TRANSMISI: JURNAL ILMIAH TEKNIK ELEKTRO, 27, (3), JULI 2025
p-ISSN 1411-0814 e-ISSN 2407-6422

SELEKSI FITUR DENGAN GENETIKA ALGORITHM UNTUK OPTIMASI RANDOM FOREST DALAM MENGKLASIFIKASI TENTANG JENIS PENGGUNAAN INTERNET

Nadio Rodicky[*)], Sopian Soim dan Suzan Zefi

Sarjana Terapan Jurusan Teknik Elektro Program Studi Teknik Telekomunikasi,
Politeknik Negeri Sriwijaya, Palembang, Indonesia

[*)]*E-mail: nadiorodicky1717@gmail.com, sopiansoim@gmail.com, suzanzefi250977@gmail.com*

Abstrak

Penggunaan internet yang semakin meluas menimbulkan kebutuhan untuk memahami pola dan jenis aktivitas yang dilakukan pengguna. Klasifikasi jenis penggunaan internet dapat membantu dalam berbagai aplikasi, seperti manajemen jaringan, analisis web, dan pengembangan layanan internet yang lebih personal. Penelitian ini bertujuan untuk mengoptimalkan model klasifikasi jenis penggunaan internet menggunakan seleksi fitur dengan algoritma genetika dan Random Forest. Dataset yang digunakan dalam penelitian ini berisi informasi tentang lalu lintas internet, seperti alamat IP, nama domain yang dikunjungi, protokol yang digunakan, durasi kunjungan, dan jenis aktivitas. Algoritma genetika digunakan untuk menyeleksi subset fitur yang optimal untuk klasifikasi, sedangkan Random Forest digunakan sebagai algoritma klasifikasi. Hasil penelitian menunjukkan bahwa seleksi fitur dengan algoritma genetika dapat meningkatkan akurasi klasifikasi Random Forest secara signifikan. Penelitian ini memberikan kontribusi dalam pengembangan metode klasifikasi jenis penggunaan internet yang lebih akurat dan efisien. Hasil penelitian ini juga dapat digunakan sebagai dasar untuk penelitian selanjutnya dalam bidang analisis jaringan dan pengembangan layanan internet.

Kata kunci: Klasifikasi jenis penggunaan internet, seleksi fitur, algoritma genetika, Random Forest, akurasi

Abstract

The increasingly widespread use of the internet has created a need to understand the patterns and types of activities carried out by users. Classification of types of internet use can help in various applications, such as network management, web analysis, and the development of more personalized internet services. This research aims to optimize the classification model for types of internet use using feature selection with genetic algorithms and Random Forest. The dataset used in this research contains information about internet traffic, such as IP addresses, domain names visited, protocols used, duration of visits, and type of activity. Genetic algorithms are used to select an optimal subset of features for classification, while Random Forest is used as a classification algorithm. The research results show that feature selection with a genetic algorithm can significantly increase the accuracy of Random Forest classification. This research contributes to the development of a more accurate and efficient classification method for types of internet use. The results of this research can also be used as a basis for further research in the field of network analysis and internet service development.

Keywords: Classification of types of internet use, feature selection, genetic algorithm, Random Forest, accuracy

1. Pendahuluan

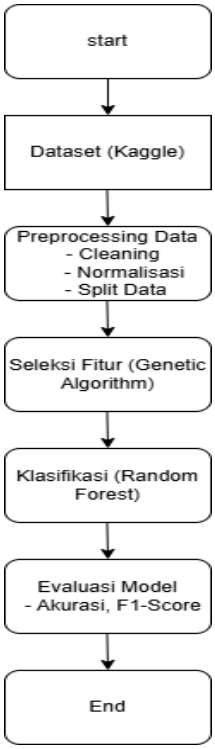
Penggunaan internet sekarang sudah menjadi bagian penting dalam kehidupan sehari-hari. Pertumbuhan pengguna internet yang pesat, terutama di era digital saat ini, penggunaan internet tidak hanya digunakan untuk mencari informasi, tetapi juga melakukan berbagai aktivitas online, Jumlah data internet yang dihasilkan pengguna cukup besar dan kompleks, juga mencakup berbagai jenis aktivitas seperti,streaming bermain game, dan transaksi e -commerce[1]. Internet adalah singkatan dari Interconnected Networking yang apabila diartikan dalam Bahasa Indonesia beararti sebuah rangkaian computer yang terhubung didalam beberapa rangkaian suatu jaringan[2]. Random Forest merupakan sebuah metode yang dikembangkan dari metode CART (Classification and Regression Trees), yang juga merupakan suatu metode atau algoritma dari Teknik pohon keputusan yang membedakan metode random forest dari metode CART ini adalah random forest menerapkan metode bootstrap aggregating (bagging) dan juga seleksi fitur random atau disebut dengan random feature selection[3]. Random Forest adalah sebuah model ansambel yang terdiri dari pohon keputusan yang dibangun dan dihasilkan menggunakan acak vector dari seluruh set fitur. Algoritma random forest sendiri paling cocok untuk menangani besar volume ulasan tekstual yang tidak terstruktur. Berdasarkan arsitektur ansambel, pertimbangkan target prediksi suara tinggi dan juga sebagai output klasifikasi akhir. Klasifikasi merupakan proses pembagian atau pengelompokkan sesuatu menjadi

kategori atau jenis didasari kesamaan sebuah pola atau sifat yang terdapat pada setiap data. Seleksi fitur dilakukan agar dapat mengurangi dimensi data dan fitur yang tidak relevan[4]. Salah satu metode yang efektif untuk memilih fitur adalah Algoritma Genetika. Algoritma Genetika merupakan metode optimasi evolusioner yang mempercepat proses seleksi alam. Algoritma ini dapat digunakan untuk menemukan kombinasi fitur terbaik yang dapat meningkatkan kinerja klasifikasi suatu model. Algoritma Genetika berkembang seiring perkembangan teknologi informasi yang sangat pesat. Kemampuan dari Algoritma Genetika ini adalah dapat menyelesaikan permasalahan yang sangat kompleks[5]. Semua pengetahuan machine learning pasti akan melibatkan data[6]. Keberhasilan pembelajaran mesin IBM telah memungkinkan Deep Blue menarik perhatian dunia . Akibatnya , pengembangan ML semakin didorong oleh kebutuhan manusia[7]. Kecerdasan Buatan (AI) yang sering dikenal dengan kecerdasan mesin merupakan suatu sistem komputer yang dilatih dengan tingkat kecerdasan hamper menyerupai manusia[8]. Dalam peggunaannya, AI dapat membantu pekerjaan manusia, namun pengguna juga perlu memahami kelebihan dan kekurangan AI sehingga dapat lebih bijak menggunakannya, AI telah membuka pintu menuju masa depan yang lebih canggih dan inovatif, dan teknologi ini terus berkembang dan memiliki potensi untuk merubah cara kita hidup dan bekerja secara fundamental[9]. Algoritma Genetika merupakan metode yang digunakan mampu menyeleksi fitur yang bisa meningkatkan hasil akurasi dari klasifikasi[10]. Algoritma Genetika akan terus melakukan proses seleksi, rekombinasi, dan mutasi pada populasi hingga ditemukan solusi optimal atau kriteria penghentian terpenuhi. Dengam cara ini, algoritma genetika bisa menemukan solusi yang terbaik dalam permasalahan yang kompleks[11]. Diharapkan dengan memanfaatkan Random Forest dan Algoritma Genetika akan diperoleh model klasifikasi yang lebih akurat dan efisien dalam menganalisis penggunaan internet. Tujuan dari penelitian ini adalah untuk mengembangkan metode yang menggabungkan Random Forest dengan Algoritma Genetika untuk mengoptimalkan pemilihan fitur dalam klasifikasi penggunaan internet. Random Forest adalah kombinasi dari masing masing teknik pohon keputusan yang ada, lalu kemudian digabung dan dikombinasikan kedalam suatu model[12]. Vektor ini didefenisikan sebagai konstruksi tree. Misalnya, dalam pemilihan acak, vektor- vektor ini terdiri dari bilangan bulat acak yang dipilih secara acak dari {1,...,N} dimana N adalah jumlah pemisahan[13]. Konsep voting yang terjadi tanpa disadari ini dikenal dengan nama majority voting[14]. Dengan demikin, Random Forest menggabungkan keluaran dari beberapa decision tree untuk mencapai satu hasil yang lebih akurat dan stabil, sehngga mengurangi resiko overfitting dan meningkatkan kinerja model dalam melakukan pengklasifikasian dan regeresi[15]. Dengan cara ini, diharapkan penelitian ini akan memberikan kontribusi signifikan terhadap analisis data dan pengembangan hipotesis berbasis data. Untuk itu penulis tertarik membahas mengenai "Seleksi fitur menggunakan algoritma genetika untuk optimasi random forest dalam mengklasifikasi jenis penggunaan internet".

2. Metode

Penelitian ini terdiri dari beberapa tahapan yang dirancang untuk menghasilkan suatu sistem yang dapat berfungsi secara optimal. Adapun tahapan-tahapan tersebut adalah sebagai berikut:

2.1. Perancangan Penelitian Model Machine Learning



Gambar 1 Blok diagram cara kerja

Kerangka tahapan perancangan model machine learning terdiri dari beberapa alur tahapan dan akan dibuat dalam bentuk diagram yang memberikan gambaran mengenai alur pembuatan model. Diagram ini bertujuan untuk mengetahui langkah-langkah penelitian secara menyeluruh mulai dari awal pengumpulan data hingga selesai pembuatan model machine learning. Tahapan untuk pembuatan model machine learning dibagi menjadi beberapa tahapan, tahapan pertama yaitu pencarian dataset yang sesuai melalui Kaggle dalam tahapan ini mencari sebuah dataset yang berkaitan dengan judul diatas. Tahapan kedua yaitu PreProcessing data dimana data yang sudah ada sebelumnya dilakukan pembersihan dan penyesuain sebelum nanti digunakan untuk melakukan pelatihan model. Kemudian dilakukan tahapan normalisasi data dan selanjutnya dilakukan proses training data dan testing, kemudian masuk ke tahapan seleksi fitur dengan

menggunakan algoritma genetika setelah itu lanjut ke proses random forest untuk melakukan proses deteksi sampe optimal.

2.1.1. Dataset Penelitian

Penelitian ini akan menggunakan data yang berkaitan dengan jenis penggunaan internet serta pastikan dataset mencakup informasi yang relevan untuk klasifikasi, seperti fitur-fitur yang menggambarkan penggunaan internet, selanjutnya mengumpulkan dataset yang berisi informasi tentang penggunaan internet seperti contoh, jenis aktivitas (browsing, streaming, gaming), durasi kunjungan serta jumlah data yang diunduh/diunggah. Sebelum digunakan harus dilakukan pre-processing data terlebih dahulu untuk mengidentifikasi kumpulan data yang relevan untuk pelatihan model machine learning. Setelah dilakukan tahapan pre-processing selesai itu dilanjutkan dengan dataset akan dibagi menjadi data training dan data testing.

2.1.2. Tahap Pre-Processing Data

Pada tahapan ini akan dilakukan pada dataset sebelum digunakan untuk melatih machine learning. Tahapan yang dilakukan dalam proses ini seperti melakukan cleaning data, normalisasi data dan pembagian dataset menjadi data training dan data testing. Data yang sudah didapatkan melalui Kaggle akan dilakukan tahapan pre-processing data dengan tahapan awal yaitu pembersihan data, berupa membersihkan data yang tidak valid atau terdapat missing value didalmnya, pembersihan data dilakukan dengan melakukan pengecekan didalam dataset. Didalam dataset tersebut masih terdapat nilai kosong seperti Nan (Not a Number). Null dan tanda tanya (?). Untuk data yang memiliki nilai seperti itu akan dilakukan penghapusan pada baris tersebut agar tidak terjadi eror pada saat dilakukan pembacaan oleh machine learning pada saat dilakukan proses pelatihan data.

Yang kedua yaitu tahap normalisasi data, normalisasi merupakan sebuah proses mengubah nilai-nilai dalam dataset agar berada dalam suatu skala yang sama, tujuan dari normalisasi ini untuk memastikan bahwa setiap fitur berkontribusi secara proporsional terhadap model yang dibangun. Dan untuk membuat suatu model data agar dapat membacanya dengan benar dan efisien. Tahap terakhir yaitu seleksi fitur atau pemilihan fitur. Fitur selection merupakan salah satu langkah penting dalam proses machine leaning yaitu mengidentifikasi subset fitur atau variable yang relevan dan menghilangkan yang tidak relevan atau berlebihan. Pemilihan fitur yang sesuai dapat meningkatkan kinerja model, mencegah overfitting, serta mempercepat proses pelatihan dan membuat model lebih interpretative. Tahapan ini akan sangat berpengaruh terhadap kinerja model, sebab semakin banyak fitur yang ditinjau maka akan membutuhkan waktu yang lebih lama, oleh karena itu dilakukan seleksi fitur agar model tidak meninjau seluruh fitur melainkan hanya beberpa fitur yang telah ditentukan.

2.1.3. Training dan Testing Data

Training data merupakan subset dari dataset digunakan untuk melatiih model machine learning. Tujuan dari data pelatihan ini adalah untuk menyediakan model dengan informasi yang akurat sehingga dapat memahami pola,hubungan, dan karakteristik yang ditemukan dalam data. Fungsi dari training data yaitu, model belajar dari training data untuk menemukan bahwa parameter tersebut sudah optimal dan dapat membantu model untuk mengidentifikasi hubungan antara input (fitur) dan output (target). Proses training dilakukan dengan model diinisialisasi dengan parameter awal, dan parameter model sudah disesuaikan berdasarkan feedback dari kesalahan yang dihitung. Sedangkan testing data merupakan subset dari dataset yang digunakan untuk memuji dan mengevaluasi sebuah kinerja model setelah dilatih dengan training data. Testing data tidak digunakan selama proses pelatihan untuk memberikan estimasi proses pelatihan tidak bias terhadap kinerja model pada data yang tidak transparan. Fungsi utama dari testing data yaitu untuk mengukur kemampuan generalisasi model ke data bar yang belum pernah dilihat selama pelatihan, serta menilai akurasi, presisi, recall, f1 score, atau metrik evaluasi lainnya.

2.2. Feature Selection menggunakan Algoritma Genetika

Algoritma Genetika merupakan teknik atau metode pengoptimalan yang menggunakan proses seleksi alam dalam teori evolusi untuk menemukan solusi optimal dari suatu masalah. Pemilihan fitur dalam GA untuk melakukan fitur seleksi melibatkan beberapa langkah utama yaitu : inisialisasi, seleksi, crossover, dan mutasi. Tujuan dari pemilihan fitur ini yaitu mengurangi jumlah fitur atau variable input dengan memilih fitur-fitur tersebut yang paling relevan terhadap model, mengurangi overfitting, dan noise dalam model, serta dapat meningkatkan akurasi dan mengurangi beban serta waktu training. Berikut merupakan langkah-langkah dalam melakukan pemilihan fitur dengan GA.

1. Inisialisasi (Initialization)
   Pada tahapan ini, dilakukan populasi awal individu(solusi potensial) dihasilkan secara acak. Setiap individu biasanya direpresentasikan sebagai kromosok yang terdiri dari sekumpulan gen atau variable.
2. Seleksi (selection)
   Untuk tahapan ini, individu-individu dalam populasi dipilih sesuai nilai fitness mereka. Individu dengan nilai fitness lebih tinggi memiliki peluang cukup besar dan dipilih sebagai orang tua yang akan dihasilkan dari keturunan.

3. Crossover
   Crossover atau rekombinasi merupakan proses dimana dua individu dipilih dan sebagian dari kromosom mereka dipertukarkan untuk membentuk individu baru atau keturunan. Crossover juga membantu menghasilkan variasi dalam sebuah populasi.
4. Mutasi (Mutation)
   Mutasi merupakan sebuah proses dimana beberapa gen dalam individu ini diubah secara acak untuk mempertahankan keragaman genetic dalam populasi dan juga mencegah konvergensi premature.
5. Evaluasi Fitness
   Untuk melakukan sebuah kinerja model dengan subset fitur yang akan dipilih oleh setiap kromosom contohnya, Random Forest. Dan setelah itu hitung nilai fitness berdasarkan akurasi yang telah dicapai atau metrik kinerja lainnya. Serta ulangi proses seleksi, crossover, dan mutasi hingga mencapai jumlah generasi yang ditentukan.

Dengan menggunakan genetic algorithm untuk melakukan feature selection, kita dapat menemukan subset ftur yang paling relevan dan juga dapat meningkatkan kinerja model dalam melakukan klasifikasi jenis penggunaan internet.

2.3. Training Menggunakan Random Forest

Training menggunakan random forest dalam melakukan klasifikasi jenis penggunan internet ini melibatkan beberapa langkah, termasuk persiapan data, pelatihan model dan evaluasi kinerja model. Berikut ini adalah penjelasan untuk melakukan training yaitu:

1. Persiapan data dimana ini dilakukan sebelum melakukan proses training dimulai dan juga dilakukan pengumpulan data, dataset yang berisikan fitur-fitur yang relevan dan label yang menetukan jenis pengguna internet. Setelah itu dilakukan Preprocessing yaitu melakukan preprocessing data seperti penanganan missing values, encoding untuk sebuah data kategorikal, dan normalisasi atau standarisasi fitur numerik.
2. Pembagian Data setelah itu dilakukan training dan testing yaitu dataset dibagi menjadi dua bagian training data untuk melatih model sementara testing data untuk mengevaluasi model.
3. Pelatihan model dilakukan sebuah Inisialisasi Model yaitu dengan cara buat instance model random forest serta dilakukan Pelatihan yaitu melatih model dengan menggunakan data training. Random forest bekerja dengan membuat banyak pohon keputusan atau decision trees dan menggabungkan untuk menghasilkan prediksi yang akurat.
4. Evaluasi Model selanjutanya gunakan model yang telah dilatih tersebut untuk membuat sebuah prediksi pada testing data. Serta dilakukan metode evaluasi dimana evaluasi kerja model menggunakan metrik seperti akurasi, precision, recall, dan f1-score
5. Visualisasi yaitu buat sebuah confusion matrix untuk melihat performa klasifikasi dan Feature Importance menampilkan bagaimana pentingnya fitur-fitur untuk melihat fitur mana yang paling berpengaruh dalam klasifikasi.

2.4. Evaluasi Model Menggunakan Data Testing

Setelah dilakukan membaca data, mempersiapkan data dan melatih model klasifikasi, selanjutnya dilakukan evaluasi terhadap performa model pada data testing. Seperti nilai pada akurasi dan laporan klasifikasi, dapat digunakan untuk menilai seberapa baik model dapat melakukan prediksi jenis penggunaan internet pada data baru (set pengujian). Evaluasi ini penting untuk memahami kekuatan dan kelemahan model, serta menentukan apakah perlu dilakukan penyesuaian atau pelatihan model ulang.

3. Hasil dan Pembahasan

hasil pengujian dan implementasi algoritma Random Forest (RF) dan algoritma genetika (GA) untuk seleksi fitur dalam konteks klasifikasi jenis penggunaan internet akan dibahas secara mendalam. Eksperimen dilakukan untuk mengevaluasi performa model RF setelah melakukan seleksi fitur dengan GA. Berikut ini hasil dari Implementasi Algoritma RF dan GA pada tahapan ini, pengujian dilakukan untuk mempreoleh implementasi sistem dengan menggunakan algoritma Random Forest dan Algoritm Genetika dalam proses seleksi fitur.

3.1. Analisis Hasil Seleksi Fitur

Hasil eksperimen menunjukkan bahwa terdapat beberapa kombinasi fitur yang memberikan nilai fitness tertinggi yaitu 1.0 yang di pengaruhi dengan jumlah data yang sangat banyak dan data yang baik untuk model random fores menggunakan seleksi fitur algoritma genetika. Hal ini menunjukkan bahwa model dengan kombinasi fitur tersebut memiliki kinerja yang sangat baik dalam mengklasifikasikan jenis penggunaan internet. Kombinasi fitur terbaik yang diperoleh pada eksperimen adalah: meningkatkan performa model klasifikasi dalam mengklasifikasikan jenis pada penggunaan internet. Fitur-fitur ini memiliki pengaruh yang signifikan terhadap kinerja model Random Forest dalam mengklasifikasikan jenis penggunaan internet. Dengan menggunakan Algoritma Genetika, dapat mengeliminasi fitur-fitur yang kurang penting dan fokus pada fitur-fitur yang memberikan kontribusi terbesar pada kinerja model.

3.2. Evaluasi Kinerja Model

Untuk mengevaluasi kinerja model, dapat menggunakan metrik akurasi. Berdasarkan hasil eksperimen, kombinasi fitur terbaik memberikan nilai akurasi yang sangat tinggi, yaitu 1.0. Hal ini menunjukkan bahwa model berhasil mengklasifikasikan jenis penggunaan internet dengan

benar untuk semua data yang diuji. Selain itu, beberapa kombinasi fitur lainnya juga memberikan nilai fitness yang bervariasi antara 0.33 hingga 1.0. Hal ini menunjukkan bahwa meskipun beberapa fitur dihilangkan, model masih dapat bekerja dengan baik namun dengan akurasi yang sedikit lebih rendah. Kemampuan untuk mengidentifikasi fitur- fitur yang paling signifikan dan relevan dalam mengklasifikasikan tipe pengguna internet. Dengan menerapkan seleksi fitur menggunakan algoritma genetika ini, membuktikan bahwa model Random Forest yang dihasilkan mencapai akurasi 100%. Hal ini menunjukkan bahwa fitur - fitur tertentu yang dianggap cukup amemberikan kontribusi

3.3. Implementasi Hasil Eksperimen

Hasil eksperimen ini menunjukkan bahwa penggunaan Algoritma Genetika untuk seleksi fitur dapat secara signifikan meningkatkan kinerja model Random Forest. Dengan mengeliminasi fitur-fitur yang tidak relevan, dapat mengurangi kompleksitas model dan meningkatkan akurasi prediksi. Penggunaan kombinasi fitur terbaik dapat membantu dalam membuat keputusan yang lebih tepat dalam berbagai aplikasi, seperti pengoptimalan jaringan, pengelolaan sumber daya, dan peningkatan kualitas layanan internet.

4. Kesimpulan

Pada bab ini akan dibahas mengenai kesimpulan serta saran berdasarkan hasil dari analisa yang telah dilakukan sebelumnya. Algoritma Genetika terbukti efektif dalam mengoptimalkan kinerja model Random Forest dengan melakukan seleksi fitur. Algoritma genetika memili kemampuan untuk mengidentifikasi fitur-fitur yang paling signifikan dan frelevan dalam mengklasifikasi tipe pengguna intenet. Dataset diambil dari Kaggle karena platform ini menyediakan data terbuka yang relevan, bersih, dan telah digunakan dalam berbagai studi sebelumnya, memungkinkan replikasi dan pembandingan hasil penelitian. Dengan menerapkan seleksi fitur menggunakan algoritma genetika ini, membuktikan bahwa model Random Forest yang dihhasilkan mencapai akuras. 100%. Hal ini menunjukan bahwa fitur-fitur tertentu yang dianggap cukup memberikan kontribusi yang signifikan terhadap kinerja model, sehingga mampu mengevaluasi berbagai jenis oengguna internet secara akurat. Memanfaatkan algoritma genetika untuk pemiliham fitur tidak hanya meningkatkan akurasi tetapi juga mengurangi kompleksitas model. Adapun saran untuk penelitian selnjutnya yaitu, Meskipun hasil nya menunjukan akurasi 100%, penting untuk memvalidasi kinerja model menggunakan dataset yang berbeda hasil kondisi yang berbeda untuk memastikan generalisasi model dan mencegah overfitting. Untuk selanjutnya dapat menggunakan kombinasi ini dalam berbagai machine learning yang lebih memerlukan klasifikasi yang akurat.

Referensi

[1]. M. T. Astuti, “Analisis Big Data Menggunakan Teknik Machine Learning Untuk Prediksi Pola Perilaku Pengguna,” J. Dunia Data, vol. 1, no. 4, pp. 1–15, 2024, [Online] Available:

[2]. Susana, H. (2022). Application of Naive Bayes Method Classification Model on Internet Access Usage. Journal of Information Systems and Information Technology Research (JURSISTEKNI), 4(1), 1-8. [2] X. Li, X.Zhang, dan J. Zhang.

[3]. E. P. Cynthia, M. Afif Rizky A., A. Nazir, and F. Syafria, “Random Forest Algorithm to Investigate the Case of Acute Coronary Syndrome,” J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 5, no. 2, pp. 369–378, 2021, doi: 10.29207/resti.v5i2.3000.

[4]. G. E. Riani and W. F. Mahmudy, “Optimasi Jangkauan Jaringan 4G Menggunakan Algoritma Genetika,” J. Teknol. Inf. dan Ilmu Komput., vol. 3, no. 2, p. 141, 2016, doi: 10.25126/jtiik.201632205.

[5]. Amini, N., Saragih, T. H., Faisal, M. R., Farmadi, A., & Abadi, F. (2022). Implementasi Algoritma Genetika untuk Seleksi Fitur pada Klasifikasi Genre Musik menggunakan Metode Random Forest. Jurnal Informatika Polinema, 9(1), 75-82.

[6]. Mahesh, Batta. "Machine learning algorithms-a review." International Journal of Science and Research (IJSR).[Internet] 9.1 (2020): 381-386

[7]. J. D. Pineda-Jaramillo, “A review of machine learning (ML) algorithms used for modeling travel mode choice•,” DYNA, vol. 86, no. 211, pp. 32–41, 2019, doi: 10.15446/dyna.v86n211.79743.

[8]. S. K. Kwofie et al., “Artificial Intelligence, Machine Learning, and Big Data for Ebola Virus Drug Discovery,” Pharmaceuticals, vol. 16, no. 3, pp. 1–12, 2023, doi: 10.3390/ph16030332.

[9]. Masrichah, S. (2023). Ancaman Dan Peluang Artificial Intelligence (AI). Khatulistiwa: Jurnal Pendidikan dan Sosial Humaniora, 3(3), 83-101.

[10]. K. A. Santoso, B. A. Setiawan, and K. Kusbudiono, “Application of Genetic Algorithm on Inclusive Labeling of a Graph,” Inpr. Indones. J. Pure Appl. Math., vol. 4, no. 1, pp. 24–32, 2022, doi: 10.15408/inprime.v4i1.24327.

[11]. E.P Cynthia, M. Afif Rizky A., A. Nazir, and F. Syafria, “Random Forest Algorithm to Investigate the Case of Acute Coronary Syndrome,” J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 5, no. 2, pp. 369–378, 2021, doi: 10.29207/resti.v5i2.3000.

[12]. Georganos et al., “Geographical random forests: a spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling,” Geocarto Int., vol. 36, no. 2, pp. 121–136, 2021, doi: 10.1080/10106049.2019.1595177.

[13]. A. More and P. Rana, “Review of Random Forest Classification Techniques to Resolve Data Imbalance: proceedings: 5-6 October 2017, Aurangabad, Maharashtra, India,” 1st Int. Conf. Intell. Syst. Inf. Manag. ICISIM-2017, pp. 72– 78, 2017.

[14]. O.- Pahlevi, A.- Amrin, and Y.- Handrianto, "Implementasi Algoritma Klasifikasi Random Forest Untuk Penilaian Kelayakan Kredit," J. Infortech, vol. 5, no. 1, pp. 71–76, 2023, doi: 10.31294/infortech.v5i1.15829.

[15]. Kurniabudi, K., Harris, A., & Rahim, A. (2020). Seleksi Fitur Dengan Information Gain Untuk Meningkatkan Deteksi Serangan DDoS menggunakan Random Forest. Techno. Com, 19(1), 56-66.