

Analisa Performa Algoritma Random Forest & Logistic Regression Dalam Sistem Credit Scoring

Bernadus Billy Riantono^a, Roni Andarsyah^b

^aProgram Studi D4 Teknik Informatika, Universitas Logistik & Bisnis Internasional, 714222072@std.ulbi.ac.id

^bProgram Studi D4 Teknik Informatika, Universitas Logistik & Bisnis Internasional, roniandarsyah@ulbi.ac.id

Submitted: 14-02-2024, Reviewed: 01-03-2024, Accepted 04-04-2024

<https://doi.org/10.47233/jteksis.v6i2.1308>

Abstract

Credit scoring plays a crucial role in the loan approval process within financial institutions. With advancements in technology, there has been a growing trend towards integrating technology into the financial industry, including credit scoring systems. Several studies have attempted to ascertain whether Machine Learning can enhance the credit scoring process, often referencing accuracy metrics as benchmarks. This raises questions about the feasibility of Machine Learning in credit scoring and whether accuracy metrics alone suffice as evaluation criteria. Other classification metrics such as F1 Score and AUC (Area Under the Curve) also warrant consideration, along with exploring the correlation between these metrics. This research aiming on evaluating the performance of Machine Learning models, specifically Random Forest and Logistic Regression, from the perspective of F1 Score and AUC Score and to validate the applicability of Machine Learning in the credit scoring process. The evaluation process is conducted using a 10-fold cross-validation method to ensure robustness. The results indicate that Random Forest outperforms Logistic Regression, demonstrating an F1 Score of 94.68 and an AUC of 0.90. Additionally, a linear correlation between F1 Score and AUC is observed, confirming the feasibility of Machine Learning in credit scoring.

Keywords: Credit Scoring, Logistic Regression, Machine Learning, Random Forest

Abstrak

Credit Scoring memegang peranan penting dalam proses pemberian pinjaman pada Institusi Keuangan, dan dengan perkembangan teknologi yang ada, mulai banyak penerapan teknologi pada industri keuangan tersebut termasuk pada sistem credit scoring, beberapa penelitian coba membuktikan apakah Machine Learning dapat membantu dalam proses Credit Scoring, dan kebanyakan mengacu kepada matrix akurasi sebagai acuan. Hal ini menimbulkan pertanyaan apakah benar Machine Learning dapat diterapkan pada proses Credit Scoring dan apakah hanya matrix akurasi saja yang dapat dijadikan acuan, bagaimana dengan matrix klasifikasi yang lain seperti F1 Score dan AUC dan bagaimana korelasi kedua matrix tersebut. Oleh karena itu penelitian ini bertujuan untuk menguji performa penerapan model Machine Learning khususnya pada model Random Forest dan Logistic Regression dari sudut pandang F1 Score dan AUC Score serta melakukan pembuktian penerapan machine learning pada proses credit scoring dengan metode uji validasi 10 k-fold cross validation. Setelah dilakukan penelitian mendalam didapatkan hasil bahwa Random Forest memberikan performa yang lebih bagus daripada Logistic Regression dengan menunjukkan nilai F1 Score 94,68 dan AUC 0,90 yang berkorelasi secara linear antara F1 Score dan AUC serta terkonfirmasi bahwa Machine Learning dapat diterapkan pada proses credit scoring.

Keywords: Credit Scoring, Logistic Regression, Random Forest, Pembelajaran Mesin

This work is licensed under Creative Commons Attribution License 4.0 CC-BY International license



PENDAHULUAN

Proses peminjaman sejumlah dana oleh institusi keuangan terhadap nasabah adalah salah satu proses tradisional yang telah lama dilakukan selama beberapa dekade terakhir pada institusi keuangan, dimana nasabah dapat melakukan pengajuan untuk mendapatkan pinjaman tersebut. Dalam proses pengajuan tersebut pihak institusi keuangan akan melakukan beberapa tahapan untuk mengetahui profil dari nasabah, seperti mengumpulkan informasi terkait pekerjaan, status pernikahan, kepemilikan rumah, hingga level pendidikan dengan tujuan untuk mengetahui faktor risiko dari seorang

nasabah apakah nasabah beresiko untuk mengalami gagal bayar atau tidak dan tahap inilah yang sering disebut sebagai Credit Scoring dimana pihak institusi keuangan akan memutuskan pemberian pinjaman berdasarkan hasil dari klasifikasi nasabah ini. Credit Scoring sendiri mulai diperkenalkan sejak tahun 1960.[1]

Seiring dengan industri keuangan yang semakin hari semakin berkembang, pemanfaatan Credit Scoring semakin meluas tidak hanya dipergunakan untuk sebatas pinjaman dana besar seperti Kepemilikan Rumah ataupun Kepemilikan Kendaraan saja, sekarang Credit Scoring sudah

dipergunakan untuk pengajuan pinjaman / kredit ritel juga seperti pengajuan kartu kredit.[2]

Credit Scoring sendiri saat ini mulai berkembang tidak hanya dilakukan dengan proses tradisional, tapi sudah mulai melibatkan teknologi didalamnya dan seiring dengan perkembangan dunia teknologi saat ini yang semakin maju, dimana salah satunya adalah Artificial Intelligence (AI), dimana Artificial Intelligence saat ini telah banyak dipergunakan dalam banyak hal. Salah satu contoh Artificial Intelligence yang saat ini sering kita jumpai dan booming di masyarakat adalah ChatGPT.[3]

Artificial Intelligence sendiri memiliki beberapa sub-bidang salah satunya adalah Machine Learning. Machine Learning sendiri menurut penelitian terakhir terbagi menjadi 3 sub bagian: Supervised, Unsupervised dan Reinforced[4] Dimana Machine Learning Supervised Learning sendiri dapat digunakan untuk memecahkan permasalahan klasifikasi dimana terdapat data yang dilengkapi atribut yang akan diprediksi misalnya label kelas, salah satu contoh Algoritma yang biasa digunakan antara lain Naive Bayes, Support Vector Machine, Logistic Regression, Random Forest dan masih banyak lagi, sedangkan untuk Unsupervised Learning dimana data yang masuk tidak memiliki atribut yang akan diprediksi, contoh dari Machine Learning Unsupervised adalah Clustering, ataupun Association.[5]

Penerapan Artificial Intelligence saat ini juga telah tersebar hampir diseluruh bidang / industri antara lain: Kesehatan[6], [7], [8], [9], Keuangan [10], Pendidikan[11], [12], Otomotif[13] dengan salah satu contohnya adalah dengan kendaraan tanpa sopir[14] bahkan hingga memasuki industri kreatif seperti pembuatan 3D Modelling, Video[15], desain interior[16] ataupun gambar.

Penerapan Artificial Intelligence dalam hal ini Machine Learning yang semakin meluas ini dan termasuk didalamnya implementasinya pada industri keuangan, menimbulkan pertanyaan bagi peneliti apakah pemanfaatan teknologi ini pada sistem Credit Scoring dapat membantu institusi keuangan dalam melakukan prediksi dan klasifikasi profil nasabah sehingga dapat mempercepat proses scoring sehingga membantu nasabah untuk mendapatkan hasil assesment pengajuan pinjaman mereka dalam waktu yang lebih cepat apakah pengajuan pinjaman mereka disetujui ataupun ditolak, dan kemudian bagaimana performa dari model machine learning tersebut. Beberapa penelitian sebelumnya telah melakukan penelitian tentang perfoma model machine learning tersebut akan tetapi penelitian tersebut berfokus pada matrix akurasi, dan hal ini menimbulkan pertanyaan untuk

peneliti bagaimana performa model tersebut apabila kita melihat dari matrix F1 Score dan AUC.[17]

Penelitian ini dilakukan untuk menemukan jawaban dari pertanyaan tersebut, oleh karena itu dalam penelitian ini dilakukan dengan melakukan pengembangan model Random Forest dan Logistic Regression dengan menggunakan dataset yang tersedia di Kaggle dimana pada saat dataset tersebut diakses pada 10 Januari 2024 memberikan data mentah sebanyak 4521 data dan bersifat imbalance[18]. Semua model nantinya akan di uji validasi menggunakan 10 K-Fold Cross Validation sebanyak sepuluh kali, dan akan dibuat sebuah sistem aplikasi sederhana untuk mengetahui jawaban dari pertanyaan apakah machine learning dapat membantu institusi keuangan dalam melakukan klasifikasi nasabah pada proses credit scoring.

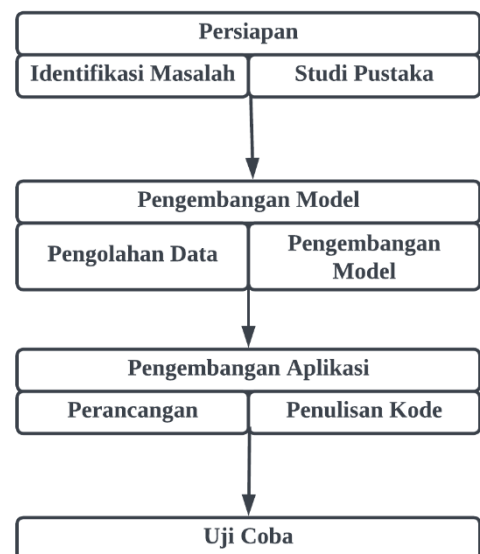
METODE PENELITIAN

Pada penelitian kali ini, metodologi yang akan digunakan adalah *research & development*, dimana peneliti melakukan ujicoba terhadap variabel kontrol input dan menganalisa hasil yang didapatkan.

Penelitian *research & development* sendiri bersifat longitudinal (bertahap dan bisa multi-years). [19]

2.1. Alur Penelitian

Berikut ini adalah diagram dari tahapan – tahapan yang dilakukan dalam penelitian ini:



Gambar 1. Diagram Alur Penelitian

2.2 Persiapan

Pada tahap ini peneliti melakukan proses identifikasi masalah untuk menentukan target eksplorasi dan studi pustaka / literatur, dimana peneliti mempelajari beberapa penelitian sebelumnya yang berkaitan dengan penerapan machine learning pada sistem credit scoring.

2.3 Pengembangan Model

Untuk pengembangan model ini peneliti menggunakan dataset yang disediakan oleh kaggle yang bernama “*credit scoring for borrowers in bank*”.[8] yang saat diakses pada 10 Januari 2024 memberikan data sebanyak 4521.

Pada tahap ini terdapat dua langkah yang dikerjakan oleh peneliti pada penelitian kali ini, yaitu proses pengolahan data dan proses pembuatan / pengembangan model itu sendiri.

Pada proses pengolahan data dilakukan tahapan – tahapan untuk melakukan data *preprocessing* untuk menyiapkan dataset yang ada dengan cara melakukan *data cleansing*, *data normalization* dan *data encoding / transformation*. Setelah itu dilakukan data analysis untuk mengetahui karakteristik dataset yang digunakan dalam penelitian.

Setelah proses *preprocessing* dan *analysis* selesai dilakukan peneliti melanjutkan dengan proses pembuatan model algoritma dimana pada penelitian kali ini model yang digunakan adalah Random Forest dan Logistic Regression.

2.4 Pengembangan Aplikasi

Pada tahap ini dilakukan pengembangan dari aplikasi sederhana yang dijadikan objek penelitian untuk pembuktian apakah benar Machine Learning dapat diterapkan pada proses credit scoring. Dimana peneliti melakukan proses perancangan dan penulisan kode. Dimana pada proses perancangan digunakan untuk menentukan apa saja variabel input yang dibutuhkan didalam aplikasi.

2.5 Uji Coba

Pada tahapan ini dilakukan ujicoba baik uji validasi model yang telah dibangun dan ujicoba aplikasi yang telah dibangun, untuk uji validasi model sendiri pada penelitian kali ini digunakan metode 10 *k-fold cross validation* untuk mengukur performa dari model yang digunakan, dan dari proses ini nantinya akan didapatkan matrix statistik klasifikasi seperti Precision, Recall, F1 Score ataupun AUC.

		Predicted	
		0	1
Actual	0	TP	FP
	1	FN	TN

Gambar 2. Ilustrasi Confusion Matrix

Dimana pada confusion matrix terdapat nilai True Positive, False Positive, False Negative dan True Negative.

True Positive (TP) : Jumlah data yang bernilai Positif dan diprediksi benar sebagai Positif.

False Positive (FP) : Jumlah data yang bernilai Negatif tetapi diprediksi sebagai Positif.

False Negative (FN) : Jumlah data yang bernilai Positif tetapi diprediksi sebagai Negatif.

True Negative (TN) : Jumlah data yang bernilai Negatif dan diprediksi benar sebagai Negatif.

dari confusion matrix juga didapatkan nilai matrix Precision, Recall dan F1 Score.

HASIL DAN PEMBAHASAN

3.1 Data Preparation & Data Analysis

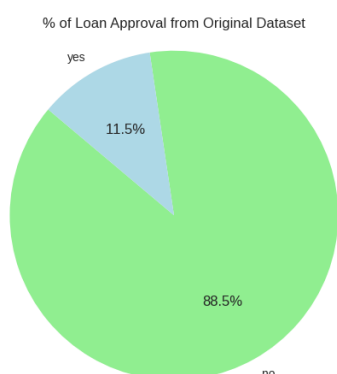
Sebelum memulai tahap lanjut, peneliti melakukan analisis terhadap dataset yang digunakan dan mendapatkan informasi bahwa dataset memuat 4521 data dan memiliki 17 kolom / *features*.

Tabel 1. Penjelasan kolom / features pada dataset

Nama	Penjelasan
age	Umur
job	Pekerjaan
marital	Status Pernikahan
education	Level Pendidikan
default	Catatan Gagal Bayar
balance	Saldo
housing	Status Kepemilikan Rumah
loan	Jumlah Pinjaman Berjalan
contact	Jalur Kontak
day	Jarak Pinjaman Sebelumnya
month	Bulan Pengajuan
duration	Lama Pinjaman

Nama	Penjelasan
campaign	Jumlah Pengajuan Pinjaman
pdays	
previous	
poutcome	
y	Approval

Target label (“y”) memuat nilai antara “yes” atau “no” dimana yes berarti pinjaman disetujui, dan dari hasil analisa didapatkan dari 4521 data yang ada terdapat 4000 data bernilai “no” dan 521 data bernilai “yes”, yang menandakan bahwa dataset ini bersifat *imbalance* atau tidak seimbang.



Gambar 3. Persentase Perbandingan Hasil

Setelah mengetahui perbandingan persentase dari pengajuan yang diterima dan ditolak, baru dilakukan beberapa tahapan *data preprocessing*, seperti: *data cleaning* dimana dataset yang dilakukan penghapusan pada data – data yang memiliki nilai null, kemudian data – data yang bersifat outlier juga dilakukan penghapusan, *data normalization* dimana data – data seperti umur yang memiliki range angka tidak wajar dilakukan scaling untuk memasukan nilai umur tersebut berada pada range / jangkauan yang normal dibandingkan dengan data – data pada kolom lain yang bersifat numerikal dan *data transformation* dimana data dilakukan encoding pada kolom yang bertipe kategorikal.

age	balance	day	duration	campaign	pdays	previous
-0.5625	1.068858	0.250000	-0.479638	-0.5	0.0	0.0
-0.3750	3.417840	-0.416667	0.158371	-0.5	340.0	4.0
-0.2500	0.726917	0.000000	0.000000	-0.5	331.0	1.0
-0.5625	0.825509	-1.083333	0.063348	1.0	0.0	0.0
1.2500	-0.329421	-0.916667	0.185520	-0.5	0.0	0.0

Gambar 4. Kondisi Data Setelah Preprocessing

3.2 Pembagian Dataset

Sebelum melakukan pembentukan model, pada penelitian kali ini membagi dataset menjadi dua bagian yaitu data latih / *train* dan uji / *testing*. Pembagian dataset ini sendiri menggunakan perbandingan 75:25 dimana 75% merupakan data training dan 25% data testing. Untuk melakukan pembagian dataset ini peneliti menggunakan fungsi `train_test_split` yang disediakan oleh *library* Scikit Learn.



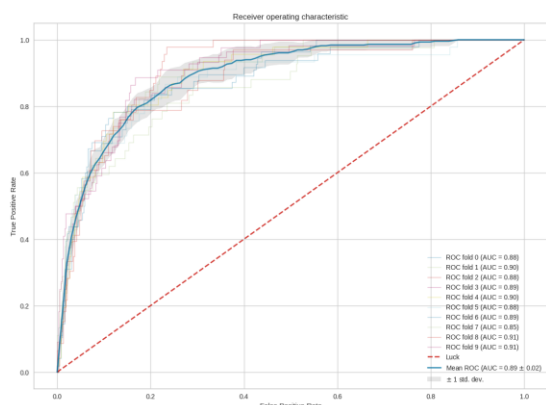
Gambar 5. Kode untuk melakukan pembagian dataset

3.3 Pembuatan Model

Setelah dataset dibagi menjadi dua bagian, maka dilanjutkan dengan proses pembuatan model itu sendiri, dimana pada penelitian kali ini model yang digunakan adalah *Logistic Regression* dan *Random Forest*.

3.4 Evaluasi Model

Setelah mendapatkan model yang akan digunakan tahap selanjutnya adalah melakukan evaluasi / uji validasi terhadap model yang telah terbentuk, dimana untuk tahap pertama dilakukan uji validasi terhadap model *Logistic Regression*, untuk uji validasi pada penelitian kali ini peneliti menggunakan metode *10 k-folds cross validation*.



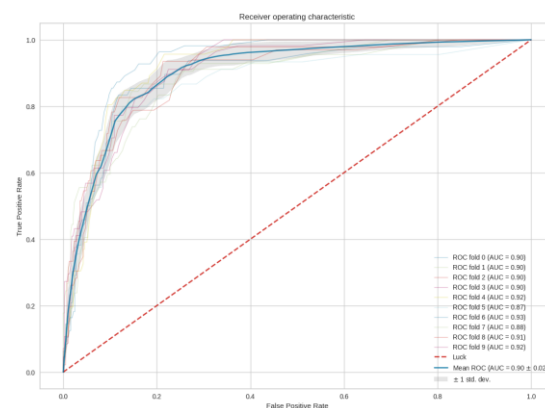
Gambar 6. ROC Curve Logistic Regression

Dari pengujian sebanyak sepuluh kali *Logistic Regression* memberikan hasil bahwa model *Logistic Regression* memiliki area cakupan di bawah kurva (AUC) yang cukup tinggi dan nilai AUC rata rata berada diangka 0.89 jika nilai AUC nya berada di bawah < 0.5 model dapat di katakan buruk jika digunakan[20] sehingga dari hasil yang didapatkan berdasarkan nilai AUC, peneliti dapat menarik kesimpulan bahwa model memiliki kemampuan untuk melakukan prediksi dengan baik. Selain itu didapatkan pula nilai precision dan recall dimana dari nilai precision dan recall ini akan didapatkan nilai F1 Score dimana dari 10 kali percobaan Logistic Regression menghasilkan nilai F1 Score rata – rata berada diangka 94,50.

Tabel 2. Statistik Logistic Regression

Test	AUC	Precision	Recall	F1
1	0.88	98.12	92.19	95.06
2	0.90	96.81	93.81	95.28
3	0.88	96.39	94.92	95.64
4	0.89	98.06	89.95	93.83
5	0.89	97.87	91.75	94.71
6	0.87	98.40	91.58	94.28
7	0.89	98.36	89.11	92.95
8	0.85	97.36	93.18	94.18
9	0.91	97.33	91.94	93.53
10	0.91	98.67	93.45	95.54

Setelah proses uji validasi pada model Logistic Regression telah dilakukan, dilanjutkan dengan melakukan uji validasi pada model Random Forest dengan metode uji validasi pada Logistic Regression yaitu 10 k-folds cross validation sebanyak 10 kali.



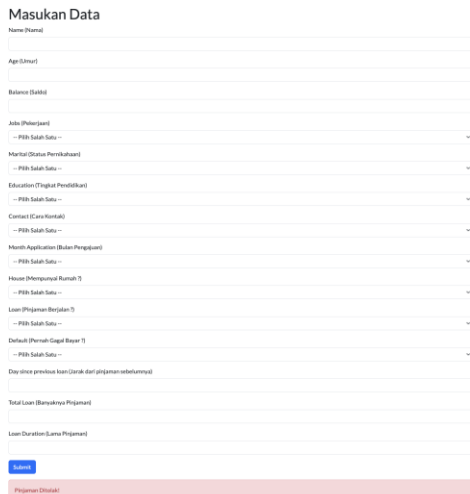
Gambar 7. ROC Curve Random Forest

Dari hasil uji validasi model *Random Forest* memberikan hasil bahwa model *Random Forest* memiliki area cakupan di bawah kurva (AUC) yang cukup tinggi dan nilai AUC rata rata berada diangka 0.90 sehingga dari nilai tersebut, menunjukan bahwa model *Random Forest* memiliki kemampuan untuk melakukan prediksi dengan baik.[10]

Tabel 3. Statistik Random Forest

Test	AUC	Precision	Recall	F1
1	0.89	98.39	90.17	94.10
2	0.89	98.40	92.96	95.60
3	0.89	98.20	94.54	96.33
4	0.90	98.90	88.70	93.52
5	0.92	97.87	91.29	94.46
6	0.87	98.14	91.79	94.85
7	0.93	98.63	88.70	93.40
8	0.88	98.42	91.65	94.91
9	0.90	97.07	91.69	94.30
10	0.91	98.67	92.29	95.37

Selain nilai AUC, didapatkan juga nilai matrix klasifikasi lainnya seperti Precision dan Recall dimana dari nilai Precision dan Recall dapat dikalkulasikan menjadi nilai F1, dimana untuk model Random Forest memberikan nilai rata – rata F1 score berada diangka 94.68.



Gambar 8. Implementasi Simple Dashboard

Setelah uji validasi dilakukan peneliti melakukan pengujian pada aplikasi yang telah dibuat. Dari hasil pengujian aplikasi dapat digunakan untuk melakukan input data secara bebas dan mendapatkan keluaran apakah status pinjaman disetujui ataupun ditolak yang memberikan fakta bahwa teknologi *machine learning* / pembelajaran mesin dapat digunakan dalam proses *credit scoring*.

SIMPULAN

Berdasarkan dari penelitian yang telah dilakukan maka didapatkan kesimpulan bahwa *Machine Learning* dapat digunakan pada proses *credit scoring*, dan dari matrix AUC dan F1 Score yang didapatkan dari pengujian didapatkan kesimpulan bahwa model *Random Forest* memberikan performa yang lebih baik dibandingkan model *Logistic Regression* dengan rata - rata nilai F1 Score di angka 94.68 dan untuk nilai AUC nya berada diangka 0.90.

Pada penelitian selanjutnya dapat ditambahkan bagaimana penanganan model terhadap perubahan data yang berlangsung sehingga tetap dapat mempertahankan performa yang telah didapat, selain itu bisa menambahkan penerapan pada aplikasi yang lebih kompleks.

DAFTAR PUSTAKA

- [1] S. K. Trivedi, "A study on credit scoring modeling with different feature selection and machine learning

- approaches," *Technol Soc*, vol. 63, p. 101413, Nov. 2020, doi: 10.1016/J.TECHSOC.2020.101413.
- [2] D. West, "Neural network credit scoring models," *Comput Oper Res*, vol. 27, no. 11–12, pp. 1131–1152, Sep. 2000, doi: 10.1016/S0305-0548(99)00149-5.
- [3] M. Rizky and R. Andarsyah, "Klasifikasi MIT-BIH Arrhythmia Database Metode Random Forest dan CNN dengan Model ResNet-50: A Systematic Literature Review," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 5, no. 3, pp. 190–196, Jul. 2023, doi: 10.47233/JTEKSIS.V5I3.825.
- [4] M. Somvanshi, P. Chavan, S. Tambade, and S. V. Shinde, "A review of machine learning techniques using decision tree and support vector machine," *International Conference on Computing Communication Control and automation*, Feb. 2016, doi: 10.1109/ICCUBEA.2016.7860040.
- [5] M. Bonzanini, "Mastering social media mining with Python : acquire and analyze data from all corners of the social web with Python," 2016.
- [6] I. M. Agus Oka Gunawan, I. D. A. Indah Saraswati, I. D. G. Riswana Agung, and I. P. Eka Putra, "Klasifikasi Penyakit Jantung Menggunakan Algoritma Decision Tree Series C4.5 Dengan Rapidminer," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 5, no. 2, pp. 73–83, Apr. 2023, doi: 10.47233/JTEKSIS.V5I2.775.
- [7] I. N. Rantaya, S. Bukhori, and J. A. Putra, "Glaucoma Identification on Retinal Fundus Image Using Random Forest Method," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 6, no. 1, pp. 1–7, Apr. 2023, doi: 10.24014/IJAIDM.V6I1.18765.
- [8] F. Nayla, R. Putri, N. Cahyo, H. Wibowo, and H. Mustofa, "Clustering of Tuberculosis and Normal Lungs Based on Image Segmentation Results of Chan-Vese and Canny with K-Means," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 6, no. 1, pp. 18–28, Apr. 2023, doi: 10.24014/IJAIDM.V6I1.21835.
- [9] J. Shedbalkar and K. Prabhushtetty, "Deep transfer learning model for brain tumor segmentation and classification using UNet and chopped VGGNet," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 33, no. 3, pp. 1405–1415, Mar. 2024, doi: 10.11591/IJEECS.V33.I3.PP1405-1415.
- [10] Y. S. Lubis, M. Rizqy Septyandy, M. Debora, and B. Barus, "Optimizing Long Short-Term Memory to Predict Currency Rates," *International Journal of Artificial Intelligence & Robotics (IJAIR)*, vol. 5, no. 2, pp. 71–80, Dec. 2023, doi: 10.25139/IJAIR.V5I2.7318.
- [11] B. W. Aji et al., "Prediction of Indonesia School Enrollment Rate by Using Adaptive Neuro Fuzzy Inference System," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 6, no. 1, pp. 40–46, Apr. 2023, doi: 10.24014/IJAIDM.V6I1.21839.
- [12] M. Munarsih and B. A. Ningsi, "Performance Comparison of Data Mining Classification Algorithms on Student Academic Achievement Prediction," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 6, no. 1, pp. 29–39, Apr. 2023, doi: 10.24014/IJAIDM.V6I1.21874.
- [13] M. Firdaus Ishak et al., "Improving night driving behavior recognition with ResNet50," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 33, no. 3, pp. 1974–1988, Mar. 2024, doi: 10.11591/IJEECS.V33.I3.PP1974-1988.
- [14] M. N. F. Salvaturi, A. Virgono, and R. E. Saputra, "Rancang Bangun Self-Driving Car Robot Berbasis Pengenalan Rambu Lalu Lintas Di Indonesia Menggunakan Convolutional Neural Network," *eProceedings of Engineering*, vol. 8, no. 6, Dec. 2021.

- [15] A. Jayanthiladevi, A. G. Raj, R. Narmadha, S. Chandran, S. Shaju, and K. Krishna Prasad, "AI in Video Analysis, Production and Streaming Delivery," *J Phys Conf Ser*, vol. 1712, no. 1, Dec. 2020, doi: 10.1088/1742-6596/1712/1/012014.
- [16] T. Mahendarto, "From Artificial Intelligence to Artificial Consciousness: An Interior Design Implication," *Journal of Artificial Intelligence in Architecture*, vol. 2, no. 1, pp. 41–52, Feb. 2023, doi: 10.24002/JARINA.V2I1.6627.
- [17] W. Gao and Z. H. Zhou, "On the Consistency of AUC Pairwise Optimization," *IJCAI International Joint Conference on Artificial Intelligence*, vol. 2015-January, pp. 939–945, Aug. 2012.
- [18] A. Kapturov, "Credit scoring for borrowers in bank." Accessed: Feb. 22, 2024. [Online]. Available: <https://www.kaggle.com/datasets/kapturovalexander/bank-credit-scoring/data>
- [19] Sugiyono. Dr, 2013 , "Metode Penelitian Kuantitatif Dan Kualitatif Dan R&D," *Bandung. Penerbit: CV Alfa Beta*.
- [20] K. H. Zou, A. J. O'Malley, and L. Mauri, "Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models," *Circulation*, vol. 115, no. 5, pp. 654–657, Feb. 2007, doi: 10.1161/CIRCULATIONAHA.105.594929.