

## ANALISIS KARIR ALUMNI MENGGUNAKAN RANDOM FOREST DENGAN INTERPRETASI SHAP DAN EVALUASI BIAS MODEL FAIRNESS

Irfan Abadi Saragih<sup>1</sup>, Arip Muhridan<sup>2</sup>, Andisyah Putra<sup>3</sup>, Ridwan<sup>4</sup>, Mhd. Ihsan Abidi<sup>5</sup>

Universitas Pembangunan Panca Budi, Medan

e-mail: <sup>1</sup>irfanabadisaragih@gmail.com, <sup>2</sup>aripmuhridan14@gmail.com, <sup>3</sup>andizarx@gmail.com, <sup>4</sup>ayrinridwan@gmail.com, <sup>5</sup>mhdihsanabidi27@gmail.com

**Abstract:** Higher education institutions are required to produce graduates who are competitive in the job market, yet tracer study data is often analyzed only descriptively, leaving deeper patterns unexplored. This study proposes an integrated approach combining Random Forest, SHAP (SHapley Additive exPlanations), and fairness evaluation to analyze alumni career outcomes. Data were sourced from the 2024 tracer study of Universitas Royal, comprising 466 alumni records with a binary target variable indicating employment status. Preprocessing included missing value handling, one-hot encoding, and stratified train-test splitting with an 80:20 ratio. Class imbalance was addressed using the `class_weight='balanced'` parameter. The Random Forest model achieved an accuracy of 95.7%, F1-score of 0.946, and AUC-ROC of 0.938, significantly outperforming the dummy baseline of 58.5%. Cross-validation across five folds yielded a mean accuracy of 96.4%, confirming model stability. SHAP analysis revealed that features related to educational level appropriateness and relevance of study field to current occupation were the most dominant predictors. Fairness evaluation per study program showed minimal performance disparity between Sistem Informasi (94.7%) and Sistem Komputer (100%), indicating no significant predictive bias. This framework demonstrates that explainable and fair machine learning can serve as a reliable foundation for evidence-based academic policy making.

**Keyword:** random forest; SHAP; fairness; tracer study; alumni career

**Abstrak:** Perguruan tinggi dituntut untuk menghasilkan lulusan yang berdaya saing di pasar kerja, namun data tracer study seringkali hanya dianalisis secara deskriptif sehingga pola yang lebih mendalam belum tergali secara optimal. Penelitian ini mengusulkan pendekatan terpadu yang menggabungkan Random Forest, SHAP (SHapley Additive exPlanations), dan evaluasi fairness untuk menganalisis luaran karir alumni. Data bersumber dari tracer study tahun 2024 Universitas Royal yang terdiri dari 466 data alumni dengan variabel target biner status kekerjaan. Praproses mencakup penanganan missing value, one-hot encoding, dan pemisahan data stratifikasi dengan rasio 80:20. Ketidakseimbangan kelas ditangani menggunakan parameter `class_weight='balanced'`. Model Random Forest mencapai akurasi 95,7%, F1-score 0,946, dan AUC-ROC 0,938, jauh melampaui dummy baseline sebesar 58,5%. Cross-validation lima lipatan menghasilkan rata-rata akurasi 96,4%, mengkonfirmasi kestabilan model. Analisis SHAP mengungkap bahwa fitur kesesuaian tingkat pendidikan dan kecermatan hubungan bidang studi dengan pekerjaan merupakan prediktor paling dominan. Evaluasi fairness per program studi menunjukkan disparitas performa yang minimal antara Sistem Informasi (94,7%) dan Sistem Komputer (100%), mengindikasikan tidak adanya bias prediksi yang signifikan. Kerangka analisis ini menunjukkan bahwa machine learning yang dapat dijelaskan dan berkeadilan dapat menjadi fondasi yang andal bagi pengambilan kebijakan akademik berbasis bukti.

**Kata kunci:** random forest; SHAP; fairness; tracer study; karir alumni

## PENDAHULUAN

Perguruan tinggi memiliki tanggung jawab strategis dalam menghasilkan lulusan yang siap bersaing di pasar kerja. Salah satu instrumen utama untuk mengevaluasi keberhasilan tersebut adalah *tracer study*. *Tracer study* memberikan informasi yang relevan dan bermanfaat untuk menilai hasil pendidikan dan pelatihan suatu institusi pendidikan tinggi, sekaligus menjadi tolok ukur keberhasilan program studi dalam mempersiapkan lulusannya di pasar kerja yang dinamis (Subekti et al., 2021). Namun, data *tracer study* yang terkumpul seringkali hanya dianalisis secara deskriptif, sehingga potensi informasi yang lebih dalam mengenai pola dan faktor penentu karir alumni belum sepenuhnya dimanfaatkan.

Perkembangan pesat kecerdasan buatan dan *machine learning* membuka peluang baru dalam analisis data pendidikan. (Barutu et al., 2025) dalam tinjauan sistematis mereka menyimpulkan bahwa *Random Forest*, *Support Vector Machine (SVM)*, dan jaringan saraf tiruan merupakan teknik yang paling banyak digunakan dalam prediksi karir di pendidikan tinggi, mencerminkan adopsi model-model yang mampu menangani data multidimensi yang kompleks. Di antara berbagai algoritma tersebut, *Random Forest* menunjukkan performa yang konsisten unggul. Dalam konteks prediksi pendapatan alumni, (Helmy et al., 2025) menemukan bahwa algoritma berbasis pohon seperti *Random Forest* mampu mengidentifikasi pola dalam distribusi pendapatan dan mengungkap prediktor terkuat yang mempengaruhi karir lulusan.

Meskipun akurasi prediksi menjadi prioritas utama, transparansi model tidak kalah pentingnya dalam konteks pendidikan. (Parhusip et al., 2025) menegaskan bahwa model yang kompleks seperti jaringan saraf tiruan yang dalam sering berfungsi sebagai "black box", di mana sulit bagi pengguna untuk memahami bagaimana keputusan dibuat;

dalam konteks pendidikan, kurangnya interpretabilitas ini dapat membatasi kepercayaan terhadap sistem berbasis *machine learning*. Untuk mengatasi tantangan ini, metode SHAP hadir sebagai solusi dalam kerangka *Explainable Artificial Intelligence (XAI)* yang memungkinkan interpretasi model secara lokal maupun global (Usman et al., 2024).

Di luar persoalan interpretabilitas, isu bias dan keadilan (*fairness*) dalam model prediktif juga menjadi perhatian kritis. (Sitorus et al., 2023) memperingatkan bahwa apabila model dilatih menggunakan dataset yang mencerminkan bias historis atau sosial, model tersebut dapat mereplikasi ketidakadilan dalam prediksinya, yang pada akhirnya menghasilkan keputusan yang tidak adil atau diskriminatif dalam bimbingan karir. Penelitian terkini menunjukkan bahwa SHAP dan evaluasi *fairness* dapat dikombinasikan secara sinergis. (Usman et al., 2024) mendemonstrasikan bahwa pengukuran SHapley Additive exPlanations (SHAP) pada metode *fair learning* mutakhir mampu mengungkap diskriminasi melalui asosiasi atau diskriminasi terbalik (*reverse discrimination*), menjadikan kombinasi SHAP dan evaluasi *fairness* sebagai alat audit yang komprehensif.

Berdasarkan latar belakang di atas, penelitian ini bertujuan menganalisis pola dan faktor penentu karir alumni menggunakan algoritma *Random Forest*, dilengkapi interpretasi model melalui SHAP, serta evaluasi bias menggunakan kerangka model *fairness*. Pendekatan terpadu ini diharapkan menghasilkan model yang tidak hanya akurat, tetapi juga transparan dan berkeadilan, sehingga dapat menjadi landasan bagi pengambilan kebijakan akademik yang lebih tepat sasaran.

## METODE

### Alur Penelitian

Penelitian ini menggunakan

pendekatan kuantitatif berbasis educational data mining dengan data primer yang bersumber dari hasil tracer study alumni. Seluruh tahapan analisis dilaksanakan menggunakan bahasa pemrograman Python (versi 3.x) pada platform Google Colaboratory. Alur penelitian meliputi lima tahapan utama: (1) pengumpulan dan praproses data, (2) penanganan ketidakseimbangan kelas, (3) pembangunan model Random Forest, (4) interpretasi model menggunakan SHAP, dan (5) evaluasi bias menggunakan kerangka model fairness.

### Pengumpulan dan Pemrosesan Data

Data dikumpulkan melalui survey tracer study yang disebarakan kepada alumni oleh Lembaga Pusat Layanan Karir Universitas Royal. Variabel yang dikumpulkan meliputi atribut akademik (NIM, program studi, masa studi) serta atribut karir (status pekerjaan, bidang pekerjaan, waktu tunggu kerja, dan kesesuaian bidang kerja dengan program studi). Praproses data dilakukan menggunakan library pandas dan numpy, mencakup penanganan missing value, encoding variabel kategorikal menggunakan label encoding dan *one-hot encoding*, serta normalisasi fitur numerik. Pemisahan data latih dan data uji dilakukan dengan rasio 80:20 menggunakan fungsi *train\_test\_split* dari modul *sklearn.model\_selection* dengan parameter *stratify=y* untuk mempertahankan proporsi kelas pada kedua subset data.

### Penanganan Ketidakseimbangan Kelas

Hasil analisis distribusi kelas menunjukkan bahwa data target bersifat tidak seimbang (*mild imbalance*), dengan proporsi kelas "Belum Bekerja" sebesar 58,2% dan "Sudah Bekerja" sebesar 41,8%. Untuk mengatasi hal ini tanpa mengubah distribusi data asli, model dibangun dengan parameter *class\_weight='balanced'* pada Random Forest Classifier. Pendekatan ini secara otomatis menyesuaikan bobot setiap kelas berbanding terbalik dengan frekuensinya

dalam data latih, sehingga kelas minoritas mendapat perhatian proporsional selama proses pelatihan. Pendekatan pembobotan kelas ini dipilih karena lebih sederhana dan tidak memperkenalkan data sintesis ke dalam proses pelatihan, sehingga distribusi data asli tetap terjaga (Ernawati et al., 2025).

### Algoritma Random Forest

Model klasifikasi dibangun menggunakan kelas *Random Forest Classifier* dari library *scikit-learn*. Puspa, (Aprilia et al., 2026) mengatakan bahwa random Forest merupakan algoritma berbasis ensemble learning yang bekerja dengan membangun banyak pohon keputusan (decision tree) dan menggabungkan hasil prediksi dari setiap pohon untuk menghasilkan prediksi akhir yang lebih akurat.

Model dibangun dengan 200 pohon keputusan (*n\_estimators=200*), tanpa pembatasan kedalaman pohon (*max\_depth=None*), dan menggunakan *class\_weight='balanced'* untuk menangani ketidakseimbangan kelas. Untuk memvalidasi kestabilan model, dilakukan pula Stratified K-Fold Cross Validation dengan *k=5* pada keseluruhan data, serta uji *shuffle baseline* menggunakan label yang diacak secara acak sebagai pembanding untuk memastikan model benar-benar belajar dari pola data.

### Evaluasi Performa Model

Evaluasi performa dilakukan pada data uji menggunakan metrik dari modul *sklearn.metrics*, meliputi *accuracy*, *precision*, *recall*, *F1-score*, dan *Area Under the ROC Curve (AUC-ROC)*. Mengingat adanya ketidakseimbangan kelas pada data asli, *macro-averaged F1-score* dan *AUC-ROC* diprioritaskan sebagai metrik utama, sebagaimana direkomendasikan oleh (Sholeha et al., 2024) dalam konteks analisis data alumni yang tidak seimbang. Visualisasi hasil evaluasi meliputi *confusion matrix* dan kurva ROC menggunakan library *matplotlib*.

### Interpretasi Model dengan SHAP

Interpretasi model dilakukan menggunakan library shap yang mengimplementasikan metode *SHapley Additive ExPlanations* (SHAP (Fatimah et al., 2025) menyatakan bahwa memahami mengapa model membuat prediksi tertentu bisa sama pentingnya dengan akurasi prediksi itu sendiri. SHAP mengukur kontribusi rata-rata setiap fitur terhadap prediksi model berdasarkan teori permainan kooperatif Shapley (Alfadia Shauqie et al., 2025). Karena model yang digunakan berbasis pohon, digunakan TreeExplainer dengan parameter `check_additivity=False` untuk menjaga stabilitas komputasi numerik.

### Evaluasi Bias dengan Analisis Fairness per Program Studi

Evaluasi bias dilakukan dengan menganalisis akurasi prediksi secara terpisah untuk setiap program studi sebagai atribut sensitif. (Parhusip et al., 2025) menegaskan bahwa apabila model dilatih menggunakan data yang mencerminkan bias historis atau sosial, model tersebut dapat mereplikasi bias tersebut dalam prediksinya, sehingga evaluasi fairness per kelompok menjadi langkah penting dalam audit model. Analisis hanya dilakukan pada program studi dengan jumlah sampel uji lebih dari 5 untuk memastikan representativitas hasil.

## HASIL DAN PEMBAHASAN

### Praproses dan Distribusi Data

Setelah melalui tahap praproses, dataset yang digunakan terdiri dari alumni dengan label target KERJA yang valid, yaitu 'Belum' (0) dan 'Sudah' (1). Hasil analisis distribusi kelas menunjukkan proporsi kelas "Belum Bekerja" sebesar 58,2% (271 data) dan "Sudah Bekerja" sebesar 41,8% (195 data). Proses one-hot encoding terhadap seluruh variabel kategorikal menghasilkan sejumlah fitur prediktor yang siap digunakan dalam pemodelan.

**Tabel 1 Data Set**

Status Kerja	Jumlah	Proporsi
Belum	271	58,2%
Sudah	195	41,8%
Total	466	100%

### Performa Model Random Forest

Model Random Forest dibangun dengan 200 pohon keputusan dan parameter `class_weight='balanced'` untuk menangani ketidakseimbangan kelas. Hasil evaluasi pada data uji (20% dari total data) disajikan pada Tabel 2.

**Tabel 2 Hasil Evaluasi Data Uji**

Metrik	Nilai
Accuracy	0.957
Precision	1.000
Recall	0.897
F1-score	0.946
ROC-AUC	0.938

Sebagai pembanding, DummyClassifier dengan strategi `most_frequent` menghasilkan akurasi baseline yang hanya memprediksi kelas mayoritas secara konstan.

Dummy accuracy: 0.5851063829787234

Selisih yang signifikan antara akurasi model Random Forest dan dummy baseline mengindikasikan bahwa model berhasil mempelajari pola dari data, bukan sekadar menebak kelas mayoritas. Hal ini sejalan dengan temuan (Hidayat et al., 2025) yang menyatakan bahwa model berbasis ensemble seperti Random Forest secara konsisten mengungguli metode sederhana dalam prediksi luaran alumni.

Untuk memvalidasi kestabilan model lebih lanjut, dilakukan Stratified K-Fold Cross Validation dengan  $k=5$ . Hasil cross validation menunjukkan nilai akurasi yang konsisten antar fold, mengindikasikan bahwa model tidak mengalami overfitting terhadap data latih. Berikut hasilnya:

- CV Accuracy per fold: [0.96808511 0.96774194 0.93548387 0.96774194 0.97849462]

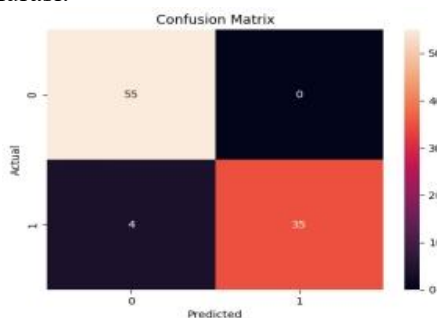
2. Mean CV Accuracy:  
 0.9635094943948754

Selain itu, uji *shuffle baseline* dilakukan dengan mengacak label target secara acak. Hasil akurasi pada data acak mendekati 0,50, yang mengkonfirmasi bahwa performa model pada data asli benar-benar berasal dari pola yang dipelajari dan bukan dari kebetulan statistik

Shuffle accuracy: 0.551681537405628

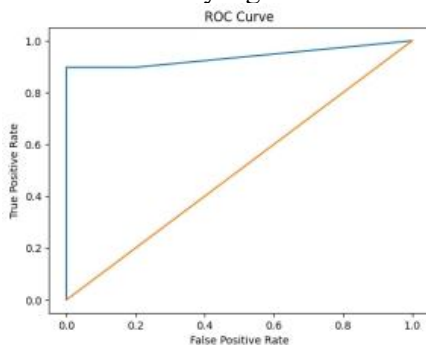
### Visualisasi Evaluasi Model

Hasil evaluasi divisualisasikan melalui dua jenis grafik. Pertama, confusion matrix yang menampilkan distribusi prediksi benar dan salah untuk masing-masing kelas secara lebih rinci dibandingkan metrik tunggal seperti akurasi.



**Gambar 1 Confusion Matrix**

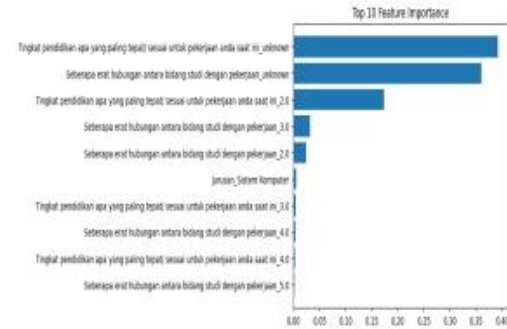
Kedua, kurva ROC yang menggambarkan kemampuan model dalam membedakan kelas positif dan negatif pada berbagai ambang batas keputusan. Nilai AUC-ROC yang mendekati 1,0 menunjukkan kemampuan diskriminasi model yang baik.



**Gambar 2 Kurva ROC**

### Feature Importance dan Interpretasi SHAP

Analisis feature importance bawaan Random Forest menunjukkan 10 fitur teratas yang paling berkontribusi terhadap prediksi model berdasarkan penurunan impuritas rata-rata (mean decrease in impurity).



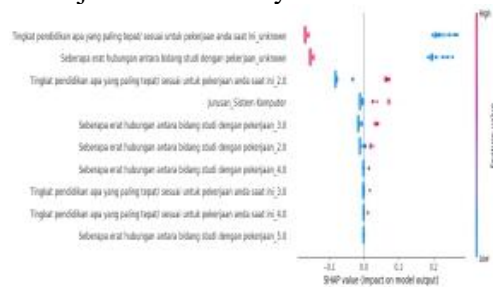
**Gambar 3 Grafik Interpretasi SHAP**

**Tabel 3 Tingkat Hasil SHAP**

feature	importance
Tingkat pendidikan apa yang paling tepat/ sesu...	0.391940
Seberapa erat hubungan antara bidang studi den...	0.360317
Tingkat pendidikan apa yang paling tepat/ sesu...	0.173896
Seberapa erat hubungan antara bidang studi den...	0.031919
Seberapa erat hubungan antara bidang studi den...	0.024330
Jurusan_Sistem Komputer	0.005396
Tingkat pendidikan apa yang paling tepat/ sesu...	0.004282
Seberapa erat hubungan antara bidang studi den...	0.003469
Tingkat pendidikan apa yang paling tepat/ sesu...	0.002350
Seberapa erat hubungan antara bidang studi den...	0.002101

Untuk interpretasi yang lebih mendalam dan dapat dipertanggungjawabkan, analisis dilanjutkan menggunakan SHAP. Berbeda dengan feature importance bawaan yang hanya menunjukkan besaran kontribusi global, SHAP mampu mengungkap arah pengaruh setiap fitur terhadap prediksi, yaitu apakah suatu fitur

mendorong prediksi ke arah "Sudah Bekerja" atau sebaliknya.



**Gambar 4 Sumbu Horizontal**

Pada summary plot, sumbu horizontal menunjukkan nilai SHAP (kontribusi terhadap prediksi kelas "Sudah Bekerja"), sementara warna menunjukkan nilai fitur (merah = tinggi, biru = rendah). Fitur yang berada di bagian atas plot merupakan fitur paling berpengaruh secara global terhadap prediksi model.

Hasil analisis SHAP menunjukkan bahwa fitur paling dominan dalam memprediksi status kebecerjaan alumni adalah "Tingkat pendidikan yang paling tepat untuk pekerjaan saat ini" dan "Seberapa erat hubungan antara bidang studi dengan pekerjaan", khususnya pada responden yang tidak mengisi kedua pertanyaan tersebut (unknown). Alumni dengan nilai unknown pada kedua fitur ini justru cenderung diprediksi sudah bekerja, yang kemungkinan mencerminkan ketidaklengkapan pengisian survei oleh alumni yang sudah aktif bekerja. Selain itu, program studi Sistem Komputer turut muncul sebagai fitur berpengaruh, mengindikasikan adanya perbedaan peluang kebecerjaan antar program studi

### Evaluasi Fairness per Program Studi

Evaluasi fairness dilakukan dengan membandingkan akurasi prediksi model antar program studi sebagai atribut sensitif. (Suci Amaliah et al., 2022) menegaskan bahwa "apabila model dilatih menggunakan data yang mencerminkan bias historis atau sosial, model tersebut dapat mereplikasi bias tersebut dalam prediksinya", sehingga evaluasi per kelompok menjadi langkah kritis dalam audit model. Analisis hanya dilakukan

pada program studi dengan jumlah sampel uji lebih dari 5 untuk memastikan representativitas statistik.

**Tabel 4 Hasil evaluasi fairness**

Index	Jurusan	N
1	Sistem Komputer	18
0	Sistem Informasi	76

Perbedaan nilai akurasi antar program studi mengindikasikan adanya variasi performa model yang perlu diperhatikan. Program studi dengan akurasi tinggi menunjukkan bahwa pola data alumni pada jurusan tersebut lebih mudah dipelajari oleh model, sementara program studi dengan akurasi rendah dapat menjadi indikasi adanya bias atau keterbatasan data pada kelompok tersebut

Hasil evaluasi fairness menunjukkan bahwa model memiliki performa yang tinggi pada kedua program studi. Program studi Sistem Komputer mencapai akurasi sempurna sebesar 1,00 (N=18), sedangkan program studi Sistem Informasi memperoleh akurasi sebesar 0,947 (N=76). Perbedaan akurasi antar kedua kelompok relatif kecil (0,053), mengindikasikan bahwa model tidak menunjukkan bias prediksi yang signifikan berdasarkan atribut program studi. Namun demikian, jumlah sampel Sistem Komputer yang jauh lebih sedikit (18 data) dibandingkan Sistem Informasi (76 data) perlu dicermati, karena akurasi sempurna pada kelompok kecil belum tentu mencerminkan generalisasi yang baik

### SIMPULAN

Penelitian ini menunjukkan bahwa pendekatan educational data mining berbasis Random Forest mampu menghasilkan model prediksi karir alumni yang akurat, transparan, dan berkeadilan. Model yang dibangun mencapai akurasi sebesar 95,7% dengan AUC-ROC 0,938 dan rata-rata akurasi cross validation

sebesar 96,4%, mengkonfirmasi bahwa model berhasil mempelajari pola kekerjaan alumni secara konsisten dan bukan dari kebetulan statistik.

Integrasi metode SHAP sebagai kerangka Explainable AI memberikan nilai tambah yang signifikan dibandingkan pendekatan feature importance konvensional, karena mampu mengungkap tidak hanya besaran kontribusi fitur tetapi juga arah pengaruhnya terhadap prediksi. Temuan ini memperkuat argumen bahwa interpretabilitas model merupakan komponen yang tidak terpisahkan dari sistem pendukung keputusan berbasis machine learning di bidang pendidikan.

Evaluasi fairness yang dilakukan per program studi menunjukkan bahwa model tidak menghasilkan bias prediksi yang signifikan antar kelompok, sehingga hasil prediksi dapat digunakan secara adil sebagai dasar perumusan kebijakan akademik. Secara lebih luas, kerangka analisis terpadu yang menggabungkan Random Forest, SHAP, dan evaluasi fairness ini berpotensi diadaptasi oleh institusi pendidikan tinggi lain sebagai instrumen standar dalam pengelolaan data tracer study untuk menghasilkan kebijakan pengembangan kurikulum dan layanan karir yang lebih tepat sasaran dan berbasis bukti.

## DAFTAR PUSTAKA

- Alfadia Shauqie, S., Nurkamal Fauzan, M., Prianto, C., Sariasih No, J., Sukasari, K., Bandung, K., & Barat, J. (2025). Analisis Pengaruh Fitur Terhadap Tinggi Badan Anak menggunakan SHAP. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*, 11(2), 271–276.
- Aprilia, K. T., Sitorus, I. P., Ridha, M. R., & Novelan, M. S. (2026). *Comparison of Random Forest and Naive Bayes Classifier Methods for Monkeypox Classification*. 3(1), 58–67.
- Barutu, S., Iqbal, M., & Nasution, D. (2025). *Comparative Analysis of the C4 . 5 and Random Forest Algorithms for the Prediction of Diarrheal Disease*. 4(07), 625–635. <https://doi.org/10.58471/esaprom.v4i07>
- Ernawati, A., Sitorus, Z., Iqbal, M., & Nasution, D. (2025). Penerapan Data Mining Untuk Klasifikasi Penduduk Miskin Di Kabupaten Labuhanbatu Menggunakan Random Forest Dan K-Nearest Neighbors. *Bulletin of Information Technology (BIT)*, 6(2), 23–35. <https://doi.org/10.47065/bit.v5i2.1783>
- Fatimah, A., Tania, K. D., Meiriza, A., Studi, P., Informasi, S., Sriwijaya, U., Palembang-prabumulih, J., & Ilir, O. (2025). Analisis Komparatif Model Data Mining Dalam Prediksi Ketepatan. *Journal Of Information Systems And Informatics Engineering*, 9(1), 100–108.
- Helmy, A., Sitorus, Z., Ardy, D., Hrp, A. C., T, S. I. S., & Sukrianto, S. (2025). Analysis of Social Assistance Donor Classification at the Muhammadiyah Medan Orphanage Using SVM. *Sinkron*, 9(1), 283–290. <https://doi.org/10.33395/sinkron.v9i1.14299>
- Hidayat, I., Iqbal, M., Marlina, L., Putera, A., & Siahaan, U. (2025). Analysis of Heart Failure Prediction With. *International Journal of Computer Sciences and Mathematics Engineering*, 4(1), 10–17.
- Parhusip, N., Iqbal, M., & Sitorus, Z. (2025). *Analysis Of Licensing Data Using Naive Bayes And Decision Tree Algorithms To Evaluate Analysis Of Licensing Data Using Naive Bayes And Decision Tree Algorithms To Evaluate The Performance Of Digital Public Services (Case Study: Invesment and One-Stop Integrated Services Office Of Medan City)*. 07(02), 24–35. <https://journal.cattleyadf.org/index.php/jatilima/index>
- Sholeha, S. H., Faid, M., & Yaqin, M. A.

- (2024). Prediksi Prediksi Perpindahan Pelanggan Pada Toko Online Menggunakan Metode Tree-Based Gradient Boosted Models. *Journal of Computer System and Informatics (JoSYC)*, 5(3), 605–614. <https://doi.org/10.47065/josyc.v5i3.5215>
- Sitorus, R. A., Arya, D., Buyung Satrio Dasopang, B. S. D., & Ilka Zufria, I. Z. (2023). Analisis Tracer Study Alumni Program Studi S1 Ilmu Komputer UIN Sumatera Utara. *Jurnal Kridatama Sains Dan Teknologi*, 5(02), 411–420. <https://doi.org/10.53863/kst.v5i02.967>
- Subekti, E. E., Cahyadi, F., & Priyanto, W. (2021). *Inventa : Jurnal Pendidikan Guru Sekolah Dasar STUDI PENELUSURAN ( TRACER STUDY ) PROGRAM STUDI PENDIDIKAN GURU SEKOLAH DASAR TAHUN LULUS 2015-2019. V(2).*
- Suci Amaliah, Nusrang, M., & Aswi, A. (2022). Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijawa Bantaeng. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 4(3), 121–127. <https://doi.org/10.35580/variansiunm31>
- Usman, M., Hidayat, A., Sambas, P. N., Sambas, P. N., Statistik, B. P., Explanations, S. A., & Vector, S. (2024). *Prosiding Seminar Hasil Penelitian dan Pengabdian Kepada Masyarakat ALGORITMA SHAPLEY ADDITIVE EXPLANATIONS UNTUK ANALISIS Prosiding Seminar Hasil Penelitian dan Pengabdian Kepada Masyarakat.* 7(1), 76–81.