IMPLEMENTASI ALGORITMA PCA UNTUK PENINGKATAN SISTEM *CLASSIFICATION MALWARE ANDROID*

TUGAS AKHIR

OLEH :
MUHAMMAD NAWAWI
09011381722104

JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2021

LEMBAR PENGESAHAN

IMPLEMENTASI ALGORITMA PCA UNTUK PENINGKATAN SISTEM KLASIFIKASI *MALWARE ANDROID*

PROPOSAL TUGAS AKHIR

Program Studi Sistem Komputer
Jenjang S1

Oleh

Muhammad Nawawi
09011381722104

Palembang,

Mengetahui,

Ketua Jurusan Sistem Koputer	Pembimbing Tugas Akhir
Dr. Ir. H. Sukemi, M.T	**Deris Stiawan, Ph. D.**
NIP. 196612032006041001	**NIP. 197806172006041002**

HALAMAN PERSETUJUAN

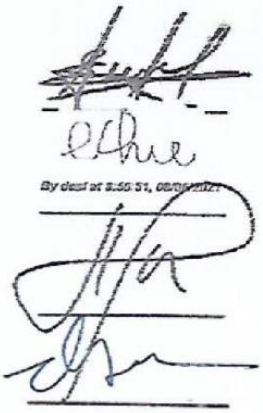
Telah diuji dan lulus pada:

Hari : Kamis

Tanggal : 01 April 2021

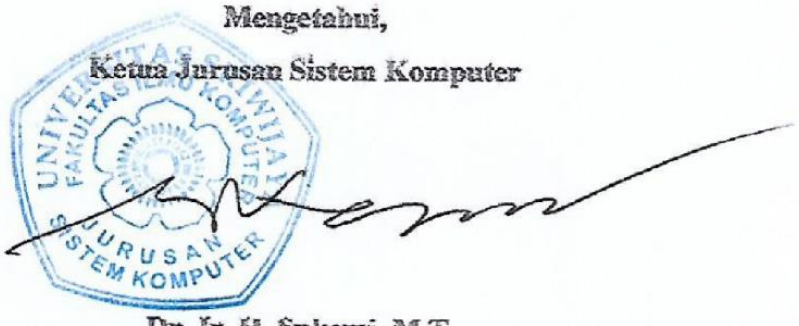
Tim Penguji:

1. Ketua : Sarmayanta Sembiring, M.T.
2. Sektretaris : Sri Desy Siswanti, M.T.
3. Anggota : Huda Ubaya, M.T.
4. Pembimbing : Deris Stiawan, Ph.D.



Mengetahui,

Ketua Jurusan Sistem Komputer



Dr. Ir. H. Sukemi, M.T.

NIP. 196612032006041001

SURAT KETERANGAN PENGECEKKAN

SIMILIARITY

Saya yang bertanda tangan dibawah ini :

Nama	: Muhammad Nawawi
NIM	: 09011381722104
Prodi	: Sistem Komputer
Fakultas	: Ilmu Komputer

Menyatakan bahwa benar hasil pengecekkan similarity Skripsi/Tesis/Disertasi/ Lap. Penelitian yang berjudul IMPLEMENTASI ALGORITMA PCA UNTUK PENINGKATAN SISTEM *CLASSIFICATION MALWARE ANDROID* adalah 19%.

Dicek oleh operator *:

1. Dosen Pembimbing
2. UPT Perpustakaan
3. Operator Fakultas .....

Demikian surat keterangan ini saya buat dengan sebenarnya dan dapat saya pertanggung jawabkan.

Indralaya, Mei 2021

Menyetujui

Dosen pembimbing

Deris Stiawan, Ph.D.
Nip. 197806172006041002

Yang menyatakan

Muhammad Nawawi
Nim. 09011381722104

*Lingkari salah satu jawaban tempat anda melakukan pengecekkan Similarity

MOTTO :

Saat semua orang menganggapmu tidak bisa apa-apa.
Jangan perdulikan,
Karena
Yang bisa mengubah nasibmu,
Adalah kamu,
Bukan orang lain.
Dan
Ketika dunia jahat kepadamu,
Maka berusahalah untuk manghadapinya,
Karena tidak aka nada orang yang membantumu
Jika kau tidak berusaha.

Ku persembahkan untuk:

- **Papa, Mama, Adik-Adik, Kakak, Nenek, Kakek, Makwo, Pakwo, Bibik, Oom, Encik, Adinda yang slalu mendukung dan memberikan semangat kepadaku.**
- **Teman-teman seperjuangan Sistem Komputer Universitas Sriwijaya angkatan 2017 yang tidak akan kulupakan**
- **Almamaterku Universitas Sriwijaya**

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadirat Allah SWT, atas segala karunia dan rahmat-Nya sehingga penulis dapat menyelesaikan proposal tugas akhir dengan judul **"Implementasi Algoritma PCA Untuk Peningkatan Sistem *Classification Malware Android* "**.

Penulisan Proposal Tugas Akhir ini dilakukan untuk melengkapi salah satu syarat memperoleh gelar Sarjana Komputer di Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya. Adapun sebagai bahan penulisan, penulis mengambil berdasarkan hasil penelitian, observasi dan beberapa sumber literatur yang mendukung dalam penulisan proposal ini. Pada kesempatan ini juga, penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu baik dari segi moril ataupun materil serta memberikan kemudahan, dorongan, saran dan kritik selama dalam proses penulisan Proposal Tugas Akhir ini. Oleh karena itu, pada kesempatan ini penulis mengucapkan rasa syukur kepada Allah SWT. dan mengucapkan terima kasih kepada yang terhormat :

1. Ke empat Orang Tua terutama untuk mama yang slalu mensuport dan mendoakan serta keluarga penulis tercinta, yang telah memberikan doa dan restu serta dukungan yang sangat besar selama mengikuti dan melaksanakan perkuliahan di Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya hingga dapat menyelesaikan Proposal Tugas Akhir ini.
2. Bapak Jaidan Jauhari, S.Pd., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Bapak DR. IR. H. Sukemi, M. T selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Deris Stiawan, Ph. D. selaku Dosen Pembimbing Tugas Akhir penulis
5. Bapak Rossi Passarella, M.ENG selaku Dosen Akademik di Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
6. dosen, staff, serta karyawan Fakultas Ilmu Komputer Universitas Sriwijaya.

7. Seluruh teman-teman seperjuangan angkatan 2017 Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya terutama Abdi, Vanisa, Ika, Adinda, Barzan, Tata, Marle, Hafidz, Hadi, Naufal, Yuan, Badruz, Tri Agung, Taufik, Ryan, Tiara, Nanda, Vira dan Fidya.
8. Almamater.

Penulis menyadari bahwa masih banyak kekurangan dalam penulisan Proposal Tugas Akhir ini. Karena sesungguhnya tak ada yang sempurna didunia ini. Untuk itu, segala saran dan kritik sangatlah penting bagi penulis. Akhir kata, semoga Proposal Tugas Akhir ini dapat bermanfaat dan berguna bagi khalayak.

Palembang, April 2021

Penulis

Muhammad Nawawi

NIM. 09011381722104

IMPLEMENTATION OF PCA ALGORITHM TO IMPROVE ANDROID MALWARE CLASSIFICATION SYSTEM

Muhammad Nawawi (09011381722104)

Department of Computer Engineering, Faculty of Computer Science, Sriwijaya University

Email : muhammadnawawi1516@gmail.com

Abstract

Dimensional reduction has been widely used in various studies in the world because of its good function in reducing data without removing the characteristics of the data. In this study, we want to see how the increase in results will be obtained for classifying if you use dimensional reduction in the data you want to use. In this research, the dimensional reduction of the PCA algorithm is used to classify andoroid malware, namely adware. The classification in this study was carried out using three classification using ANN (Artificial Neural Network). Differentiate each component from the result of dimensionality reduction using PCA and using optimal functions such as Adam, Adadelta and SGD (Stochastic Gradient Descent). Get pretty good results with a value of 99.90% accuracy, 99.99% precision, 99.84% sensitivity, 99.98% specificity and 99.91% for the F1-Score.

Keywords : *Malware Andorid, Adware,* Reduksi dimensi, algoritma PCA, *Classification, Artificial Neural Network.*

IMPLEMENTASI ALGORITMA PCA UNTUK PENINGKATAN SISTEM *CLASSIFICATION MALWARE ANDROID*

Muhammad Nawawi (09011381722104)

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email : muhammadnawawi1516@gmail.com

Abstrak

Reduksi dimensi sudah banyak dipakai berbagai penelitian didunia karena fungsinya yang baik dalam memperkecil data tanpa menghilangkan karakteristik dari data trsebut. Pada penelitian ini ingin melihat bagaimana peningkatan hasil yang akan didapatkan untuk melakukan klasifikasi jika menggunakan reduksi dimensi pada data yang ingin digunakan. Pada penelitian kali ini menggunakan reduksi dimensi algoritma PCA untuk melakukan klasifikasi pada malware andoroid yaitu adware. Klasifikasi pada penelitian ini dilakukan dengan menggunakan tiga buah *classification* dengan mengggunakan ANN (*Artificial Neural Network*). Membedakan tiap komponen dari hasil reduksi dimensi menggunakan PCA dan menggunakan fungsi optimal seperti Adam, Adadelta dan SGD (*Stochastic Gradient Descent*). Mendapatkan hasil yang cukup baik yaitu nilai 99.90% akurasi, 99.99% presisi, 99.84% sensitivitas, 99.98% spesifitas dan 99.91% untuk F1-Score.

Kata Kunci : *Malware Andorid*, *Adware*, Reduksi dimensi, algoritma PCA, Klasifikasi, *Artificial Neural Network*.

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1. Latar Belakang

Malware adalah sebuah *software* yang bisa menyusup ke sebuah sistem operasi yang dapat merusak sistem serta mencuri *file-file* pribadi penting yang ada pada system yang disusupi. Beberapa oknum menggunakan *malware* untuk mendapatkan data-data penting dari korban dengan banyak tujuan. *Smartphone* dengan sistem operasi *Android*pun juga menjadi target serangan oknum yang memanfaatkan *malware*. Salah satunya *Adware* yang dapat menyerang *smartphone* dengan melalui iklan yang ada pada aplikasi yang dijalankan ataupun pada *website* yang kita telusuri.

Iklan digunakan untuk mempromosikan, menjual produk atau layanan. Dengan munculnya internet strategi pemasaran bergerak cepat menuju periklanan digital, dan tren ini tidak menunjukkan tanda-tanda mereda. Iklan digital muncul di situs web *online* ataupun aplikasi seluler. Saat ini, kebanyakan *smartphone* menggunakan *platform android*, yang sudah mendorong pertumbuhan ekosistem aplikasi *Android* dengan sangat pesat. Akibatnya, *malware* dapat digunakan untuk melakukan aktivitas berbahaya saat menggunakan hak istimewa yang diberikan oleh sistem operasi *Android*[1]. Penyedia iklan membayar pengembangan aplikasi dan situs web dengan jumlah tertentu untuk setiap tayangan iklan yang ditampilkan kepada pengguna. Perangkat lunak yang melakukan penipuan berupa iklan disebut dengan *adware*[2].

Adware adalah suatu iklan yang difungsikan untuk menguntungkan semua pihak, tetapi sekarang *adware* lebih diartikan sebagai penyedia dan sumber utama dari *malware*. *Adware* umumnya dikenal sebagai "perangkat lunak yang secara otomatis menampilkan atau mengunduh materi iklan ketika pengguna sedang online ataupun menjalankan aplikasi[3]. Jenis serangan *adware* berbahaya yang paling sering ialah menyusup serta mengontrol perangkat *Android* pengguna menggunakan *adware* berbahaya[4]. *Adware* juga dapat menyertakan aktivitas berbahaya lainnya seperti, mencuri informasi *sensitive* dari pengguna dan

menyebarkannya kepada pihak ketiga. *Adware* akan terus menerus menampilkan iklan bahkan ketika pengguna tidak menjalankan program yang dikehendaki sebelumnya[5].

Pengembangan *malware* telah menjadi aktivitas yang serius akhir-akhir ini karena jumlah *platform* target meningkat dari tahun ke tahun, yang secara signifikan meningkatkan pengembangan teknik yang memadai untuk mendeteksi malware[6]. *Adware* juga menjadi dasar untuk me-*rooting smartphone* pengguna agar mendapatkan akses ke perangkat[2]. Ada dua cara penipuan iklan yaitu pada seluler dan web. Pada seluler iklan akan tampil melaluin aplikasi *Android* yang telah berkerjasama dengan iklan tersebut ataupun aplikasi yang telah disisipkan *malware,* sedangkan pada web akan muncul ditampilan atau halaman web dan ada juga saat kita akan mengklik sesuatu tetapi yang muncul adalah iklan.

Ini adalah bukti bahwa *malware* merupakan ancaman yang serius dalam sistem dan sering kali melampaui kapasitas analis *malware*. Oleh karena itu, diperlukan upaya yang cukup besar untuk melindungi dari serangan *malware*[6]. Penelitian mengenai kalsifikasi *malware* sudah banyak dilakukan dalam beberapa tahun ini. Sehingan penelitian mengenai kalsifikasi *malware* bukan hal yang baru lagi, namun pengembangan penelitian ini masih terus berjalan hingga sekarang. Ha1 ini dapat dilihat dari seberapa banyaknya penelitian yang telah dilakukan sebelumnya. Pada penelitian[7] mereka menggunakan algoritma *Random Forest* (RF), *K-Nearest Neighbor* (KNN), *Decision Tre* (DT), *Random Tree* (RT), *Regression* (R) dengan tingkat akurasi sebesar 94%. Dari penelitian tersebut masih memiliki kekurangan dengan proses memakan waktu yang lama dan banyak data yang terlewatkan. Sehingga perlu dilakukan penelitian yang lebih baik dan tidak membutuhkan waktu lama untuk klasifikasi *malware Android* agar dapat mengatasi masalah diatas.

1.2. Rumusan Masalah

Rumusan masalah yang akan dibahas pada tugas akhir ini adalah :

1. Bagaimana mengklasifikasi *malware Android* dengan menggunakan algoritma ANN?

2. Bagaimana menerapkan sistem klasifikasi *Binary* dengan menggunakan algoritma ANN?
3. Bagaimana algoritma PCA mendukung sistem klasifikasi *malware Android* menggunakan ANN?

1.3. Batasan Masalah

Batasan masalah untuk tugas akhir ini sebagai berikut :

1. Dataset yang pakai pada penelitian ini berasal dari *University of New Brunswick* yaitu CIC&Mal2017[8]
2. Algoritma yang digunakan pada klasifikasi *malware Android* adalah algoritma ANN
3. Proses *dimentionality reduction* yang digunakan adalah *Principal Component Analysis*
4. Dalam penelitian ini tidak membahas deteksi dan pencegahan

1.4. Tujuan

Adapun tujuan yang hendak dicapai dari dilakukannya penelitian ini adalah :

1. Menerapkan algoritma ANN untuk klasifikasi data *malware Android*
2. Menerapkan sistem klasifikasi *Binary* dan melakukan peningkatan nilai akurasi dari penelitian sebelumnya
3. Mengetahui pengaruh algoritma *Principal Component Analysis* pada kinerja sistem klasifikasi *malware Android*

1.5. Manfaat

Adapun beberapa manfaat yang didapat dari di1akukannya penelitian ini adalah :

1. Dapat mengklasifikasi malware Android dengan digunakan algoritma ANN
2. Algoritma *Principal Component Analysis* sangat berpengaruh untuk melakukan peningkatan kinerja system klasifikasi dengan digunakan algoritma ANN
3. Hasil yang telah didapatkan dari penelitian ini bisa menjadi referensi untuk peningkatan nilai akurasi, spesifitas, sensitivitas, F1-score dan presisi untuk

melakukan klasifikasi dengan menerapkan algoritma *Principal Component Analysis* dan ANN

1.6. Metodologi Penelitian

Penelitian ini akan melewati beberapa tahapan :

1. Tahap Pertama (Studi Pustaka / Literatur)

   Tahap ini dilakukan dengan cara mencari dan membaca literatur dan referensi tentang "Klasifikasi *malware* menggunakan ANN" sehingga dapat menunjang penulisan laporan Tugas Akhir.
2. Tahap Kedua (Prosesing pengubahan data PCAP ke csv menggunakan CICFlowMeter)

   Tahap ini dilakukan untuk mempersiapkan data yang akan digunakan.
3. Tahap Ketiga (Pembersihan File)

   Pada tahap ini dilakukan proses dengan melakukan reduksi dimensi
4. Tahap Keempat (Eksperimen)

   Pada tahap ini menyederhanakan data agar lebih maksimum, kemudian melakukan klasifikasi data normal dan malware.
5. Tahap Kelima (Penarikan Kesimpulan dan Saran)

   Pada tahap ini bisa ditarik kesimpulan hasil klasifikasi dan setudi literatur agar untuk penulis selanjutnya yang akan menjadikan bahan referensi.

1.7. Sistematika Penulisan

Agar memudahkan untuk menyusun tugas akhir ini serta memperjelaskan isi dari tiap bab yang ada pada laporan ini, dibuatlah penulisan sistematik dari penelitian ini sebagai berikut.

BAB I PENDAHULUAN

Bab ini berisikan uraian latar belakang dengan sistematik topik yang diambil

BAB II TINJAUAN PUSTAKA

Bab ini berisikan tentang literatur yang relevan, bagan teori dan bagan berfikir

BAB III METODOLOGI

Bab ini menjelaskan step by step dan terperinci tentang apa yang akan dilakukan dan digunakan dengan mencari kemudian menganalisa tema yang sudah dibuat penulis untuk tugas akhir.

BAB IV PENGUJIAN DAN ANALISA

Bab ini menjelaskan dari hasil pengujian yang sudah dilakukan dan di analisa dengan hasil klasifikasi yang sudah dibuat.

BAB V KESIMPULAN

Bab ini berisikan tentang kesimpulan yang telah didapatkan oleh penulis serta merupakan jawaban dari semua tujuan yang akan dicapai pada bab I (Pendahuluan).

DAFTAR PUSTAKA

[1] A. Pekta\cs and T. Acarman, "Ensemble machine learning approach for android malware classification using hybrid features," in *International Conference on Computer Recognition Systems*, 2017, pp. 191–200.

[2] S. Suresh, F. Di Troia, K. Potika, and M. Stamp, "An analysis of Android adware," *J. Comput. Virol. Hacking Tech.*, vol. 15, no. 3, pp. 147–160, 2019.

[3] J. Gao, L. Li, P. Kong, T. F. Bissyandé, and J. Klein, "Should you consider adware as malware in your study?," in *2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, 2019, pp. 604–608.

[4] K. Lee and H. Park, "Malicious Adware Detection on Android Platform using Dynamic Random Forest," in *International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, 2019, pp. 609–617.

[5] J. Y. Ndagi and J. K. Alhassan, "Machine Learning Classification Algorithms for Adware in Android Devices: A Comparative Evaluation and Analysis," in *2019 15th International Conference on Electronics, Computer and Computation (ICECCO)*, 2019, pp. 1–6.

[6] F. O. Catak, A. F. Yaz\i, O. Elezaj, and J. Ahmed, "Deep learning based Sequential model for malware analysis using Windows exe API Calls," *PeerJ Comput. Sci.*, vol. 6, p. e285, 2020.

[7] M. Murtaz, H. Azwar, S. B. Ali, and S. Rehman, "A framework for Android Malware detection and classification," in *2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, 2018, pp. 1–5.

[8] A. H. Lashkari, A. F. A. Kadir, L. Taheri, and A. A. Ghorbani, "Toward developing a systematic approach to generate benchmark android malware datasets and classification," in *2018 International Carnahan Conference on Security Technology (ICCST)*, 2018, pp. 1–7.

[9] K. H. Walse, R. V Dharaskar, and V. M. Thakare, "PCA Based Optimal ANN Classi fi ers for Human Activity Recognition Using Mobile Sensors Data," *in Proceedings of First International Conference on Information and Communication Technology*, vol. 1, pp. 429-436, 2016.

[10] M. K. Alzaylaee, S. Y. Yerima, and S. Sezer, "DL-Droid: Deep learning based android malware detection using real devices," *Comput. Secur.*, vol. 89, p. 101663, 2020.

[11] O. T. Suryati and A. Budiono, "Impact Analysis of Malware Based on Call Network API with Heuristic Detection Method," *Int. J. Adv. Data Inf. Syst.*, vol. 1, no. 1, pp. 1–8, 2020.

[12] M. B. B. de Robles, J. A. C. Hermocilla, and J. P. Pabico, "Characterization

and Classification of Malware Traffic over the Tor Network." *in Proceedings of the 20th Philippine Computing Science Congress* (PCSC 2020), pp. 78-87, 2020.

[13] J. Eliyanto, "Reduksi Dimensi untuk Meningkatkan Performa Metode Fuzzy Klastering pada Big Data," *in Public Knowledge Project,* vol. 1, no. 1, pp. 27–36, 2019.

[14] H. Lu, S. Liu, H. Wei, and J. Tu, "Expert Systems with Applications Multi-kernel fuzzy clustering based on auto-encoder for fMRI functional network," *Expert Systems with Applications*, vol. 159, 2020.

[15] P. Hartono, "Mixing Autoencoder With Classifier : Conceptual Data Visualization," *IEEE Access*, vol. 8, no. 1, pp. 105301-105310, 2020.

[16] D. Arivudainambi, V. K. KA, P. Visu, and others, "Malware traffic classification using principal component analysis and artificial neural network for extreme surveillance," *Comput. Commun.*, vol. 147, pp. 50–57, 2019.

[17] Y.-M. Kwon, J.-J. An, M.-J. Lim, S. Cho, and W.-M. Gal, "Malware Classification Using Simhash Encoding and PCA (MCSP)," *Symmetry (Basel).*, vol. 12, no. 5, p. 830, 2020.

[18] R. S. Suhartanto, C. Dewi, and L. Muflikhah, "Implementasi Jaringan Syaraf Tiruan Backpropagation untuk Mendiagnosis Penyakit Kulit pada Anak," Jurnal Masyarakat Informatika (JMASIF), vol. 1, no. 7, pp. 555–562, 2017.

[19] I. M. Nasser, M. Al-Shawwa, and S. S. Abu-Naser, "Artificial Neural Network for Diagnose Autism Spectrum Disorder," *in International Journal of Academic Information System Research* (IJAISR), vol. 3, 2019.

[20] S. A. Aditya, "Inovasi Metode ANN-S Untuk Pengecekan Baterai Berbasis Labview," in *Seminar Nasional Teknik Elektro*, 2020, vol. 5, no. 2, pp. 235–238.

[21] J. Teknovasi, E. B. Nababan, and M. Zarlis, "BIPOLAR DALAM ALGORITMA BACKPROPAGATION PADA PREDIKSI KEMAMPUAN SISWA," Jurnal Ilmiah Informatika (JIF), vol. 02, pp. 103–116, 2015.

[22] Y. Yu *et al.*, "RMAF : Relu-Memristor-Like Activation Function for Deep Learning," *IEEE Access*, vol. 8, pp. 72727–72741, 2020.

[23] T. K. R. Arvind, M. Brand, C. Heidorn, S. Boppu, F. Hannig, and J. Teich, "Hardware Implementation of Hyperbolic Tangent Activation Function for Floating Point Formats," *IEEE Access*, p. 121, 2020.

[24] R. Adam, D. O. Melinte, and L. Vladareanu, "Facial Expressions Recognition for Human – Robot Interaction Using Deep Convolutional Neural," *Sensors*, vol. 20, 2020.

[25] V. N. Sewdien, R. Preece, J. L. R. Torres, E. Rakhshani, and M. Van Der Meijden, "Assessment of critical parameters for arti fi cial neural networks based short-term wind generation forecasting," *Renew. Energy*, vol. 161, pp.

878–892, 2020.

[26] K. Rahman, “Training Sensitivity in Graph Isomorphism Network,” ACM *Digital Library* pp. 2181–2184, 2020.

[27] G. Perin and S. Picek, “On the Influence of Optimizers in Deep Learning-based Side-channel Analysis,” *Cryptology ePrint Archive*, pp. 13–28, 2020.

[28] I. Kandel and M. Castelli, “Comparative Study of First Order Optimizers for Image Classification Using Convolutional Neural Networks on Histopathology Images,” *J. Imagin*, vol. 6, 2020.

[29] Q. Wang, Y. Ma, K. Zhao, and Y. Tian, “A Comprehensive Survey of Loss Functions in Machine Learning,” *Ann. Data Sci*, 2020.

[30] H. Qin, R. Gong, X. Liu, X. Bai, J. Song, and N. Sebe, “Binary neural networks : A survey,” *Pattern Recognit*, vol. 105, 2020.'

[31] A. Luque, A. Carrasco, A. Mart\’\in, and A. de las Heras, “The impact of class imbalance in classification performance metrics based on the binary confusion matrix,” *Pattern Recognit.*, vol. 91, pp. 216–231, 2019.

[32] A. H. Lashkari, A. F. A. Kadir, H. Gonzalez, K. F. Mbah, and A. A. Ghorbani, “Towards a network-based framework for android malware detection and characterization,” in *2017 15th Annual Conference on Privacy, Security and Trust (PST)*, 2017, pp. 233–23309.

[33] M. Noorani, S. Mancoridis, and S. Weber, “*On the Detection of Malware on Virtual Assistants Based on Behavioral Anomalies*,” *Drexel University*, 2019.