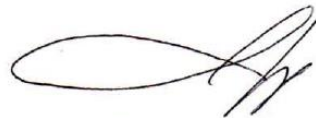
Noor Falih, S.Kom., M.T.

Pembimbing II



Dr. Ermatifa, M.Kom.

Dekan



Desta Sandya Prasvita, S.Kom., M.Kom.

Ketua Program Studi

Ditetapkan : Jakarta

Tanggal Pengesahan : 21 Juli 2022



KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadiran Allah SWT atas nikmat, rahmat, dan karunia-Nya sehingga skripsi sebagai syarat kelulusan dengan judul “Klasifikasi Malware Berdasarkan Fitur API Call dan Android Permissions Menggunakan Metode Radial Basis Function Network Dengan K-Means Clustering Sebagai Metode Pemilihan Centroidnya” dapat diselesaikan.

Ucapan terima kasih penulis sampaikan atas dukungan, bimbingan, bantuan, serta motivasi yang telah diberikan selama proses penyusunan skripsi, kepada:

1. Ibu Dr. Ermatita, M.Kom. selaku Dekan Fakultas Ilmu Komputer
2. Bapak Dr. Didit Widiyanto, S.Kom., M.Si. selaku Dosen Pembimbing I
3. Bapak Noor Falih, S.Kom., M.T. selaku Dosen Pembimbing II
4. Bapak Desta Sandya Prasvita, M.Kom. selaku Ketua Program Studi S1 Informatika sekaligus Pembimbing Akademik
5. Serta orangtua, keluarga, kerabat, dan teman-teman sekalian.

Semoga skripsi ini dapat memberikan berbagai manfaat dan ilmu yang berguna di masa mendatang.

KLASIFIKASI MALWARE BERDASARKAN FITUR API CALL DAN ANDROID *PERMISSIONS* MENGGUNAKAN *RADIAL BASIS FUNCTION NETWORK*

Bagas Aditya Wibisono

ABSTRAK

Malware telah menjadi ancaman utama bagi pengguna teknologi saat ini. *Malware* atau *Malicious Software* adalah sebuah perangkat lunak yang bersifat intrusif dengan tujuan menginfeksi, menjelajah, mencuri, atau merusak perangkat. Berbagai metode pendeteksian *malware* telah dikembangkan untuk mengantisipasi perkembangan *malware*. Penelitian ini berfokus pada klasifikasi *malware* berdasarkan fitur *API call* dan *Android permissions* menggunakan *Radial Basis Function Network* dengan *K-Means Clustering* sebagai metode pemilihan centroid. *Radial Basis Function Network* merupakan bagian dari jaringan syaraf tiruan yang menggunakan fungsi *Gaussian* sebagai fungsi aktivasinya, sedangkan *K-Means Clustering* merupakan algoritma *unsupervised learning* dalam *machine learning* atau algoritma pengelompokan. Dataset yang digunakan adalah *malgenome-215-dataset* yang dapat diunduh pada repositori *figshare*. Data *split* dilakukan dengan menggunakan *K-Fold*. Pengujian yang dilakukan berdasarkan hyperparameter *learning rate*, jumlah *epoch*, jumlah *hidden unit*, dan jumlah *K* pada *K-Fold* nya. Akurasi, *precision*, *recall*, dan *F1 score* dihitung berdasarkan *confusion matrix*. Hasil eksperimen menunjukkan akurasi 98,41%, *precision* 99,3%, *recall* 97,92%, dan *F1 score* 98,6%.

Kata Kunci : *Klasifikasi Malware, API call, Android Permissions, Radial Basis Function Network, K-Means Clustering*

MALWARE CLASSIFICATION BASED ON API CALL FEATURE AND ANDROID PERMISSIONS USING RADIAL BASIS FUNCTION NETWORK

Bagas Aditya Wibisono

ABSTRACT

Malware has become a major threat to technology users today. Malware or Malicious Software is intrusive software with the purpose of infecting, browsing, stealing, or damaging a device. Various malware detection methods have been developed to anticipate the development of malware. This study focuses on malware classification based on API calls and Android permissions using Radial Basis Function Network with K-Means Clustering as the centroid selection method. Radial Basis Function Network is part of an artificial neural network that uses a Gaussian function as its activation function, while K-Means Clustering is an unsupervised learning algorithm in machine learning or clustering algorithms. The dataset used is malgenome-215-dataset which can be downloaded from the figshare repository. Data split is done with K-Fold. The tests were carried out based on the hyperparameters of the learning rate, the number of epochs, the number of hidden units, and the number of K in the K-Fold. Accuracy, precision, recall, and F1 scores were calculated based on the confusion matrix. The experimental results showed 98.41% accuracy, 99.3% precision, 97.92% recall, and 98.6% F1 score.

Keywords : *Malware Classification, API call, Android Permissions, Radial Basis Function Network, K-Means Clustering*

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN ORISINALITAS	ii
PERNYATAAN PERSETUJUAN	iii
LEMBAR PERSETUJUAN	iv
LEMBAR PENGESAHAN	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Ruang Lingkup Penelitian	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Luaran Yang Diharapkan	4
1.7 Sistematika Penulisan	4
BAB II	6
TINJAUAN PUSTAKA	6
2.1 Landasan Teori	6
2.1.1 API Call	6
2.1.2 Android Permissions	6
2.1.3 Malware	7
2.1.4 Machine Learning	8
2.1.5 Radial Basis Function Network	15
2.1.6 K-Means Clustering	17
2.1.7 Cross-Validation	18
2.2 Studi Literatur	20
BAB III	29

METODE PENELITIAN	29
3.1 Kerangka Pikir	29
3.1.1 Identifikasi Permasalahan	30
3.1.2 Studi Literatur.....	30
3.1.3 Pengumpulan Data.....	30
3.1.4 Praproses data.....	31
3.1.5 Klasifikasi Malware.....	31
3.1.6 Evaluasi Hasil.....	35
3.1.7 Kesimpulan	36
3.2 Perangkat Pendukung.....	37
BAB IV	38
HASIL DAN PEMBAHASAN	38
4.1 Pengumpulan Data.....	38
4.2 Praproses Data	40
4.2.1 Label Encoding.....	40
4.2.2 Data Balancing	41
4.2.3 Feature Selection	42
4.3 Pemilihan Centroid dengan K-Means.....	49
4.4 Pengujian Model Radial Basis Function Network	53
4.4.1 Pengujian Berdasarkan K-Fold	54
4.4.2 Pengujian Berdasarkan <i>Learning Rate</i>	55
4.4.3 Pengujian Berdasarkan <i>Epoch</i>	56
4.4.4 Pengujian Berdasarkan <i>Hidden Unit</i>	57
4.5 Evaluasi Hasil.....	58
BAB V.....	60
PENUTUP	60
5.1 Kesimpulan.....	60
5.2 Saran	61
DAFTAR PUSTAKA	62

DAFTAR TABEL

Tabel 2.2 Ringkasan studi literatur	23
Tabel 3.1 Rancangan percobaan berdasarkan <i>K-Fold</i>	32
Tabel 3.2 Rancangan percobaan berdasarkan <i>learning rate</i>	32
Tabel 3.3 Rancangan percobaan berdasarkan <i>epoch</i>	32
Tabel 3.4 Rancangan percobaan berdasarkan jumlah <i>hidden unit</i>	33
Tabel 3.5 <i>Confusion matrix</i>	36
Tabel 4.1 Keterangan jumlah fitur dan <i>class</i>	38
Tabel 4.2 Urutan kemunculan fitur pada dataset	39
Tabel 4.3 Keterangan fitur dataset	39
Tabel 4.4 Dataset dengan target <i>class</i> sebelum <i>encoding</i>	40
Tabel 4.5 Dataset dengan target <i>class</i> sesudah <i>encoding</i>	41
Tabel 4.6 Sampel data <i>mutual information</i>	43
Tabel 4.7 Keterangan fitur <i>mutual information</i>	43
Tabel 4.8 Probabilitas <i>class</i>	44
Tabel 4.9 Probabilitas fitur terhadap <i>class</i>	44
Tabel 4.10 Probabilitas fitur dan <i>class</i>	45
Tabel 4.11 Daftar 100 fitur relevan	46
Tabel 4.12 Sampel data dengan 10 fitur	50
Tabel 4.13 Keterangan fitur <i>K-Means Clustering</i>	50
Tabel 4.14 <i>Cluster centroid</i> awal	51
Tabel 4.15 Data masukan	51
Tabel 4.16 <i>Centroid</i> dengan nilai baru	53
Tabel 4.17 Hyperparameter <i>K-Fold</i>	54
Tabel 4.18 Hasil pengujian berdasarkan <i>k-fold</i>	54
Tabel 4.19 Hyperparameter <i>learning rate</i>	55
Tabel 4.20 Hasil pengujian berdasarkan <i>learning rate</i>	55
Tabel 4.21 Hyperparameter <i>epoch</i>	56
Tabel 4.22 Hasil pengujian berdasarkan <i>epoch</i> dengan <i>learning rate</i> 0.001	56
Tabel 4.23 Hasil pengujian berdasarkan <i>epoch</i> dengan <i>learning rate</i> 0.01	57
Tabel 4.24 Hyperparameter <i>hidden unit</i>	57

Tabel 4.25 Hasil pengujian berdasarkan <i>hidden unit</i>	58
Tabel 4.26 <i>Confusion matrix</i>	59

DAFTAR GAMBAR

Gambar 2.1 Supervised learning	9
Gambar 2.2 Unsupervised learning	10
Gambar 2.3 Reinforcement learning	11
Gambar 2.4 Neural network	12
Gambar 2.5 Supervised neural network	13
Gambar 2.6 Unsupervised neural network	13
Gambar 2.7 Struktur <i>Radial Basis Function Network</i>	15
Gambar 2.8 Algoritma <i>Radial Basis Function Network</i>	17
Gambar 2.9 K-Means Clustering	17
Gambar 2.10 Ilustrasi <i>K-Fold</i>	19
Gambar 3.2 Kerangka pikir	29
Gambar 3.3 Arsitektur Radial Basis Function Network	33
Gambar 3.4 Model <i>Radial Basis Function</i>	35
Gambar 4.1 Proses <i>One Hot Encoding</i>	40
Gambar 4.2 Grafik distribusi data sebelum <i>balancing</i>	41
Gambar 4.3 Proses <i>Random Undersampling</i>	42
Gambar 4.4 Grafik perbandingan <i>class</i> setelah <i>balancing</i>	42

DAFTAR LAMPIRAN

Lampiran A. Daftar fitur <i>malgenome-215-dataset</i>	A-2
Lampiran B. <i>Source code</i> klasifikasi <i>malware</i>	B-1
Lampiran C. <i>Similarity index</i> skripsi	C-Error! Bookmark not defined.