SISTEM KLASIFIKASI SERANGAN SQL *INJECTION* & XSS PADA RAMA *REPOSITORY* DENGAN METODE *RANDOM FOREST* (RF)

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat

Memperoleh Gelar Sarjana Komputer

Oleh:

AMARTYA BIMANTARA

09011281722041

JURUSAN SISTEM KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2021

SISTEM KLASIFIKASI SERANGAN SQL *INJECTION* & XSS PADA RAMA *REPOSITORY* DENGAN METODE *RANDOM FOREST* (RF)

TUGAS AKHIR
Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

OLEH:
AMARTYA BIMANTARA
09011281722041

JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA

2021

LEMBAR PENGESAHAN

SISTEM KLASIFIKASI SERANGAN SQL *INJECTION* & XSS PADA RAMA *REPOSITORY* DENGAN METODE *RANDOM FOREST* (RF)

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

AMARTYA BIMANTARA
09011281722041

Indralaya, Juli 2021

Pembimbing I	Pembimbing II
Deris Stiawan, M.T., Ph.D., IPU	**Ali Bardadi, S.SI., M.Kom.**
NIP. 197806172006041002	**NIP. 198806292019031007**

Ketua Jurusan Sistem Komputer

Dr. Ir. H. Sukemi, M.T.
NIP. 196612032006041001

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Kamis

Tanggal : 15 Juli 2021

Tim Penguji :

1. **Ketua Sidang** : **Sarmayanta Sembiring, M.T**
2. **Sekretaris Sidang** : **Iman Saladin B. Azhar, M.MSI**
3. **Penguji Sidang** : **Ahmad Heryanto, M.T**
4. **Pembimbing I** : **Deris Stiawan, M.T., Ph.D., IPU**
5. **Pembimbing II** : **Ali Bardadi, S.SI., M.Kom.**

Mengetahui,

Ketua Jurusan Sistem Komputer

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

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Amartya Bimantara

NIM : 09011281722041

Judul : Sistem Klasifikasi Serangan SQL *Injection* & XSS Pada RAMA *Repository* Dengan Metode *Random Forest* (RF)

Hasil Pengecekan *Software iThenticate/Turnitin* : 2%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan tugas akhir ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya.

Demikian, pernyataan ini saya buat dalam keadaan sadar dan tanpa paksaan dari siapapun.

Indralaya, Juli 2021

Amartya Bimantara
NIM. 09011281722041

HALAMAN PERSEMBAHAN

"Usahamu tidak akan pernah mengkhianatimu"

"Tidak masalah seberapa lambat engkau berjalan selama engkau tidak berhenti"

"Pemenang tidak percaya pada takdir, mereka tidak bisa menerima kekalahan" -David Silva

"Bila kau cemas dan gelisah akan sesuatu, masuklah kedalamnya sebab ketakutan menghadapinya lebih mengganggu daripada sesuatu yang kau takuti sendiri" -Ali bin Abi Thalib

"Karya ini kupersembahkan untuk dua orang yang paling berharga dalam hidupku, yang selalu mengajariku arti dari kehidupan, selalu menuntunku melalui ucapan, dan selalu menyertaiku melalui doa, IBU dan AYAHKU"

KATA PENGANTAR

Assalamu'alaikum Warahmatullah Wabarakatuh

Puji syukur Alhamdulillah penulis panjatkan atas kehadirat Allah SWT yang telah memberikan karunia dan rahmat-Nya, sehingga penulis dapat menyelesaikan penulisan Tugas Akhir ini yang berjudul **"Sistem Klasifikasi Serangan SQL *Injection* XSS pada RAMA *Repository* dengan Metode *Random Forest* (RF)".**

Pada kesempatan ini, penulis ingin mengucapkan terima kasih kepada beberapa pihak atas ide dan saran serta bantuannya dalam menyelesaikan penulisan Tugas Akhir ini. Oleh karena itu, penulis ingin mengucapkan rasa syukur dan terima kasih kepada Allah SWT yang telah memberikan berkah serta nikmat Kesehatan dan kesempatan kepada penulis dalam menyusun Tugas Akhir ini:

- Allah SWT, yang telah memberikan rahmat dan karunia-Nya sehingga saya dapat menyelesaikan penulisan Proposal Tugas Akhir ini dengan baik dan lancar.
- Orang Tua saya tercinta yang telah membesarkan saya dengan penuh kasih saying dan selalu mengajarkan saya dalam berbuat hal yang baik. Terimakasih untuk segala do'a, motivasi dan dukungannya baik moril, materil, maupun spiritual selama ini.
- Bapak Jaidan Jauhari, S.Pd., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
- Bapak Dr. Ir. H. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
- Bapak Deris Stiawan, M.T., Ph.D., selaku Dosen Pembimbing I Tugas Akhir di Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.

- Bapak Ali Bardadi, S.SI., M.Kom., selaku Dosen Pembimbing II Tugas Akhir di Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
- Bapak Ahmad Fali Oklilas, M.T., selaku Pembimbing Akademik Jurusan Sistem Komputer.
- Mbak Nurul Afifah, S.Kom., M.Kom., yang telah membantu membimbing dalam menyelesaikan Tugas Akhir.
- Mbak Renny Virgasari selaku admin Jurusan Sistem Komputer yang telah membantu mengurus seluruh berkas.
- Tia Hermita, Meutia Zamieyus, Febi Rusmiati, Lisa Melinda, Ahmad Afidin, Agung Setiawan, dan teman – teman seperjuangan grup Riset Comnets lainnya yang telah banyak membantu.
- Civitas Akademika Fakultas Ilmu Komputer Universitas Sriwijaya.
- Dan Seluruh pihak yang telah membantu.

Penulis menyadari bahwa laporan ini masih sangat jauh dari kata sempurna. Untuk itu, kritik dan saran yang membangun sangat diharapkan penulis. Akhir kata penulis berharap, semoga Tugas Akhir ini bermanfaat dan berguna bagi khalayak.

Wassalamualaikum Warahmatullah Wabarakatuh

Indralaya, Juli 2021

Penulis,

CLASSIFICATION SYSTEM OF SQL INJECTION & XSS ON RAMA REPOSITORY USING RANDOM FOREST (RF)

Amartya Bimantara (09011281722041)
Dept. of Computer Engineering, Faculty of Computer Science, Sriwijaya University
Email : amartyabimantara3@gmail.com

ABSTRACT

SQL Injection is an attack on a website database system that exploits security gaps located in the database layer of a website application. XSS is an attack on a website application that takes resources from victims of website visitors such as cookies and credit card numbers by entering malicious scripts on the website. In this study using RAMA Repository datasets derived from RAMA Repository website. Principal Component Analysis (PCA) algorithm is used to gain important values on features for the classification process. Additionally, a Random Forest (RF) algorithm is used to perform SQL Injection and XSS attack classifications. The result of this study showed that Random Forest (RF) by utilizing Principal Component Analysis (PCA) algorithm can classify SQL Injection and XSS attacks quite well with accuracy results of 98.95%, precision of 98.95%, sensitivity of 98.95%, F1-Score of 98.95%.

***Keywords*:**

Classification, SQL Injection, Cross Site Scripting, XSS, RAMA Repository, Principal Component Analysis (PCA), Random Forest (RF)

SISTEM KLASIFIKASI SERANGAN SQL *INJECTION* & XSS PADA RAMA *REPOSITORY* DENGAN METODE *RANDOM FOREST* (RF)

Amartya Bimantara (09011281722041)
Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya
Email : amartyabimantara3@gmail.com

ABSTRAK

SQL *Injection* merupakan serangan pada sistem database suatu website yang memanfaatkan celah keamanan yang berada pada layer basis data sebuah aplikasi website. XSS merupakan serangan pada aplikasi website yang mengambil sumber daya dari korban pengunjung website seperti *cookies* dan nomor kartu kredit dengan memasukkan *script* berbahaya pada website. Pada penelitian ini menggunakan dataset RAMA *Repository* yang berasal dari website repositori RAMA. Digunakan algoritma *Principal Component Analysis* (PCA) untuk mendapatkan nilai – nilai penting pada fitur untuk proses klasifikasi. Selain itu, digunakan algoritma *Random Forest* (RF) untuk melakukan klasifikasi serangan SQL *Injection* dan XSS. Hasil pada penelitian ini menunjukkan bahwa *Random Forest* (RF) dengan memanfaatkan algoritma *Principal Component Analysis* (PCA) dapat melakukan klasifikasi serangan SQL *Injection* dan XSS dengan cukup baik dengan hasil akurasi sebesar 98.95%, presisi sebesar 98.95%, sensitivitas sebesar 98.95%, F1-Score sebesar 98.95%.

Kata Kunci :

Classification, SQL Injection, Cross Site Scripting,* XSS, *RAMA Repository, Principal Component Analysis (PCA), Random Forest (RF)

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB I
PENDAHULUAN

1.1 Latar Belakang

SQL *injection* merupakan serangan kode injeksi yang dilakukan dengan menginjeksi kueri SQL berbahaya ke dalam sistem suatu aplikasi web dan SQL *injection* merupakan salah satu ancaman yang sangat serius [1]. Sintaks SQL memungkinkan sintaks perintah database untuk dicampur dengan data pengguna. Jika pengembang tidak berhati-hati, data pengguna dapat diartikan sebagai perintah, sehingga penyerang tidak hanya dapat memasukkan data ke aplikasi Web, tetapi juga menjalankan perintah secara *illegal* pada database [2].

Cross Site Scripting (XSS) termasuk jenis serangan bersifat injeksi dengan menyuntikan skrip berbahaya kedalam web. XSS bekerja pada saat skrip pada sisi browser disuntikan dengan menggunakan aplikasi web kepada *end user* yang berbeda. Serangan ini secara tidak langsung menargetkan korban, namun harus mengeksploitasi kerentanan website yang dikunjungi korban. XSS masuk dalam TOP 10 *Vulnerability Web Attack* oleh *Open Web Application Security Project* (OWASP) [3].

RAMA *Repository* merupakan repositori nasional untuk laporan dari hasil penelitian seperti proyek mahasiswa (diploma), tugas akhir, skripsi, tesis (S2), disertasi (S3), termasuk laporan penelitian dosen atau peneliti yang bukan merupakan publikasi di jurnal, konferensi, maupun buku yang diintegrasikan dari repositori Lembaga Penelitian dan Perguruan Tinggi di Indonesia [4].

Pada penelitian [5], membahas bagaimana mengklasifikasikan serangan *Cross Site Scipting* (XSS) menggunakan 10 algoritma *Machine Learning*. Penelitian ini menggunakan dataset yang diperoleh dari sumber internet terpercaya yaitu XSSed, Alexa, dan Elgg berjumlah 1000 dimana terbagi menjadi 400 data normal dan 600 data serangan. Dari proses klasifikasi yang dilakukan, algoritma *Random Forest* mendapatkan hasil terbaik yaitu nilai akurasi 97.2%, nilai presisi 97.7%, nilai sensitivitas 97.1%, dan nilai F1-score 97.4%.

Selanjutnya, pada penelitian [6], membahas mengenai deteksi serangan XSS dengan menggunakan algoritma *Machine Learning*. Penelitian ini menggunakan, dataset yang berasal dari XSSed, yaitu arsip online terbesar dari situs web yang rentan akan serangan XSS. Kemudian dilakukan proses deteksi dengan menggunakan tiga algoritma *Machine Learning*, yaitu algoritma *Random Forest*, *Support Vector Machine* (SVM), dan *K-Nearest Neighbor*. Hasil akurasi yang didapatkan algoritma *Random Forest* cukup tinggi yaitu 99,50%.

Pada penelitian [7], digunakan SMOTE untuk mengatasi data yang *imbalanced* dan memiliki banyak *value* yang kosong. Hasil dari penelitian ini yaitu SMOTE berpengaruh dalam hasil klasifikasi yang dilakukan dengan menghasilkan akurasi yang cukup baik karena data yang digunakan telah menjadi *balance*.

Pada penelitian [8], algoritma *Principal Component Analysis* (PCA) dan *Linear Discriminant Analysis* (LDA) digunakan untuk mereduksi fitur ukuran data tanpa menghilangkan informasi – informasi penting data yang selanjutnya akan dilakukan klasifikasi menggunakan algoritma *Decision Tree*.

Berdasarkan ulasan di atas, penulis akan membahas mengenai sistem klasifikasi serangan SQL *injection* dan XSS pada RAMA *Repository* menggunakan dataset berformat sql yang dikonversi terlebih dahulu sehingga dapat di proses untuk dilakukan seleksi fitur dengan algoritma *Principal Component Analysis* (PCA). Selanjutnya hasil dari seleksi fitur tersebut akan diklasifikasi menggunakan algoritma *Random Forest* untuk mendapatkan model terbaik.

1.2 Perumusan Masalah

Perumusan masalah dalam penelitian Tugas Akhir ini adalah sebagai berikut:

1. Bagaimana menerapkan seleksi fitur untuk proses klasifikasi serangan SQL *injection* dan XSS?
2. Bagaimana melakukan klasifikasi serangan SQL *injection* dan XSS?
3. Bagaimana pengaruh hasil performa klasifikasi dengan metode *random forest* terhadap nilai akurasi, presisi, sensitivitas, F1-Score, BACC, dan MCC?

1.3 Batasan Masalah

Batasan masalah pada penelitian Tugas Akhir ini yaitu sebagai berikut:

1. Dataset yang digunakan pada peneltian ini berasal dari repositori RAMA.
2. Algoritma yang digunakan untuk melakukan klasifikasi yaitu algoritma *Random Forest*.
3. Proses seleksi fitur dilakukan dengan menggunakan algoritma PCA.
4. Jenis serangan yang akan diklasifikasikan yaitu SQL *injection* dan XSS.
5. Penelitian ini tidak membahas mengenai bagaimana cara melakukan pencegahan seranan SQL *injection* dan XSS pada repositori RAMA.

1.4 Tujuan

Tujuan pada penelitian Tugas Akhir ini yaitu sebagai berikut:

1. Menerapkan algoritma *Principal Component Analysis* (PCA) untuk melakukan seleksi fitur pada proses klasifikasi serangan SQL *injection* dan XSS.
2. Menerapkan metode *Random Forest* untuk klasifikasi serangan SQL *injection* dan XSS dan mendapatkan model terbaik.
3. Mengukur hasil performa akurasi, presisi, sensitivitas, F1-Score, BACC, dan MCC.

1.5 Manfaat

Manfaat dari penelitian Tugas Akhir ini yaitu sebagai berikut:

1. Mengoptimalkan hasil dari klasifikasi
2. Dapat menerapkan metode *Random Forest* dalam mengklasifikasikan serangan SQL *injection* dan XSS.
3. Mendapatkan performa yang baik pada proses klasifikasi dengan metode *Random Forest*.

1.6 Metodologi Penelitian

Metodologi penelitian pada penelitian ini dilakukan melalui tahapan – tahapan yang meliputi:

1. Metode Literatur dan Studi Pustaka

   Dalam tahap ini, peneliti mencari informasi – informasi mengenai sistem klasifikasi serangan dengan algoritma *Random Forest* melalui beberapa jurnal ilmiah, internet, buku, serta artikel terkait yang mendukung dalam penulisan tugas akhir.

2. Metode Konsultasi

   Dalam tahap ini, peneliti melakukan konsultasi kepada pihak – pihak yang memiliki pengetahuan, informasi, serta wawasan yang baik dalam mengatasi permasalahan yang ditemui pada penulisan tugas akhir.

3. Metode Pengumpulan Data

   Dalam tahap ini, dilakukan pengumpulan dan pengambilan data serangan SQL *injection* dan XSS pada repositori RAMA.

4. Metode Pengujian

   Dalam tahap ini, dilakukan pembangunan atau perancangan sistem yang untuk melatih serta mendapatkan hasil klasifikasi serangan SQL *injection* dan XSS.

5. Metode Analisa dan Kesimpulan

   Dalam tahap ini, dilakukan analisa terhadap hasil dari proses klasifikasi dan membuat beberapa kesimpulan dari penelitian ini.

1.7 Sistematika Penulisan

Sistematika penulisan pada penelitian tugas akhir ini yaitu sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini, dijelaskan latar belakang, perumusan masalah, batasan masalah, tujuan, dan manfaat dari topik yang diangkat yaitu sistem klasifikasi serangan SQL *injection* dan XSS pada RAMA *Repository* dengan metode *Random Forest*.

BAB II. TINJAUAN PUSTAKA

Pada bab ini, dijelaskan mengenai beberapa *literature review* yang berhubungan dengan permasalahan klasifikasi serangan SQL *injection* dan XSS dengan algoritma *Random Forest*.

BAB III. METODOLOGI PENELITIAN

Pada bab ini, dijelaskan secara bertahap mengenai bagaimana tahap – tahap yang dilakukan pada penelitian, meliputi tahapan mempersiapkan dataset, penerapan algoritma PCA dan *Random Forest*, dan model klasifikasi yang akan digunakan sehingga tujuan dari penelitian dapat tercapai.

BAB IV. HASIL DAN ANALISA

Pada bab ini, dijelaskan mengenai hasil yang telah diperoleh dari tahap sebelumnya. Adapun hasil yang telah didapatkan dianalisa dan hasil yang didapatkan divalidasi.

BAB V. KESIMPULAN

Pada bab ini, dijelaskan mengenai kesimpulan dari hasil yang telah didapatkan dan merupakan jawaban yang didapatkan dari tujuan.

DAFTAR PUSTAKA

[1] L. Zhang, D. Zhang, C. Wang, J. Zhao, and Z. Zhang, "ART4SQLi: The ART of SQL Injection Vulnerability Discovery," *IEEE Trans. Reliab.*, vol. 68, no. 4, pp. 1470–1489, 2019, doi: 10.1109/TR.2019.2910285.

[2] L. Ma, D. Zhao, Y. Gao, and C. Zhao, "Research on SQL Injection Attack and Prevention Technology Based on Web," *Proc. - 2nd Int. Conf. Comput. Network, Electron. Autom. ICCNEA 2019*, pp. 176–179, 2019, doi: 10.1109/ICCNEA.2019.00042.

[3] S. Gupta and B. B. Gupta, "Cross-Site Scripting (XSS) attacks and defense mechanisms: classification and state-of-the-art," *Int. J. Syst. Assur. Eng. Manag.*, vol. 8, pp. 512–530, 2017, doi: 10.1007/s13198-015-0376-0.

[4] P. Tinggi and L. Penelitian, "RAMA REPOSITORY."

[5] S. Rathore, P. K. Sharma, and J. H. Park, "XSSClassifier: An efficient XSS attack detection approach based on machine learning classifier on SNSs," *J. Inf. Process. Syst.*, vol. 13, no. 4, pp. 1014–1028, 2017, doi: 10.3745/JIPS.03.0079.

[6] F. A. Mereani and J. M. Howe, "Detecting Cross-Site Scripting Attacks Using Machine Learning," *Adv. Intell. Syst. Comput.*, vol. 723, pp. 200–210, 2018, doi: 10.1007/978-3-319-74690-6_20.

[7] S. F. Abdoh, M. Abo Rizka, and F. A. Maghraby, "Cervical cancer diagnosis using random forest classifier with SMOTE and feature reduction techniques," *IEEE Access*, vol. 6, pp. 59475–59485, 2018, doi: 10.1109/ACCESS.2018.2874063.

[8] M. A. Akbar *et al.*, "An Empirical Study for PCA- and LDA-Based Feature Reduction for Gas Identification," *IEEE Sens. J.*, vol. 16, no. 14, pp. 5734–5746, 2016, doi: 10.1109/JSEN.2016.2565721.

[9] F. M. M. Mokbal, W. Dan, A. Imran, L. Jiuchuan, F. Akhtar, and W. Xiaoxi, "MLPXSS: An Integrated XSS-Based Attack Detection Scheme in Web Applications Using Multilayer Perceptron Technique," *IEEE Access*, vol. 7, pp. 100567–100580, 2019, doi: 10.1109/access.2019.2927417.

[10] Q. Li, F. Wang, J. Wang, and W. Li, “LSTM-Based SQL Injection Detection Method for Intelligent Transportation System,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4182–4191, 2019, doi: 10.1109/TVT.2019.2893675.

[11] Y. Li and B. Zhang, “Detection of SQL Injection Attacks Based on Improved TFIDF Algorithm,” *J. Phys. Conf. Ser.*, vol. 1395, no. 1, 2019, doi: 10.1088/1742-6596/1395/1/012013.

[12] P. Tang, W. Qiu, Z. Huang, H. Lian, and G. Liu, “Knowledge-Based Systems Detection of SQL injection based on artificial neural network,” *Knowledge-Based Syst.*, vol. 190, p. 105528, 2020, doi: 10.1016/j.knosys.2020.105528.

[13] B. K. Ayeni, J. B. Sahalu, and K. R. Adeyanju, “Detecting Cross-Site Scripting in Web Applications Using Fuzzy Inference System,” *J. Comput. Networks Commun.*, vol. 2018, 2018, doi: 10.1155/2018/8159548.

[14] Y. Fang, “WOVSQLI : Detection of SQL Injection Behaviors Using Word Vector and LSTM,” pp. 170–174, 2018.

[15] A. Luo, W. Huang, and W. Fan, “A CNN-based Approach to the Detection of SQL Injection Attacks,” *2019 IEEE/ACIS 18th Int. Conf. Comput. Inf. Sci.*, pp. 320–324, 2019.

[16] M. Hasan, Z. Balbahaith, and M. Tarique, “Detection of SQL Injection Attacks: A Machine Learning Approach,” *2019 Int. Conf. Electr. Comput. Technol. Appl. ICECTA 2019*, 2019, doi: 10.1109/ICECTA48151.2019.8959617.

[17] S. O. Uwagbole, W. J. Buchanan, and L. Fan, “Applied Machine Learning predictive analytics to SQL Injection Attack detection and prevention,” *Proc. IM 2017 - 2017 IFIP/IEEE Int. Symp. Integr. Netw. Serv. Manag.*, pp. 1087–1090, 2017, doi: 10.23919/INM.2017.7987433.

[18] A. K. Dalai and S. K. Jena, “Neutralizing SQL injection attack using server side code modification in web applications,” *Secur. Commun. Networks*, vol. 2017, 2017, doi: 10.1155/2017/3825373.

[19] Z. C. S. S. Hlaing and M. Khaing, “A Detection and Prevention Technique on SQL Injection Attacks,” *2020 IEEE Conf. Comput. Appl. ICCA 2020*, pp. 1–6, 2020, doi: 10.1109/ICCA49400.2020.9022833.

[20] O. C. Abikoye, A. Abubakar, A. H. Dokoro, O. N. Akande, and A. A. Kayode, “A novel technique to prevent SQL injection and cross-site scripting attacks using Knuth-Morris-Pratt string match algorithm,” *Eurasip J. Inf. Secur.*, vol. 2020, no. 1, 2020, doi: 10.1186/s13635-020-00113-y.

[21] B. Nagpal, N. Chauhan, and N. Singh, “SECSIX: security engine for CSRF, SQL injection and XSS attacks,” *Int. J. Syst. Assur. Eng. Manag.*, vol. August, pp. 631–644, 2017, doi: 10.1007/s13198-016-0489-0.

[22] X. Chen and X. Hao, “Feature reduction method for cognition and classification of IoT devices based on artificial intelligence,” *IEEE Access*, vol. 7, pp. 103291–103298, 2019, doi: 10.1109/ACCESS.2019.2929311.

[23] H. Zhao, J. Zheng, J. Xu, and W. Deng, “Fault diagnosis method based on principal component analysis and broad learning system,” *IEEE Access*, vol. 7, pp. 99263–99272, 2019, doi: 10.1109/ACCESS.2019.2929094.

[24] W. Lei, J. Fang, and Y. Zhao, “Based on principal component analysis odor character classification and application,” *Proc. 28th Chinese Control Decis. Conf. CCDC 2016*, pp. 3757–3761, 2016, doi: 10.1109/CCDC.2016.7531638.

[25] Y. Yan, R. Liu, Z. Ding, X. Du, J. Chen, and Y. Zhang, “A parameter-free cleaning method for SMOTE in imbalanced classification,” *IEEE Access*, vol. 7, pp. 23537–23548, 2019, doi: 10.1109/ACCESS.2019.2899467.

[26] I. Ahmad, M. Basheri, M. J. Iqbal, and A. Rahim, “Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection,” *IEEE Access*, vol. 6, no. c, pp. 33789–33795, 2018, doi: 10.1109/ACCESS.2018.2841987.

[27] S. Waskle, L. Parashar, and U. Singh, “Intrusion Detection System Using PCA with Random Forest Approach,” *Proc. Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2020*, no. Icesc, pp. 803–808, 2020, doi: 10.1109/ICESC48915.2020.9155656.

[28] Y. Liu, Y. Wang, and J. Zhang, “New machine learning algorithm: Random forest,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7473 LNCS, pp. 246–252, 2012, doi: 10.1007/978-3-642-34062-8_32.

[29] A. Tharwat, “Classification assessment methods,” *Appl. Comput. Informatics*, 2018, doi: 10.1016/j.aci.2018.08.003.

[30] M. Kabir, S. Ahmad, M. Iqbal, Z. N. Khan Swati, Z. Liu, and D. J. Yu, “Improving prediction of extracellular matrix proteins using evolutionary information via a grey system model and asymmetric under-sampling technique,” *Chemom. Intell. Lab. Syst.*, vol. 174, no. December 2017, pp. 22–32, 2018, doi: 10.1016/j.chemolab.2018.01.004.

[31] M. Bach, A. Werner, J. Żywiec, and W. Pluskiewicz, “The study of under- and over-sampling methods’ utility in analysis of highly imbalanced data on osteoporosis,” *Inf. Sci. (Ny).*, vol. 384, pp. 174–190, 2017, doi: 10.1016/j.ins.2016.09.038.

[32] D. Ding, S. Han, H. Zhang, Y. He, and Y. Li, “Predictive biomarkers of colorectal cancer,” *Comput. Biol. Chem.*, vol. 83, no. January, 2019, doi: 10.1016/j.compbiolchem.2019.107106.