

REVIEW PERBANDINGAN PERFORMANSI METODE DECISION TREE, RANDOM FOREST, LOGISTIC REGRESSION, DAN XGBOOST PADA DATASET PENIPUAN KARTU KREDIT

LAPORAN PROGRAM INTERNSHIP II

Laporan ini dibuat untuk memenuhi persyaratan kelulusan

Matakuliah Program Internship II

TYPE : TOPIK BARU

SKEMA : INDUSTRI



Oleh:

Dinda Anik Masruro

1184003

PROGRAM DIPLOMA IV TEKNIK INFORMATIKA

POLITEKNIK POS INDONESIA

TAHUN 2022

***PERFORMANCE COMPARISON REVIEW OF DECISION
TREE, RANDOM FOREST, LOGISTIC REGRESSION, AND
XGBOOST METHODS ON CREDIT CARD FRAUD DATASETS***

REPORT OF THE INTERNSHIP II PROGRAM

This report is made to meet the graduation requirements

For the internship II program courses

TYPE : NEW TOPIC

SCHEMA : INDUSTRY



By:

Dinda Anik Masruro

1184003

PROGRAM DIPLOMA IV TEKNIK INFORMATIKA

POLITEKNIK POS INDONESIA

TAHUN 2022

ABSTRAK

Peningkatan penggunaan kartu kredit terjadi secara signifikan seiring dengan perkembangan teknologi yang semakin canggih. Namun peningkatan penggunaan tersebut tak terlepas dari kenaikan terjadinya penipuan pada transaksi kartu kredit (Zamachsari & Puspitasari, 2021). Hal ini tentunya sangat merugikan perusahaan pengelola keuangan setiap tahun. Dengan perkembangan teknologi yang ada maka kasus penipuan transaksi pada kartu kredit dapat diminimalisir dengan cara mendeteksi penipuan pada penggunaan kartu kredit. Pada penelitian ini penulis telah melakukan analisa pada penggunaan kartu kredit untuk meminimalisir adanya penipuan pada kartu kredit. Selain itu juga pada penelitian ini digunakan beberapa metode yang digunakan untuk mengetahui hasil akurasi yang baik, sehingga dapat diketahui metode mana yang paling baik digunakan untuk mendeteksi penipuan pada kartu kredit (Kalid et al., 2020). Dengan menggunakan empat metode penelitian yaitu *Decision tree*, *Random forest*, *Logistic regression*, dan *Xgboost*. Dari hasil *Dataset Credit Card* dengan menggunakan algoritma *Decision tree* hasil *accuracy* sebesar 99%, *sensitivity* sebesar 0%, *specificity* sebesar 100%, dan *AUC* sebesar 83%, dengan menggunakan algoritma *Random Forest* hasil *accuracy* sebesar 99%, *sensitivity* sebesar 0%, *specificity* sebesar 100%, dan *AUC* sebesar 99%, dengan menggunakan *Logistic Regression* hasil *accuracy* sebesar 99%, *sensitivity* sebesar 56%, *specificity* sebesar 99%, dan *AUC* sebesar 97%, sedangkan menggunakan *Xgboost* hasil *accuracy* sebesar 99%, *sensitivity* sebesar 67%, *specificity* sebesar 99%, dan *AUC* sebesar 99%. Sedangkan dari hasil *Dataset Credit Card Taiwan* dengan menggunakan algoritma *Decision tree* hasil *accuracy* sebesar 83%, *sensitivity* sebesar 39%, *specificity* sebesar 96%, dan *AUC* sebesar 76%, dengan menggunakan algoritma *Random Forest* hasil *accuracy* sebesar 82%, *sensitivity* sebesar 40%, *specificity* sebesar 95%, dan *AUC* sebesar 77%, dengan menggunakan *Logistic Regression* hasil *accuracy* sebesar 77%, *sensitivity* sebesar 0%, *specificity* sebesar 100%, dan *AUC* sebesar 64%, sedangkan dengan menggunakan *Xgboost* hasil *accuracy* sebesar 81%, *sensitivity* sebesar 40%, *specificity* sebesar 93%, dan *AUC* sebesar 73%. Maka dari hasil penelitian ini dapat disimpulkan bahwa metode *Xgboost* pada *dataset Credit Card* memiliki nilai akurasi yang tinggi. Sedangkan pada *dataset Credit Card Taiwan* metode *Decision Tree* memiliki akurasi yang lebih tinggi. Sehingga kedua metode tersebut dapat direkomendasikan digunakan untuk mendeteksi penipuan pada kartu kredit.

Kata Kunci: *Machine Learning*, *Decision Tree*, *Random Forest*, *Xgboost*, *Logistic regression*, Kartu Kredit

ABSTRACT

The use of credit cards has increased significantly in line with the development of increasingly sophisticated technology. However, the increase in use is inseparable from the increase in the occurrence of fraud in credit card transactions (Zamachsari & Puspitasari, 2021). This is certainly very detrimental to financial management companies every year. With the development of existing technology, cases of fraudulent transactions on credit cards can be minimized by detecting fraud in the use of credit cards. In this study the author has analyzed the use of credit cards to minimize credit card fraud. In addition, in this study several methods were used to determine good accuracy results, so that it can be seen which method is best used to detect fraud on credit cards (Kalid et al., 2020). By using four research methods, namely Decision tree, Random forest, Logistic regression, and Xgboost. From the results of the Credit Card Dataset using the Decision tree algorithm, the accuracy is 99%, sensitivity is 0%, specificity is 100%, and AUC is 83%, using the Random Forest algorithm, the accuracy is 99%, sensitivity is 0%, specificity is 100%, and AUC of 99%, by using Logistic Regression the results of accuracy are 99%, sensitivity is 56%, specificity is 99%, and AUC is 97%, while using Xgboost the results are 99% accuracy, sensitivity 67%, specificity of 99%, and AUC of 99%. Meanwhile, from the results of the Taiwan Credit Card Dataset using the Decision Tree algorithm, the accuracy is 83%, sensitivity is 39%, specificity is 96%, and AUC is 76%, using the Random Forest algorithm, the accuracy is 82%, sensitivity is 40%, specificity is 95%, and AUC is 77%, using Logistic Regression the accuracy is 77%, sensitivity is 0%, specificity is 100%, and AUC is 64%, while using Xgboost the accuracy is 81%, sensitivity is 40 %, specificity is 93%, and AUC is 73%. So from the results of this study it can be concluded that the Xgboost method on the Credit Card dataset has a high accuracy value. Meanwhile, in the Taiwan Credit Card dataset, the Decision Tree method has higher accuracy. So that both methods can be recommended to be used to detect fraud on credit cards.

Keywords: Machine Learning, Decision Tree, Random Forest, Xgboost, Logistics Regression, Credit Card

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Allah Swt. atas ridanya saya dapat menyelesaikan laporan Internship II ini. Adapun judul Internship II saya adalah “Review Perbandingan Performansi Metode Decision Tree, Random Forest, Logistic Regression, dan Xgboost pada Dataset Penipuan Kartu Kredit”.

Laporan ini diajukan untuk memenuhi syarat kelulusan mata kuliah Internship II di Diploma IV Politeknik Pos Indonesia. Tidak dapat disangkal bahwa butuh usaha yang keras dalam penyelesaian pengerjaan laporan ini. Namun, karya ini tidak akan selesai tanpa orang-orang tercinta di sekeliling saya yang mendukung dan membantu. Terima kasih saya sampaikan kepada:

1. Kedua orang tua tercinta yang senantiasa memberikan dukungan dan doanya.
2. Dr. Ir. Agus Purnomo, M.T. selaku Direktur Politeknik Pos Indonesia.
3. M.Yusril Helmi Setyawan, S. Kom., M. Kom. selaku Ketua Program Studi DIV Teknik Informatika.
4. Syafrial Fachri Pane, S.T., M.T.I., EBDP. CDSP, SFPC selaku Koordinator Internship II
5. Rolly Maulana Awangga, S.T., M.T., CAIP., SFPC selaku Pembimbing Internship II yang telah memberikan pengarahan dan membimbing kepada penulis.
6. Rolly Maulana Awangga, S.T., M.T., CAIP., SFPC selaku dosen wali kelas D4 TI 4 C
7. Serta semua pihak yang tidak dapat penulis sebutkan satu persatu.

Akhir kata penulis berharap semoga laporan ini dapat bermanfaat bagi seluruh mahasiswa khususnya bagi penulis. Penulis sadar bahwa proposal ini masih jauh dari kata sempurna. Oleh karena itu penulis sangat mengharapkan adanya kritik dan saran yang membangun dari para pembaca untuk kesempurnaan proposal ini. Semoga Allah SWT selalu memberikan rahmat dan hidayah-Nya kepada kita semua. Aamiin.

Bandung, 7 Juli 2022

Penulis

DAFTAR ISI

ABSTRAK.....	iii
ABSTRACT	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL.....	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah	2
1.3 Tujuan dan Manfaat.....	2
1.4 Ruang Lingkup	2
1.5 Sistematika Penulisan	2
BAB II TINJAUAN PUSTAKA.....	4
2.1 State of The Art (SOTA)	4
2.2 Landasan Teori	5
2.2.1 Kartu Kredit.....	5
2.2.2 Fraud.....	5
2.2.3 Machine Learning	5
2.2.4 Logistic Regression	6
2.2.5 Decision Tree	6
2.2.6 Random Forest	7
2.2.7 Xgboost	7
2.2.8 Accuracy, Sensitivity, Specificity, AUC.....	7
2.2.9 Python.....	7
2.3 Penelitian Terkait.....	8
BAB III ANALISA PERUSAHAAN	11
3.1 Sejarah Dinas ESDM Provinsi Jawa Barat.....	11
3.2 Visi dan Misi	12
3.3 Strategi Dinas ESDM Provinsi Jawa Barat	12
3.4 Struktur Organisasi Perusahaan.....	13
3.5 Deskripsi dan Ruang Lingkup Internship.....	13

BAB IV METODE PENELITIAN.....	14
4.1 Diagram Alur Metodologi Penelitian	14
4.2 Indikator Capaian Penelitian	15
4.3 Alur Metodologi Random Forest.....	16
4.4 Alur Metodologi Logistic Regression	17
4.5 Alur Metodologi Decision Tree.....	18
4.6 Alur Metodologi Xgboost.....	19
BAB V HASIL PENELITIAN	20
5.1 Hasil Penelitian.....	20
5.1.1 Dataset.....	20
5.1.2 Hasil Eksperimen Model.....	25
BAB VI PENUTUP	27
6.1 Kesimpulan.....	27
DAFTAR PUSTAKA	28
LAMPIRAN.....	32

DAFTAR GAMBAR

Gambar 3.4. 1 Struktur Organisasi Dinas ESDM Provinsi Jawa Barat	13
Gambar 4. 1 Alur Metodologi Penelitian	14
Gambar 5. 1 Perbandingan Spesifikasi dataset Credit Card Eropa	20
Gambar 5. 2 Perbandingan Spesifikasi dataset Credit Card Taiwan.....	20
Gambar 5. 3 Histogram Dataset Credit Card Eropa	22
Gambar 5. 4 Histogram Dataset Credit Card Taiwan	23
Gambar 5. 5 Diagram Dataset Credit Card Eropa.....	24
Gambar 5. 6 Diagram Dataset Credit Card Taiwan	24

DAFTAR TABEL

Tabel 2.3. 1 Penelitian Terkait	8
Tabel 4.2. 1 Indikator Capaian	15
Tabel 4.3. 1 Alur Metodologi Random Forest	16
Tabel 4.4. 1 Alur Metodologi Logistic Regression	17
Tabel 4.5. 1 Alur Metodologi Decision Tree	18
Tabel 4.6. 1 Alur Metodologi Xgboost	19
Tabel 5. 1 Dataset Credit Card Eropa	21
Tabel 5. 2 Dataset Credit Card Taiwan	23
Tabel 5. 3 Jumlah Data.....	24
Tabel 5. 4 Hasil Eksperimen Dengan Algoritma Decission Tree	25
Tabel 5. 5 Hasil Eksperimen Dengan Algoritma Random Forest.....	25
Tabel 5. 6 Hasil Eksperimen Dengan Algoritma Logistic Regression	25
Tabel 5. 7 Hasil Eksperimen Dengan Algoritma Xgboost.....	26

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kasus penipuan merupakan tindakan kriminal yang menyebabkan kerugian, salah satu kasus penipuan yang terjadi adalah penipuan pada kartu kredit. Kasus penipuan ini sangat merugikan baik individu maupun lembaga keuangan atau bank. Penipuan kartu kredit biasanya terjadi ketika pemilik kartu kredit tidak mengetahui bahwa kartu kredit tersebut digunakan oleh orang lain atau bisa disebut dengan penipu. Penipu dapat bebas melakukan transaksi atau memanipulasi data dengan menggunakan kartu kredit sehingga menimbulkan kerugian. Penipuan ini diklasifikasikan ke dalam beberapa jenis seperti penipuan secara langsung, penipuan menggunakan aplikasi, penipuan internal, dan penipuan palsu.

Dengan perkembangan zaman yang semakin canggih penipuan tersebut dapat dideteksi menggunakan *machine learning*. *Machine learning* sendiri memainkan peran penting di berbagai area yang efisien untuk pemrosesan data, salah satunya adalah identifikasi penipuan kartu (Trivedi et al., 2020). Dari beberapa kasus penipuan kartu kredit yang tergolong sulit untuk dideteksi karena memakan waktu serta metode yang canggih adalah penipuan palsu. Penipuan palsu merupakan kasus penipuan yang dilakukan dengan cara jarak jauh dengan penipu yang sudah mengetahui data pribadi, nomor kartu kredit serta pin dari pengguna. Dalam dekade terakhir, telah terjadi peningkatan e-commerce, yang telah meningkatkan penggunaan kartu kredit secara signifikan. Peningkatan penggunaan kartu kredit telah membawa peningkatan konstan dalam transaksi penipuan (Randhawa et al., 2018). Penipuan transaksi kartu kredit telah sangat berdampak pada industri keuangan. Sebuah laporan baru-baru ini menunjukkan bahwa sekitar 27,85 miliar dolar hilang karena penipuan kartu kredit pada 2018, meningkat 16,2% dibandingkan dengan 23,97 miliar dolar yang hilang pada 2017, dan diperkirakan mencapai 35 miliar dolar pada 2023 (Tingfei et al., 2020). Kerugian ini dapat dikurangi melalui pemantauan dan pencegahan penipuan yang efisien. Sementara itu, *machine learning* (ML) telah diterapkan untuk mengembangkan beberapa sistem deteksi penipuan kartu kredit (Kalid et al., 2020). Namun, deteksi penipuan kartu kredit tetap menjadi tantangan dari perspektif pembelajaran karena ketidakseimbangan. Oleh karena itu pada penelitian ini, peneliti melakukan deteksi terhadap penipuan palsu pada kartu kredit. Algoritma *machine learning* yang digunakan pada penelitian ini adalah *Decision tree*, *Random forest*, *Logistic regression*, dan *Xgboost*.

Algoritma *decision tree* adalah teknik berbasis pohon di mana setiap jalur mulai dari root dijelaskan dengan cara memisahkan data urutan sampai hasil boolean pada simpul keputusan tercapai (Charbuty & Abdulazeez, 2021), sedangkan *random forest* merupakan salah satu metodologi ensemble yang digunakan hanya untuk meningkatkan keselarasan dalam algoritma *machine learning* pada kecerdasan buatan. Salah satu pengklasifikasi *random forest* yang disarankan oleh Breiman, seorang peneliti. Metode *random forest* juga dapat membantu mengidentifikasi variabel independen yang benar-benar cocok sehingga sistem dapat memilih fitur. Selain itu, banyak dari temuan menunjukkan pentingnya dalam memilih beberapa kemungkinan untuk setiap semak, tetapi dalam studi empiris temuan ini juga yang terbaik dalam hal prediksi akurasi (Trivedi et al., 2020), kemudian *logistic regression* adalah metode lain yang memutuskan untuk meninjau data statistik dengan *machine learning*. Ini juga merupakan proses masuk ke isu-isu mengenai kategorisasi biner (kesulitan dengan lebih dari dua nilai moral kelas). Regresi logistik ke mean digunakan untuk memodelkan hasil kelas

seperti lulus / gagal total, positif dan konstruktif / negatif atau netral lagi dan dalam kasus deteksi ancaman penipuan kartu kredit maka kami menggunakan kelas distribusi probabilitas sebagai penipuan dan bukan penipuan (Trivedi et al., 2020), dan *xgboost* merupakan algoritma yang ditingkatkan berdasarkan gradient boosting *decision tree* dan dapat membangun *boosted trees* secara efisien dan beroperasi secara parallel (Muslim & Karo, 2020). *Machine learning* memberi komputer kemampuan untuk belajar tanpa diprogram secara eksplisit. Proses *machine learning* mirip dengan proses *data mining* (Dairu & Shilong, 2021). Algoritma *machine learning* digunakan untuk menganalisis semua transaksi dan melaporkan hal yang mencurigakan. Laporan-laporan ini kemudian diselidiki oleh para profesional yang menghubungi pemegang kartu untuk mengonfirmasi apakah transaksi itu asli atau curang. Para penyelidik memberikan umpan balik ke sistem otomatis yang digunakan untuk melatih dan memperbarui algoritma hingga akhirnya meningkatkan kinerja deteksi penipuan dari waktu ke waktu. Tujuan dari penelitian ini adalah untuk mengetahui hasil akurasi dari beberapa algoritma yang digunakan sehingga dapat diketahui hasil akurasi yang akurat dari algoritma yang digunakan untuk mendeteksi penipuan pada kartu kredit.

1.2 Identifikasi Masalah

Adapun identifikasi masalah pada penelitian ini sebagai berikut,

1. Bagaimana cara mendeteksi penipuan pada sebuah transaksi dengan menggunakan algoritma *Decision tree*, *Random forest*, *Logistic regression*, dan *Xgboost*?
2. Bagaimana hasil dari model klasifikasi algoritma *Decision tree*, *Random forest*, *Logistic regression*, dan *Xgboost* untuk mendeteksi penipuan?

1.3 Tujuan dan Manfaat

Adapun tujuan pada penelitian ini sebagai berikut,

1. Mengetahui bagaimana mendeteksi penipuan pada sebuah transaksi dengan menerapkan algoritma *Decision tree*, *Random forest*, *Logistic regression*, dan *Xgboost*.
2. Menganalisis model klasifikasi algoritma *Decision tree*, *Random forest*, *Logistic regression*, dan *Xgboost* yang menghasilkan akurasi optimal dalam mendeteksi penipuan.

Adapun manfaat pada penelitian ini sebagai berikut,

1. Memberikan informasi jumlah penipuan yang terjadi pada kegiatan transaksi yang dilakukan.
2. Mengetahui hasil paling tepat dari hasil akurasi optimal dalam mendeteksi penipuan.

1.4 Ruang Lingkup

Adapun ruang lingkup pada penelitian ini sebagai berikut,

1. Pada penelitian ini data yang digunakan didapatkan dari dataset credit card.
2. Periode waktu 3 bulan.
3. Bahasa pemrograman menggunakan *Python*.
4. Software yang digunakan yaitu *Jupyter* atau *Google Colab*.

1.5 Sistematika Penulisan

Berdasarkan latar belakang dan perumusan masalah diatas, maka penyusunan laporan ini dibuat dalam suatu sistematika yang terdiri dalam enam BAB, yaitu:

BAB I PENDAHULUAN

Bab ini berisi penjelasan terkait dengan State of The Art (SOTA) yang menjelaskan mengenai pemaparan teori umum dengan topik yang dibahas secara global dan mengaitkan dengan referensi yang ada. Identifikasi masalah menjelaskan mengenai masalah dalam judul penelitian dengan pendekatan *machine learning* dan memberikan solusi atas masalah tersebut. Tujuan menjelaskan tentang solusi dari masalah yang ada. Ruang lingkup menjelaskan mengenai batasan dalam pemodelan dan aplikasi tersebut. Serta sistematika penulisan menjelaskan tentang isi laporan internship II.

BAB II LANDASAN TEORI

Bab ini berisi penjelasan mengenai konsep dasar dan pendukung dari sistem yang akan dibangun dengan menggunakan metode tertentu, antara lain State of The Art (SOTA), diagram alur metodologi penelitian, dan penelitian sebelumnya yang berhubungan dengan tema yang diambil.

BAB III ANALISA PERUSAHAAN

Bab ini berisi tentang sejarah perusahaan, visi dan misi perusahaan, strategi perusahaan, struktur organisasi perusahaan dan job description, serta dekripsi dan ruang lingkup perusahaan.

BAB IV METODOLOGI PENELITIAN

Bab ini berisi penjelasan diagram alur metodologi penelitian beserta tahapan-tahapan diagram alur penelitian untuk menyelesaikan penelitian yang sedang dilakukan sehingga bisa mencapai tujuan yang diharapkan.

BAB V HASIL PENELITIAN

Bab ini akan menjelaskan mengenai tahapan proses pengujian yang dilakukan dan analisis hasil pengujian untuk mendapatkan kesimpulan dari setiap proses pengujian.

BAB VI KESIMPULAN

Pada bab ini akan dijelaskan mengenai kesimpulan dari seluruh kegiatan yang telah dilakukan.

BAB II

TINJAUAN PUSTAKA

2.1 State of The Art (SOTA)

Pesatnya perkembangan teknologi informasi yang disertai dengan peningkatan aktivitas masyarakat dalam melakukan transaksi secara elektronik, disatu sisi proses ini akan memberikan kenyamanan bagi masyarakat akan tetapi disertai dengan terjadinya penipuan transaksi keuangan. Penipuan yang tersembunyi dan beragam jenis transaksi konsumen menyebabkan sulitnya mendeteksi transaksi yang akan merugikan konsumen dan perbankan (Zamachsari & Puspitasari, 2021). Dengan perkembangan zaman yang semakin canggih penipuan tersebut dapat dideteksi menggunakan *machine learning*. *Machine learning* sendiri memainkan peran penting di berbagai area yang efisien untuk pemrosesan data, salah satunya adalah identifikasi penipuan kartu (Trivedi et al., 2020). Dari beberapa kasus penipuan kartu kredit yang tergolong sulit untuk dideteksi karena memakan waktu serta metode yang canggih adalah penipuan palsu. Penipuan palsu merupakan kasus penipuan yang dilakukan dengan cara jarak jauh dengan penipu yang sudah mengetahui data pribadi, nomor kartu kredit serta pin dari pengguna. Dalam dekade terakhir, telah terjadi peningkatan e-commerce, yang telah meningkatkan penggunaan kartu kredit secara signifikan. Peningkatan penggunaan kartu kredit telah membawa peningkatan konstan dalam transaksi penipuan (Randhawa et al., 2018). Penipuan transaksi kartu kredit telah sangat berdampak pada industri keuangan. Sebuah laporan baru-baru ini menunjukkan bahwa sekitar 27,85 miliar dolar hilang karena penipuan kartu kredit pada 2018, meningkat 16,2% dibandingkan dengan 23,97 miliar dolar yang hilang pada 2017, dan diperkirakan mencapai 35 miliar dolar pada 2023 (Tingfei et al., 2020). Kerugian ini dapat dikurangi melalui pemantauan dan pencegahan penipuan yang efisien. Sementara itu, *machine learning* (ML) telah diterapkan untuk mengembangkan beberapa sistem deteksi penipuan kartu kredit (Kalid et al., 2020).

Pada penelitian lainnya menggunakan teknik oversampling untuk mengatasi *imbalance* data. *Oversampling* diterapkan pada dataset hipertensi karena kelas hipertensi memiliki jumlah record yang sedikit bila dibandingkan dengan jumlah record untuk kelas tidak hipertensi. Evaluasi penggunaan oversampling dilakukan dengan mengimputasi *missing value*, melakukan *oversampling*, dan mentransformasi data ke dalam *range* yang sama (Neural, 2020), kemudian menerapkan algoritma klasifikasi. Hasilnya dengan menerapkan *oversampling*, *Artificial Neural Network* (ANN) paling baik dengan akurasi sebesar 0,91, disusul *Decision Tree* sebesar 0,86, dan *Naïve Bayes* sebesar 0,71 (Yuanika et al., 2021). Sedangkan pada penelitian ini menerapkan algoritma *Random Forest* untuk melakukan klasifikasi guna memperoleh model prediksi penerimaan pengajuan pinjaman dengan akurasi yang optimal (Sistem, 2021). Pengujian yang telah dilakukan menghasilkan performa dari klasifikasi *bank marketing dataset* menggunakan algoritma *Random Forest* dengan akurasi sebesar 88,30%, AUC (+) sebesar 0,500 dan AUC (-) sebesar 0,000. Penggunaan optimasi *Bagging* dan algoritma Genetika belum mampu meningkatkan performa dari algoritma *Random Forest* untuk klasifikasi *bank marketing dataset*.

Kemudian penelitian lain mengusulkan algoritma *k-means-SMOTE* dibandingkan dengan beberapa algoritma klasifikasi, di antaranya adalah algoritma C4.5, *Support Vector Machines* (SVM), dan *Naïve Bayes*. Metode algoritma klasifikasi SVM memiliki akurasi 82% dan sensitivitas 77% sedangkan metode klasifikasi *Naïve Bayes* menghasilkan nilai spesifikasi terbaik yaitu sebesar 89% (Hairani & Bumigora, 2020). Namun, deteksi penipuan kartu kredit

tetap menjadi tantangan dari perspektif pembelajaran karena ketidakseimbangan. Oleh karena itu pada penelitian ini, peneliti melakukan deteksi terhadap penipuan palsu pada kartu kredit. Algoritma *machine learning* yang digunakan pada penelitian ini adalah *Decision tree*, *Random forest*, *Logistic regression*, dan *Xgboost*. Penggunaan empat model ini diharapkan dapat menentukan hasil akurasi yang optimal.

2.2 Landasan Teori

2.2.1 Kartu Kredit

Kartu kredit umumnya mengacu pada kartu yang diberikan kepada pelanggan (*cardholders*), biasanya memungkinkan mereka untuk melakukannya membeli barang dan jasa dalam batas kredit atau menarik uang tunai di muka. Kartu kredit memberikan *cardholders* sebuah memanfaatkan waktu, yaitu, memberikan waktu bagi pelanggan mereka untuk membayar nanti dalam waktu yang ditentukan, dengan membawanya ke siklus penagihan berikutnya (Nath, 2020).

2.2.2 Fraud

Laporan keuangan juga menjadi alat referensi oleh pelaku bisnis, namun tak jarang beberapa orang memanipulasi laporan keuangan dengan cara mengabaikan angka yang salah atau merubah angka dengan sengaja untuk mencapai tujuan tertentu (Melo et al., 2019).

Berdasarkan pernyataan Auditing Standarts (SAS) no.99 mendefinisikasikan fraud adalah tindak kecurangan yang disengaja dan menyebabkan mistatement pada laporan keuangan (Perwara & Bachtiar, 2020).

2.2.3 Machine Learning

Artificial Intelligence (AI) atau Kecerdasan buatan dewasa ini berkembang begitu pesat. Semua hal berkaitan dengan kecerdasan buatan (Ramdhani, 2019) yang terus dikembangkan sesuai dengan perkembangan zaman.

Dalam Kecerdasan buatan memiliki banyak cabang salah satunya *Machine Learning* (ML), Sebenarnya cabang dari kecerdasan buatan itu ada tujuh yaitu *Machine Learning* itu sendiri, *natural language processing*, *expert system*, *vision*, *speech*, *planning* dan *robotics*. Dalam kecerdasan buatan di berikan cabang cabang tersebut sebenarnya hanya untuk mempersempit lingkup dalam pengembangan maupun dalam pembelajaran AI (Brunton et al., 2020) hal ini didasarkan karena kecerdasan buatan memiliki lingkup yang luas.

Dikembangkannya *Machine Learning* didasarkan disiplin ilmu lainnya seperti statistika, matematika dan data mining. Dimana mesin dikembangkan untuk dapat belajar secara mandiri tanpa perlu diarahkan oleh penggunaanya dan mesin dapat belajar dengan menganalisa data tanpa perlu memprogram ulang atau memberi perintah ulang. Dasar-dasar *Machine Learning* dan konsep *Machine Learning* pertama ditemukan oleh beberapa ilmuwan matematika yaitu Adrien Marie Legendre, Thomas Bayes dan Andrey Markov pada tahun 1920-an. Dan mulai dari situlah *Machine Learning* mulai dikembangkan. Penerapan *Machine Learning* yang terkenal pada tahun 1996 adalah *Deep Blue* yang dibuat oleh IBM. *Machine Learning* ini di kembangkan untuk bisa belajar serta bermain catur. Pada masa itu *Deep Blue* diuji coba untuk bermain catur dengan juara catur profesional dengan pemenangnya adalah *Machine Learning deep blue*. Manfaat *Machine Learning* pada kehidupan kita sehari hari dapat kita rasakan bersama (Roscher et al., 2020). Seperti pada kunci layar hp yang menggunakan *face unlock* untuk membuka kunci layar Smartphone. Atau bahkan iklan iklan yang di sesuaikan dengan trend saat ini. itu merupakan hasil dari pengolahan data menggunakan *Machine Learning*.

2.2.4 Logistic Regression

Regresi linier adalah model yang mengestimasi koefisien regresi yang secara linier dapat menyatakan hubungan antara variabel bebas X dan variabel terikat Y dengan nilai kontinu. Jika variabel dependen Y adalah variabel biner (Yun, 2021), digunakan regresi logistik karena regresi linier tidak dapat diterapkan secara langsung. Beberapa algoritma regresi dapat digunakan untuk klasifikasi, dan regresi logistik secara luas digunakan untuk memperkirakan probabilitas bahwa sampel termasuk dalam kelas tertentu (Thabtah et al., 2019). Jika probabilitas yang diestimasi melebihi 0,5 maka sampel diprediksi termasuk dalam kelas tersebut, dan jika kurang dari 0,5 digunakan sebagai pengklasifikasi biner untuk memprediksi bahwa sampel tersebut tidak termasuk dalam kelas tersebut (Pesantez-Narvaez et al., 2019). Untuk memperkirakan probabilitas, regresi logistik menghitung jumlah tertimbang dari karakteristik input, tetapi alih-alih mengeluarkan hasilnya segera seperti regresi linier, ia mengeluarkan logistik dari nilai hasil. Logistik adalah fungsi sigmoid yang menghasilkan nilai antara 0 dan 1 (Marshall & Wallace, 2019). Fungsi logistik memiliki efek membatasi hasil keluaran untuk selalu antara 0 dan 1 untuk nilai numerik x , dan ekspresinya didefinisikan sebagai berikut.

$$y = \frac{1}{1 + e^{-f(x)}}$$

Dalam (1), (x) dapat berupa fungsi linier sederhana atau fungsi linier berganda (Fernández & Provost, 2019). Untuk masalah klasifikasi dengan dua kategori, jika $(x) > 0$ diklasifikasikan sebagai $\rightarrow 1$ dan $(x) < 0$ diklasifikasikan sebagai $\rightarrow 0$. Batas keputusan model regresi logistik adalah $(x) = 0$ dalam hyperplane dan menjadi $= 0,5$. Kesalahan dalam prediksi biasanya terjadi di sekitar batas keputusan (Lee et al., 2020).

2.2.5 Decision Tree

Decision Tree merupakan teknik klasifier dengan *root node* untuk mengumpulkan data, *inner node* berisi pertanyaan dan *leaf node* sebagai pengambil keputusan untuk sejumlah data yang belum diketahui kelasnya kedalam kelas-kelas yang ada. *Decision tree* classifier dikenal karena kinerjanya yang telah ditingkatkan. Karena presisi yang kuat, splitting parameter yang dioptimalkan, dan teknik *tree pruning* yang telah disempurnakan (ID3, C4.5, CART, CHAID, dan QUEST) biasanya digunakan oleh semua pengklasifikasi data yang dikenal luas. Dataset yang terpisah digunakan untuk melatih sampel dari kumpulan data yang sangat besar, yang pada gilirannya mempengaruhi presisi test set (Charbuty & Abdulazeez, 2021). Kumpulan dari beberapa *tree* terbaik membentuk satu model yang dikenal dengan *Random Forest* (RF). Adanya *Random forest* diyakini dapat memberikan solusi untuk permasalahan *overlap* pada metode *decision tree* ketika menggunakan kriteria dan kelas yang sangat banyak (Kurniawan & Yulianingsih, 2021). Konsep dari pohon keputusan adalah mengubah data menjadi *decision tree* dan aturan-aturan keputusan. Manfaat utama dari penggunaan *decision tree* adalah kemampuannya untuk mem-*break down* proses pengambilan keputusan yang kompleks menjadi lebih simple, sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan. Nama lain dari *decision tree* adalah CART (*Classification and Regression Tree*). Dimana metode ini merupakan gabungan dari dua jenis pohon, yaitu *classification tree* dan juga *regression tree*. Untuk memudahkan, berikut ilustrasi dari keduanya. *Decision tree* juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. *Decision tree* memadukan antara eksplorasi data dan pemodelan.

2.2.6 Random Forest

Random Forest adalah pengembangan dari metode CART, analisis yang dilakukan pada n amatan dan p peubah penjelas, dengan tahapan *bootstrap* dan *random feature selection* yang dilakukan secara berulang sebanyak k kali, sehingga terbentuk sebuah hutan yang terdiri atas k pohon (Tyralis & Papacharalampous, 2019). Hasil pertumbuhan sekelompok pohon dan pemilihan kelas yang paling populer dianggap sebagai peningkatan yang cukup besar dalam performa dalam melakukan prediksi (Cabanac et al., 2021). Pada tahapan *bootstrap* kondisi yang harus diperhatikan adalah data yang digunakan tidak terlalu kecil, sehingga tidak memiliki banyak *outlier* dan bukan data *time series* (Kurniawan & Yulianingsih, 2021). *Random Forest* adalah algoritma pembelajaran ensemble berdasarkan *Decision Tree*. RF menggunakan teknologi *bootstrap* untuk mengekstrak sampel acak dari sampel asli dan membangun pohon keputusan tunggal. Pada setiap *node* pohon keputusan, pilih titik pengurutan menggunakan *subspace* fitur acak. Terakhir, gabungkan pohon keputusan ini untuk mendapatkan hasil prediksi akhir dengan suara terbanyak (Zhu et al., 2021). *Random Forest* adalah salah satu algoritma *Supervised Learning*. Ini digunakan baik untuk klasifikasi dan regresi. Algoritma *Random Forest* memilih sampel acak dari dataset yang diberikan. Kemudian membangun *Decision Tree* untuk setiap sampel dan kemudian mendapatkan prediksi dari masing-masing *Decision Tree*. Pemungutan suara dilakukan untuk setiap hasil prediksi dan keputusan dengan suara terbanyak dipilih sebagai prediksi akhir (Bagga et al., 2020).

2.2.7 Xgboost

Algoritma *XGBoost* digunakan untuk memprediksi karena *XGBoost* adalah salah satu algoritma prediksi paling populer untuk gradient boosting machine (GBM) (Jiang et al., 2019). *Xgboost* telah banyak digunakan di industri karena kinerjanya yang tinggi dalam pemecahan masalah dan persyaratan minimal untuk rekayasa fitur. Kedua, dibandingkan dengan algoritma *deep learning*, *XGBoost* diakui lebih mudah digunakan untuk kumpulan data kecil yang berjalan di CPU (Dong et al., 2020).

2.2.8 Accuracy, Sensitivity, Specificity, AUC

Specificity mengukur proporsi ‘true negative’ yang diidentifikasi dengan benar pada pendeteksi penipuan kartu kredit. Sensitivity mengukur proporsi ‘true positive’ yang diidentifikasi dengan benar pada pendeteksi penipuan kartu kredit. Accuracy adalah tingkat kedekatan antara nilai yang didapat terhadap nilai sebenarnya pada pendeteksi penipuan kartu kredit. AUC adalah akurasi klasifikasi ROC dilakukan dengan cara menghitung luas daerah di bawah kurva ROC pada pendeteksi penipuan kartu kredit (Trivedi et al., 2020).

2.2.9 Python

Python adalah bahasa scripting tingkat tinggi, ditafsirkan, interaktif, dan berorientasi objek yang kuat yang dibuat oleh Guido Van Rossum pada akhir 1980-an. *Python* adalah bahasa yang sangat cocok untuk programmer tingkat pemula dan mendukung pengembangan berbagai aplikasi mulai dari pemrosesan teks sederhana hingga browser www hingga pengembangan game (Melo et al., 2019). *Python* saat ini adalah bahasa pemrograman dengan pertumbuhan tercepat di dunia, berkat kemudahan penggunaan, kurva belajar yang cepat, dan berbagai paket berkualitas tinggi untuk data science dan *machine learning* (Vallat, 2018).

2.3 Penelitian Terkait

Tabel 2.3. 1 Penelitian Terkait

No.	Area Penelitian	Tahun	Karakteristik Data	Metode	Model	Hasil Penelitian
1.	Penerapan <i>Deep Learning</i> dalam Deteksi Penipuan (Randhawa et al., 2018)	2021	Data transaksi banksim	<i>Machine Learning</i>	<i>Machine Learning with Ensemble, Deep Learning.</i>	Nilai akurasi transaksi normal dalam <i>dataset</i> adalah 98.789% yang merupakan nilai akurasi dasar untuk membangun model.
2.	Deteksi Penipuan dengan <i>Machine Learning</i> (Prajapati et al., 2021)	2021	Dataset kartu kredit	<i>Machine Learning</i>	<i>Light GBM, Adaboost, Random Forest Classifier</i>	Hasil terbaik dalam hal akurasi yaitu , <i>AdaBoost</i> yang memberikan hasil tertinggi dengan 0,9613. Dalam hal presisi, <i>Light BGM</i> menghasilkan hasil tertinggi dengan 0,986. Dalam hal <i>recall</i> , <i>Adaboost</i> memberikan <i>recall</i> tertinggi dengan 0.889. Dari segi waktu pelaksanaan,
3.	Model Deteksi Penipuan Kartu Kredit yang Efisien (Trivedi et al., 2020)	2020	Dataset kartu kredit	<i>Machine Learning</i>	<i>Random Forest, Naïve Baiyes. Logistic Regression, SVM, kNN, Decision Tree, GBM</i>	Hasil eksperimen menunjukkan persentase parameter penilaian yang berbeda hanya untuk kumpulan data penipuan kartu kredit untuk teknik <i>machine learniang</i> yang berbeda. Hasil menunjukkan bahwa teknik RF menunjukkan persentase akurasi dengan 95,988 persen, meskipun SVM 93,228 persen, LR 92,89 persen, NB 91,2 persen, DC 90,9 persen serta GBM 93,99 persen menunjukkan persentase presisi identifikasi

						penipuan kartu kredit <i>machine learning</i> ULB.
4.	Klasifikasi pada Kebakaran Hutan dan Lahan (Muslim & Karo, 2020)	2020	Data kebakaran hutan	<i>XGBoost, Feature Importance</i>	<i>Klasifikasi</i>	hasil klasifikasi diperoleh akurasi sebesar 82.51 % dengan running time 10 menit 35 detik. Kami percaya bahwa hasil tersebut tidaklah efektif, mengingat jumlah data yang diolah tidak besar dengan akurasi yang tidak menyentuh angka 85% terlebih lagi memakan waktu yang lama. Selanjutnya akan dipilih beberapa variabel terbaik dengan menggunakan feature importance
5.	Evaluasi Kinerja pada Pendeteksi Penipuan (Mittal & Tyagi, 2019)	2019	Dataset Kartu kredit	<i>Machine Learning, Supervised Learning, Unsupervised Learning.</i>	<i>Random Forest, Neural Networks (NN) and Artificial Neural Networks (ANN), Deep Learning (DL), Support Vector Machine (SVM), Nave Bayes, Logistic Regression, Extended Gradient Boosted Tree (XGBT), Quadratic Discriminant Analysis (QDA), K-Nearest Neighbours</i>	Ada beberapa nilai NaN dalam tabel hasil di mana pengklasifikasi tidak dapat mendeteksi bahkan satu nilai positif atau negatif
6.	Perbandingan Aktifitas System	2021	Data Kartu Kredit	<i>Machine Learning</i>	<i>Decision Tree, Support vector</i>	Hasil yang didapat pada penelitian ini adalah mengetahui metode terbaik

	Deteksi Penipuan dengan Menggunakan Metode Analisis Komprehensif (Veeramani, 2022)				<i>machine, Neural Network, Logistic Regression</i>	sesuai dengan kebutuhan data masing-masing
7.	Deteksi Penipuan Kartu Kredit Menggunakan AdaBoost dan Suara Mayoritas (Randhawa et al., 2018)	2018	Dataset kartu kredit	<i>Machine Learning, AdaBoost</i>	<i>Bayesia, Trees, Neural Network, Linear Regression, Logistic Regression, Support Vector Machine</i>	Hasil semua akurasi masih diatas 90% dan tidak ada yang berada di bawah 30%

BAB III

ANALISA PERUSAHAAN

3.1 Sejarah Dinas ESDM Provinsi Jawa Barat

Dinas Energi dan Sumber Daya Mineral Provinsi Jawa Barat yang saat ini berlokasi di Jalan Sukarno-Hatta Nomor 576 Bandung, pertama kali dibentuk pada tahun 1978 berdasarkan Peraturan Daerah Provinsi Jawa Barat Nomor 4/DP/040/DP/1978 dan kelembagaannya diatur dengan Peraturan Daerah Nomor 5/DP/040/DP/1978. Awalnya dinas ini bernama "Dinas Pertambangan Daerah Provinsi Jawa Barat". Peraturan Daerah ini baru dapat dilaksanakan pada tahun 1986 setelah ditetapkan Peraturan Pemerintah Nomor 37 Tahun 1986 tentang Penyerahan sebagian urusan pemerintahan di bidang pertambangan kepada Pemerintah Daerah Tingkat I, mengingat makin sulitnya mengontrol kegiatan usaha penambangan di daerah terutama untuk jenis bahan galian golongan C yang dianggap telah banyak menyebabkan kerusakan serta mengganggu daya dukung dan fungsi lingkungan, sehingga peran Pemerintah Daerah perlu lebih dipertegas dalam pengaturan dan pengendaliannya, melalui penyerahan urusan. Karena Peraturan Daerah yang ditetapkan pada tahun 1978 tersebut dinilai sudah tidak sesuai lagi pada saat pengesahan PP tersebut, maka pada tahun 1988 dilakukan pencabutan Peraturan Daerah Nomor 5/DP/040/DP/1978 yang diganti dengan Peraturan Daerah Nomor 12 Tahun 1988 tentang perubahan Struktur Organisasi dan Tata Kerja, dengan susunan organisasi.

Pada tahun 2000, seiring bergulirnya Otonomi Daerah dan makin bertambahnya peran yang menjadi tanggungjawab oleh lembaga ini, maka dilakukan perubahan Struktur Organisasi dan Tata Kerja (SOTK) Dinas Pertambangan, sehingga nomenklatur dinas pun disesuaikan menjadi "Dinas Pertambangan dan Energi Provinsi Jawa Barat" sebagaimana tertuang dalam Peraturan Daerah No.15 Tahun 2000 tentang Struktur Organisasi dan Tata Kerja Dinas-Dinas Daerah Provinsi Jawa Barat, termasuk penyesuaian struktur organisasi dan perangkat yang ada di bawahnya. Adapun penyesuaian nama perangkat dinas tersebut, meliputi : 1 (satu) Bagian Tata Usaha dan 5 (lima) Sub Dinas yaitu Sub Dinas Bina Program, Sub Dinas Eksplorasi dan Pemetaan, Sub Dinas Bina Usaha, Sub Dinas Bina Teknik, Sub Dinas Konservasi dan Pengendalian. Selain perangkat tersebut di atas, Dinas Pertambangan dan Energi Provinsi Jawa Barat juga dilengkapi dengan Unit Pelaksanaan Teknis Dinas (UPTD) yang ditetapkan dengan Peraturan Daerah Nomor 5 Tahun 2002.

UPTD bidang Pertambangan dan Energi tersebut terdiri dari 5 (lima) balai yaitu, Balai Konservasi dan Pemanfaatan Sumberdaya Pertambangan dan Energi Wilayah I Cianjur, Balai Konservasi dan Pemanfaatan Sumberdaya Pertambangan dan Energi Wilayah II Purwakarta, Balai Konservasi dan Pemanfaatan Sumberdaya Pertambangan dan Energi Wilayah III Tasikmalaya, Balai Konservasi dan Pemanfaatan Sumberdaya Pertambangan dan Energi Wilayah IV Cirebon, serta Balai Penelitian Pertambangan dan Energi. Selain itu, keberadaan Dinas Pertambangan dan Energi Provinsi Jawa Barat merupakan konsekuensi dari 10 diberlakukannya Undang-undang Nomor 22 Tahun 1999 jo. Undang-undang Nomor 34 Tahun 2002 Tentang Pemerintahan Daerah dan Peraturan Pemerintah Nomor 25 Tahun 2000 tentang Kewenangan pemerintah dan Kewenangan Provinsi sebagai Daerah Otonom.

Di akhir tahun 2008 bersamaan dengan masa pergantian Gubernur dilakukan reorganisasi di lingkungan Pemerintah Provinsi Jawa Barat yang merupakan pelaksanaan dari Peraturan Pemerintah No. 38 Tahun 2007 tentang Pembagian Urusan Pemerintah antara Pemerintah, Pemerintah Daerah Provinsi dan Pemerintah Daerah Kabupaten/Kota serta Peraturan

Pemerintah Nomor 41 Tahun 2007 tentang Organisasi Perangkat Daerah. Atas kebijakan tersebut, maka nomenklatur Dinas Pertambangan dan Energi berubah menjadi "Dinas Energi dan Sumber Daya Mineral Provinsi Jawa Barat". Perubahan nomenklatur dinas tersebut turut merubah struktur organisasi yang ada di lingkungan Dinas Energi dan Sumber Daya Mineral yang awalnya terdiri dari 1 (satu) Bagian Tata Usaha dan 5 (lima) Sub Dinas menjadi 1(satu) Sekretariat dan 4 (empat) Bidang yaitu : Bidang Listrik dan Pemanfaatan Energi, Bidang Mineral Geologi dan Air, Bidang Panas Bumi dan Migas dan Bidang Bina Usaha dan Kerjasama sebagaimana diatur dalam Peraturan Daerah Nomor 21 Tahun 2008 tentang Organisasi dan Tata Kerja Dinas Daerah Provinsi Jawa Barat.

3.2 Visi dan Misi

1. VISI

Terwujudnya Jawa Barat termaju dalam pembangunan pertambangan dan energi yang berkelanjutan dan berpihak kepada kepentingan rakyat.

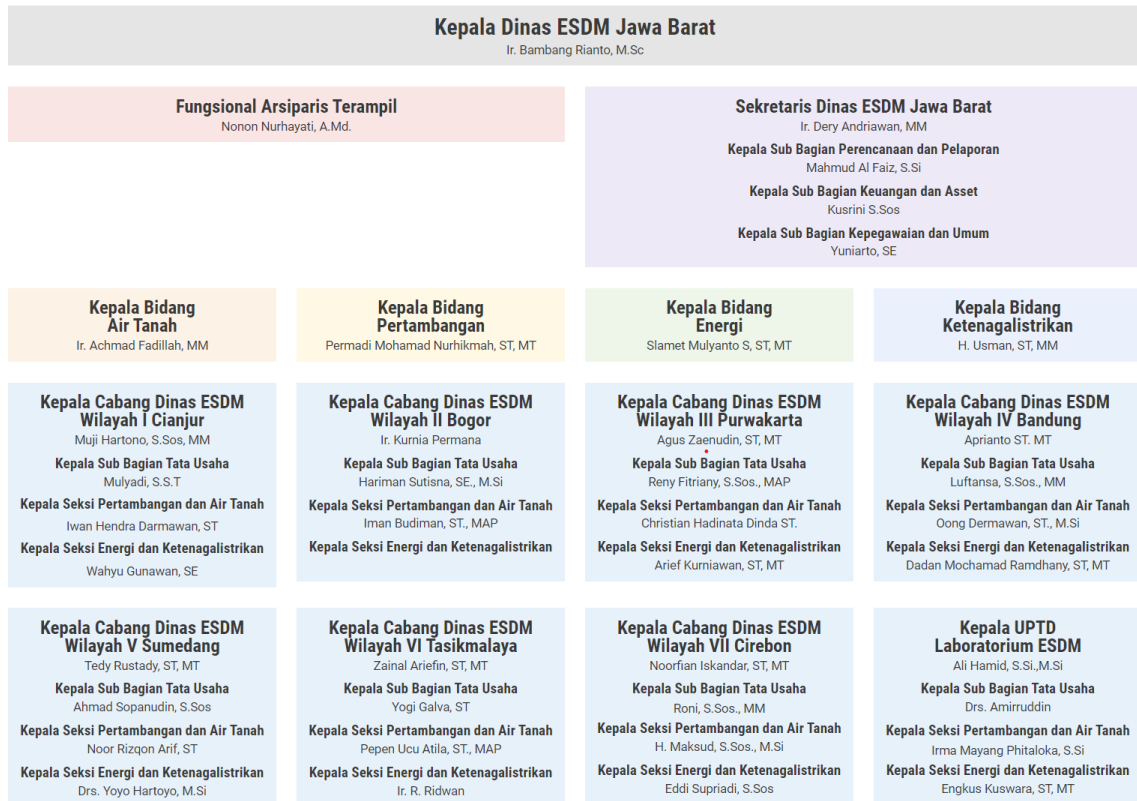
2. MISI

- a. Meningkatkan pelayanan publik untuk membangun pertambangan energi yang maju.
- b. Mendorong peningkatan investasi dan kontribusi usaha di bidang pertambangan dan energi terhadap perekonomian regional.
- c. Mendorong terwujudnya pengelolaan pertambangan dan energi yang berkelanjutan.

3.3 Strategi Dinas ESDM Provinsi Jawa Barat

- a. Meningkatkan ketersediaan dan pemanfaatan energi baru terbarukan.
- b. Meningkatkan konservasi sumber daya energi.
- c. Meningkatkan akses elektrifikasi dan daya untuk menunjang produktivitas masyarakat.
- d. Melakukan penataan perusahaan ketenagalistrikan.
- e. Meningkatkan keamanan ketenagalistrikan melalui SLO dan tenaga teknik bersertifikat.
- f. Melakukan pemantauan muka air tanah secara real time.
- g. Optimalisasi pengimmbuhan air tanah.
- h. Meningkatkan pengelolaan perusahaan sumber daya pertambangan.
- i. Melaksanakan Pembinaan Teknis Terhadap Pelaku Usaha Pertambangan.
- j. Inventarisasi Data dan Pengembangan Usaha Pertambangan.
- k. Meningkatkan tingkat keikutsertaan aparatur dalam diklat/bimbingan teknis.
- l. Meningkatkan kualitas Sistem Pelayanan Perencanaan dan pelaporan Dinas Energi dan Sumber Daya Mineral.
- m. Menyediakan sarana dan prasarana untuk membantu pelaksanaan tugas aparatur.

3.4 Struktur Organisasi Perusahaan



Gambar 3.4. 1 Struktur Organisasi Dinas ESDM Provinsi Jawa Barat

3.5 Deskripsi dan Ruang Lingkup Internship

Pada tahapan persiapan kegiatan *Internship* meliputi beberapa kegiatan yaitu:

- Internship* adalah salah satu matakuliah wajib yang harus diikuti oleh penulis.
- Program ini dilakukan dan dimulai dengan sosialisasi I dan II selanjutnya. Sosialisasi ini bertujuan untuk memberikan informasi *update* terkait masalah pelaksanaan *internship* yang akan dilaksanakan.
- Setelah kegiatan sosialisasi, penulis selanjutnya mengajukan surat permohonan kepada perusahaan dengan syarat melampirkan surat dari kampus atau prodi, surat izin orang tua, dan surat siap melaksakan *internship* sampai dengan selesai.

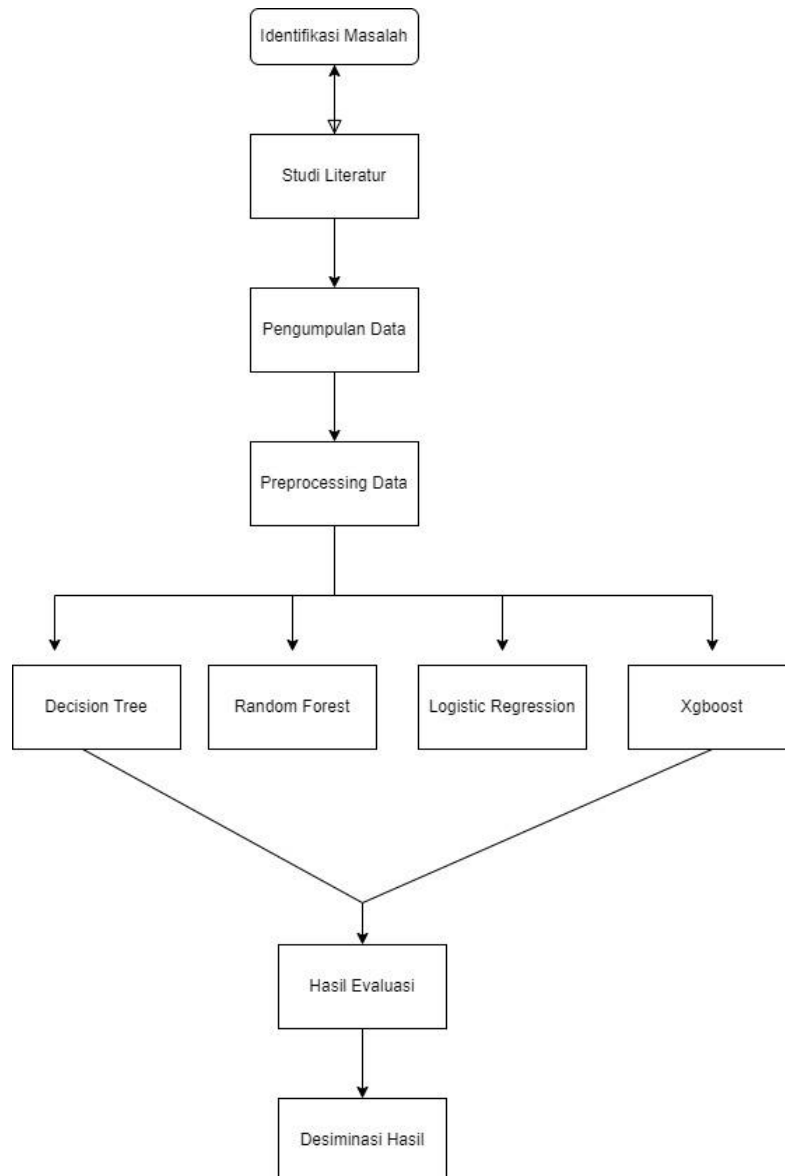
Adapun ruang lingkup *internship* kegiatan *Internship* di Dinas ESDM Provinsi Jawa Barat yaitu:

- Merancang Website Sistem Informasi Prioritas pemenuhan Sarana dan Prasarana Kerja Aparatur Ruang Kerja Dinas ESDM Provinsi Jawa Barat.

BAB IV

METODE PENELITIAN

4.1 Diagram Alur Metodologi Penelitian



Gambar 4. 1 Alur Metodologi Penelitian

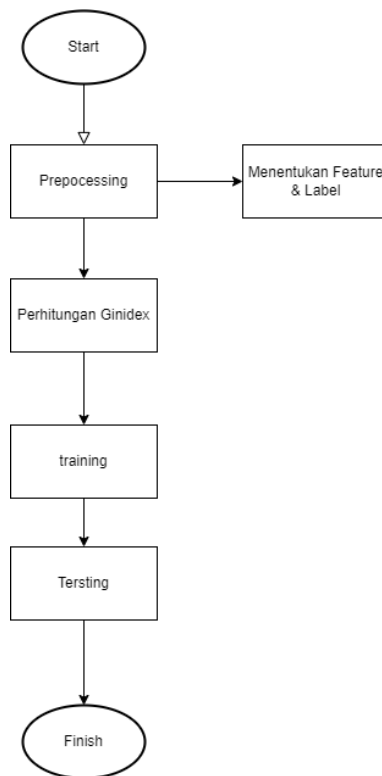
4.2 Indikator Capaian Penelitian

Berdasarkan diagram alur metodologi penelitian diatas, terdapat indikator capaian sebagai berikut.

Tabel 4.2. 1 Indikator Capaian

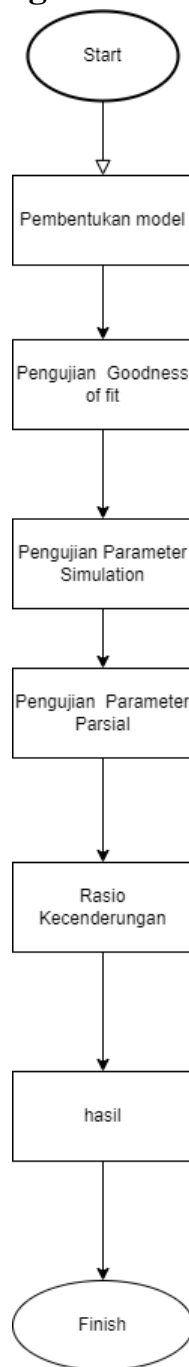
No.	Tahapan		Indikator capaian
1.	Identifikasi Masalah	→	1. Rancangan Researce question.
2.	Studi Literatur	→	2. Deskripsi teori, temuan dan bahan penelitian terkait
3.	Pengumpulan Data	→	3. Pengumpulan data transaksi kartu kredit pada kaggle dataset yang dijadikan bahan penelitian
4.	<i>Preprocessing</i> Data	→	4. <i>Preprocessing</i> data merupakan tahap mempersiapkan data untuk diselanjutnya dilakukan analisis. Beberapa tahap yang dilakukan pada prepoessing yaitu mempersiapkan 4 metode yaitu <i>decision tree</i> , <i>random forest</i> , <i>logistic regression</i> , dan <i>xgboost</i> .
6.	Evaluasi Hasil	→	5. Hasil penelitian.
7.	Diseminasi Hasil	→	6. Laporan internship II dan draft jurnal.

4.3 Alur Metodologi Random Forest



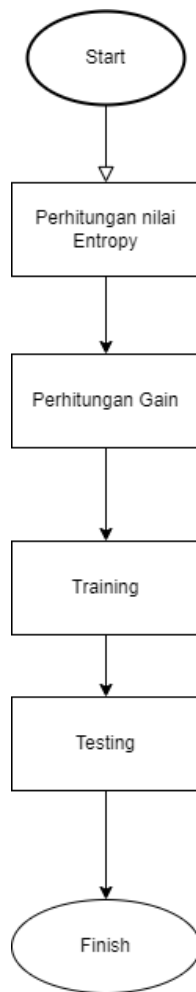
Tabel 4.3. 1 Alur Metodologi Random Forest

4.4 Alur Metodologi Logistic Regression



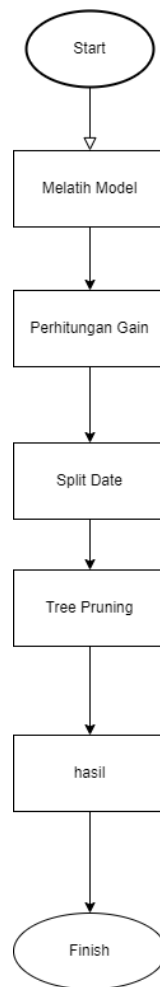
Tabel 4.4. 1 Alur Metodologi Logistic Regression

4.5 Alur Metodologi Decision Tree



Tabel 4.5. 1 Alur Metodologi Decision Tree

4.6 Alur Metodologi Xgboost



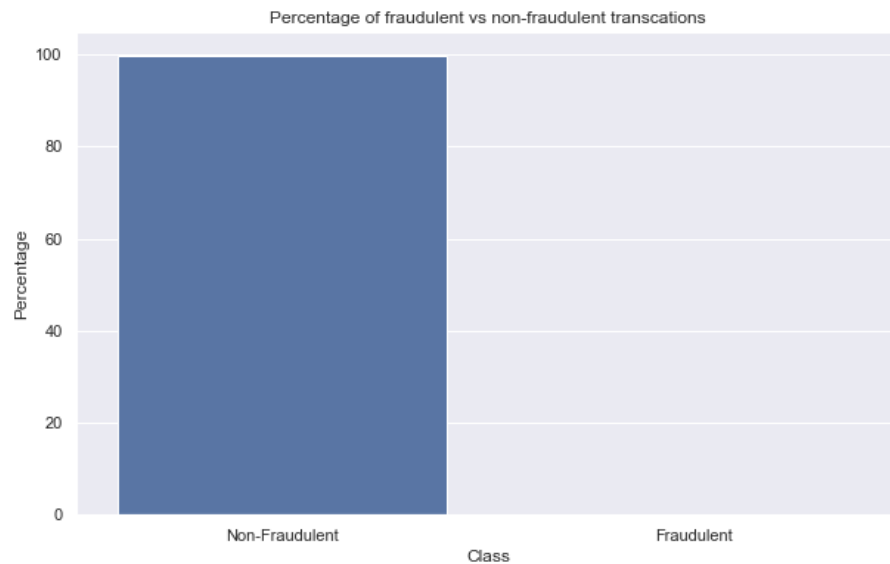
Tabel 4.6. 1 Alur Metodologi Xgboost

BAB V

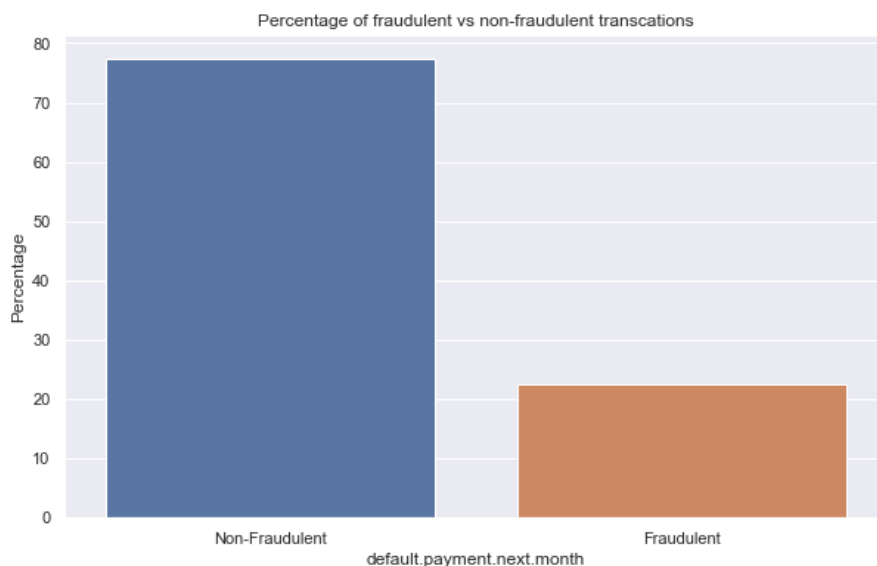
HASIL PENELITIAN

5.1 Hasil Penelitian

Dari hasil penelitian dapat dilihat dari kedua dataset yang digunakan maka hasil perbandingan hasil pendeteksi penipuan seperti gambar dibawah ini:



Gambar 5. 1 Perbandingan Spesifikasi dataset Credit Card Eropa



Gambar 5. 2 Perbandingan Spesifikasi dataset Credit Card Taiwan

5.1.1 Dataset

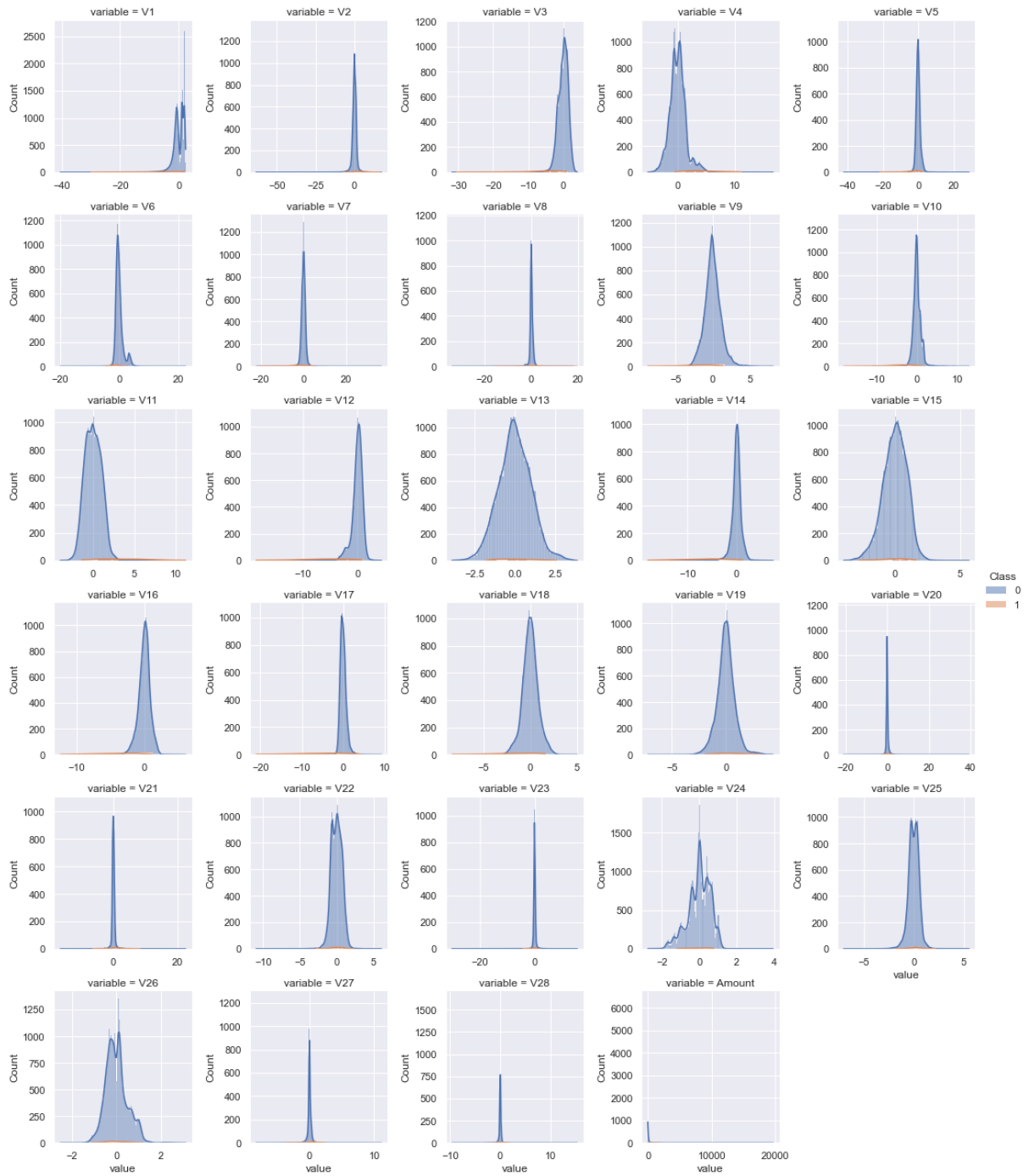
Penelitian ini menggunakan dataset deteksi penipuan kartu kredit yang terkenal yang bersumber dari kaggle <https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud> dan juga <https://www.kaggle.com/code/sanjoy1110/credit-card-taiwan-using-support-vector-machine/data>. Kumpulan data disiapkan oleh Grup Pembelajaran Mesin Université Libre de

Bruxelles (ULB) tentang penambahan data besar dan deteksi penipuan. Dataset berisi transaksi kartu kredit yang dilakukan dalam dua hari pada bulan September 2021 oleh klien kartu kredit Eropa. Dataset tidak seimbang, dengan hanya 492 transaksi penipuan dari 284.807. Sementara itu, semua atribut kecuali "Waktu" dan "Jumlah" adalah numeric karena transformasi yang dilakukan pada dataset, dan dikodekan sebagai V_1, V_2, \dots, V_{28} untuk alasan kerahasiaan. Atribut "Amount" adalah biaya transaksi dan atribut "Time" adalah detik yang berlalu antara transaksi dan transaksi pertama dalam kumpulan data. Terakhir, atribut "Kelas" adalah variabel dependen, dan memiliki nilai 1 untuk transaksi penipuan dan 0 untuk transaksi yang sah. Dari dataset yang digunakan maka dapat dilakukan deteksi penipuan untuk mengetahui jumlah penipuan pada transaksi. Dataset dapat dilihat pada tabel 5.1 - 5.2 seperti berikut ini.

Tabel 5. 1 Dataset Credit Card Eropa

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
0	0.0	-1.359807	-0.072781	2.536347	1.378155	-0.338321	0.462388	0.239599	0.098698	0.363787	0.090794	-0.551600
1	0.0	1.191857	0.266151	0.166480	0.448154	0.060018	-0.082361	-0.078803	0.085102	-0.255425	-0.166974	1.612727
2	1.0	-1.358354	-1.340163	1.773209	0.379780	-0.503198	1.800499	0.791461	0.247676	-1.514654	0.207643	0.624501
3	1.0	-0.966272	-0.185226	1.792993	-0.863291	-0.010309	1.247203	0.237609	0.377436	-1.387024	-0.054952	-0.226487
4	2.0	-1.158233	0.877737	1.548718	0.403034	-0.407193	0.095921	0.592941	-0.270533	0.817739	0.753074	-0.822843
		V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	
		-0.617801	-0.991390	-0.311169	1.468177	-0.470401	0.207971	0.025791	0.403993	0.251412	-0.018307	
		1.065235	0.489095	-0.143772	0.635558	0.463917	-0.114805	-0.183361	-0.145783	-0.069083	-0.225775	
		0.066084	0.717293	-0.165946	2.345865	-2.890083	1.109969	-0.121359	-2.261857	0.524980	0.247998	
		0.178228	0.507757	-0.287924	-0.631418	-1.059647	-0.684093	1.965775	-1.232622	-0.208038	-0.108300	
		0.538196	1.345852	-1.119670	0.175121	-0.451449	-0.237033	-0.038195	0.803487	0.408542	-0.009431	
		V22	V23	V24	V25	V26	V27	V28	Amount	Class		
		0.277838	-0.110474	0.066928	0.128539	-0.189115	0.133558	-0.021053	149.62	0		
		-0.638672	0.101288	-0.339846	0.167170	0.125895	-0.008983	0.014724	2.69	0		
		0.771679	0.909412	-0.689281	-0.327642	-0.139097	-0.055353	-0.059752	378.66	0		
		0.005274	-0.190321	-1.175575	0.647376	-0.221929	0.062723	0.061458	123.50	0		
		0.798278	-0.137458	0.141267	-0.206010	0.502292	0.219422	0.215153	69.99	0		

Selain dari tabel maka data dapat dilihat dari hasil histogram masing-masing data seperti gambar 5.3 dibawah ini:



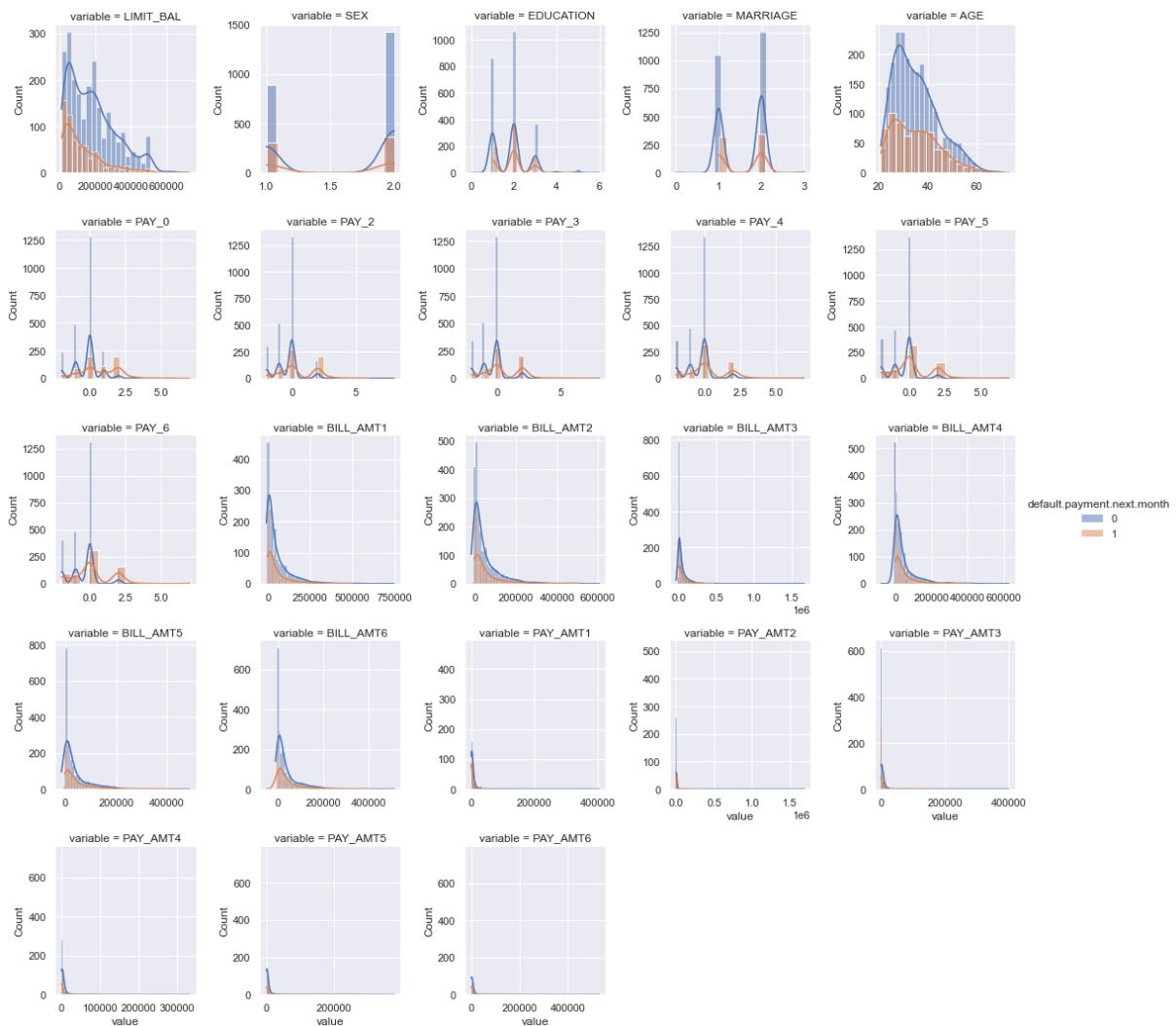
Gambar 5. 3 Histogram Dataset Credit Card Eropa

Tabel 5. 2 Dataset Credit Card Taiwan

	ID	LIMIT_BAL	SEX	EDUCATION	MARRIAGE	AGE	PAY_0	PAY_2	PAY_3	PAY_4	PAY_5	PAY_6	BILL_AMT1	BILL_AMT2	BILL_AMT3	BILL_AMT4
0	1	20000.0	2	2	1	24	2	2	-1	-1	-2	-2	3913.0	3102.0	689.0	0.0
1	2	120000.0	2	2	2	26	-1	2	0	0	0	2	2682.0	1725.0	2682.0	3272.0
2	3	90000.0	2	2	2	25	0	0	0	0	0	0	29239.0	14027.0	13559.0	14331.0
3	4	50000.0	2	2	1	25	0	0	0	0	0	0	46990.0	48233.0	49291.0	28314.0
4	5	50000.0	1	2	1	57	-1	0	-1	0	0	0	8617.0	5670.0	35835.0	20940.0

BILL_AMT5	BILL_AMT6	PAY_AMT1	PAY_AMT2	PAY_AMT3	PAY_AMT4	PAY_AMT5	PAY_AMT6	default.payment.next.month
0.0	0.0	0.0	689.0	0.0	0.0	0.0	0.0	1
3455.0	3261.0	0.0	1000.0	1000.0	1000.0	0.0	2000.0	1
14948.0	15549.0	1518.0	1500.0	1000.0	1000.0	1000.0	5000.0	0
28959.0	29547.0	2000.0	2018.0	1200.0	1100.0	1069.0	1000.0	0
19146.0	19131.0	2000.0	36681.0	10000.0	9000.0	689.0	679.0	0

Selain dari tabel maka data dapat dilihat dari hasil histogram masing-masing data seperti gambar 5.4 dibawah ini:



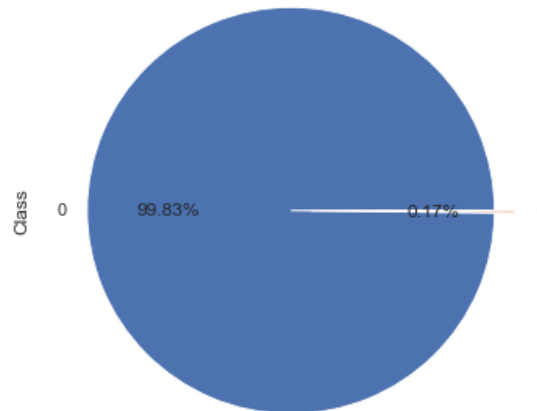
Gambar 5. 4 Histogram Dataset Credit Card Taiwan

Adapun Tabel 5.3 menunjukkan jumlah data kartu kredit tanpa penipuan dan penipuan yang ditunjukkan sebagai berikut:

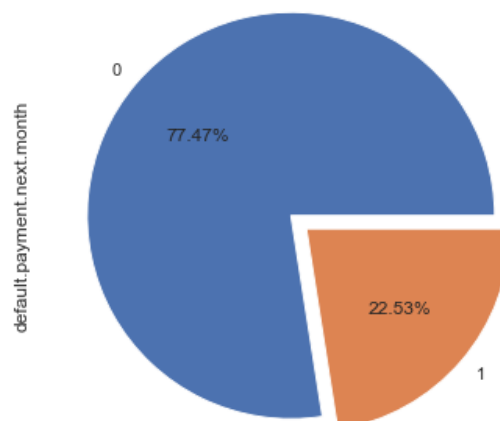
Tabel 5. 3 Jumlah Data

<i>Dataset</i>	Jumlah Data	
	Tanpa Penipuan	Penipuan
Credit Card Eropa	28432	49
Credit Card Taiwan	2324	676

Perbandingan jumlah data juga dapat dilihat dari hasil diagram berikut ini:



Gambar 5. 5 Diagram Dataset Credit Card Eropa



Gambar 5. 6 Diagram Dataset Credit Card Taiwan

5.1.2 Hasil Eksperimen Model

A. Hasil Implementasi *Decission Tree*

Dari penelitian yang sudah dilakukan pada dataset Credit Card dengan menggunakan algoritma *Decision tree* hasil *accuracy* sebesar 99%, *sensitivity* sebesar 0%, *specificity* sebesar 100%, dan *AUC* sebesar 83%, sedangkan dari hasil Dataset Credit Card Taiwan dengan menggunakan algoritma *Decision tree* hasil *accuracy* sebesar 83%, *sensitivity* sebesar 39%, *specificity* sebesar 96%, dan *AUC* sebesar 76%, maka hasil akurasi serta eksperimen dapat dilihat seperti dibawah ini:

Tabel 5. 4 Hasil Eksperimen Dengan Algoritma *Decission Tree*

Dataset	Accuracy	Sensitivity	Specificity	AUC
Credit Card	99%	0%	100%	83%
Credit Card Taiwan	83%	39%	96%	76%

B. Hasil Implementasi *Random Forest*

Dari penelitian yang sudah dilakukan pada dataset Credit Card dengan menggunakan algoritma *Random Forest* hasil *accuracy* sebesar 99%, *sensitivity* sebesar 0%, *specificity* sebesar 100%, dan *AUC* sebesar 99%, sedangkan dari hasil Dataset Credit Card Taiwan dengan menggunakan algoritma *Random Forest* hasil *accuracy* sebesar 82%, *sensitivity* sebesar 40%, *specificity* sebesar 95%, dan *AUC* sebesar 77%, maka hasil akurasi serta eksperimen dapat dilihat seperti dibawah ini:

Tabel 5. 5 Hasil Eksperimen Dengan Algoritma *Random Forest*

Dataset	Accuracy	Sensitivity	Specificity	AUC
Credit Card	99%	0%	100%	99%
Credit Card Taiwan	82%	40%	95%	77%

C. Hasil Implementasi *Logistic regression*

Dari penelitian yang sudah dilakukan pada dataset Credit Card dengan menggunakan algoritma *Random Forest* hasil *accuracy* sebesar 99%, *sensitivity* sebesar 56%, *specificity* sebesar 99%, dan *AUC* sebesar 97%, sedangkan dari hasil Dataset Credit Card Taiwan dengan menggunakan algoritma *Random Forest* hasil *accuracy* sebesar 77%, *sensitivity* sebesar 0%, *specificity* sebesar 100%, dan *AUC* sebesar 64%, maka hasil akurasi serta eksperimen dapat dilihat seperti dibawah ini:

Tabel 5. 6 Hasil Eksperimen Dengan Algoritma *Logistic Regression*

Dataset	Accuracy	Sensitivity	Specificity	AUC
Credit Card	99%	56%	99%	97%
Credit Card Taiwan	77%	0%	100%	64%

D. Hasil Implementasi *Xgboost*

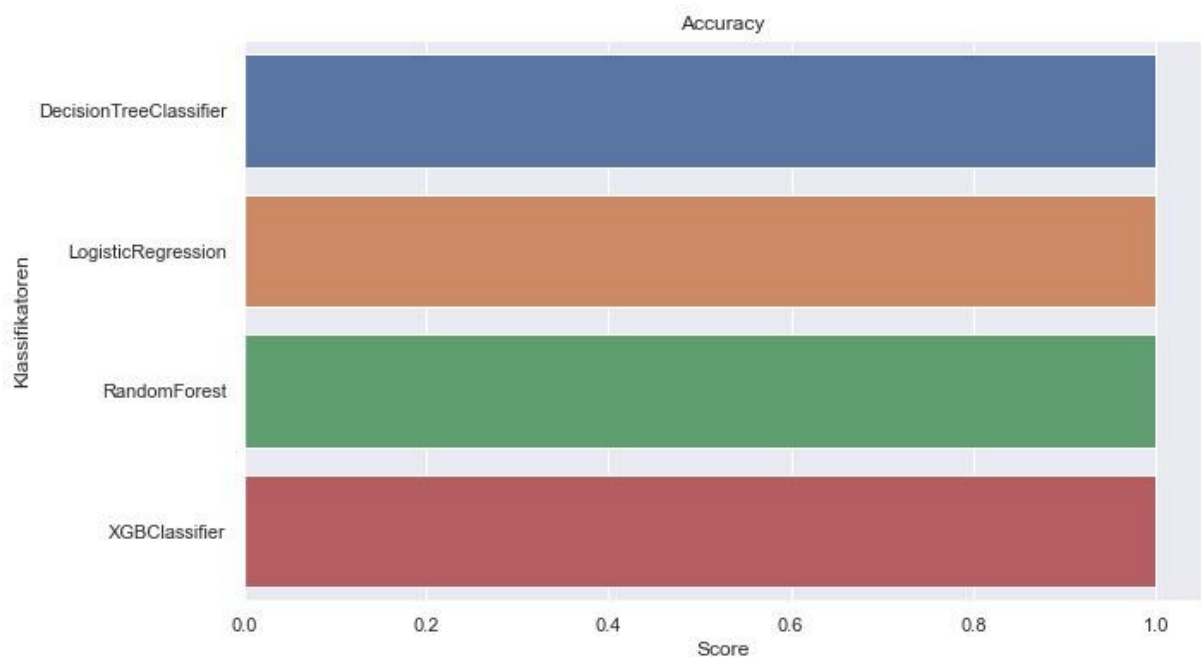
Dari penelitian yang sudah dilakukan pada dataset Credit Card dengan menggunakan algoritma *Random Forest* hasil *accuracy* sebesar 99%, *sensitivity* sebesar 67%, *specificity* sebesar 99%, dan *AUC* sebesar 99%, sedangkan dari hasil Dataset Credit Card Taiwan dengan menggunakan algoritma

Random Forest hasil *accuracy* sebesar 81%, *sensitivity* sebesar 40%, *specificity* sebesar 93%, dan *AUC* sebesar 73%, maka hasil akurasi serta eksperimen dapat dilihat seperti dibawah ini:

Tabel 5. 7 Hasil Eksperimen Dengan Algoritma Xgboost

Dataset	Accuracy	Sensitivity	Specificity	AUC
Credit Card	99%	67%	99%	99%
Credit Card Taiwan	81%	40%	93%	73%

Hasil penelitian dengan menggunakan perbandingan empat algoritma yaitu *decision tree*, *random forest*, *logistic regression*, dan *xgboost* yang menunjukan bahwa dengan perbandingan dua dataset yang digunakan dapat dilihat bahwa dengan menggunakan model algoritma *xgboost* pada *dataset credit card* yang memiliki hasil paling akurat daripada tiga model algoritma yang lain. Sedangkan dengan menggunakan model algoritma *xgboost* pada *dataset credit card taiwan* yang memiliki hasil paling akurat daripada tiga model algoritma yang lain. Sehingga hasil perbandingan keempat metode ini dapat dilihat dari diagram berikut ini :



BAB VI

PENUTUP

6.1 Kesimpulan

Dari hasil penelitian yang dilakukan dari perbandingan antara 2 dataset dengan menggunakan algoritma *decision tree*, *random forest*, *logistic regression*, dan *xgboost*. Maka dapat dilihat bahwa hasil dari dataset *credit card* hasilnya lebih baik karena besar penipuan hanya 0.17% banding 99.83% tanpa penipuan dengan jumlah data tanpa penipuan terdapat 28.432 dan dengan penipuan 49, sedangkan dataset pembanding *credit card* Taiwan yang lebih banyak data penipuan dengan besar 22.53% banding 77.47% tanpa penipuan dengan jumlah data 676 penipuan dan 2324 tanpa penipuan. Sedangkan dari ke 4 model algoritma *Xgboost* hasil *accuracy* sebesar 99%, *sensitivity* sebesar 67%, *specificity* sebesar 99%, dan *AUC* sebesar 99%. Maka dapat ditarik kesimpulan bahwa algoritma model *xgboost* lebih optimal hasil akurasiya dibandingkan ke tiga model yang lainnya.

DAFTAR PUSTAKA

- Bagga, S., Goyal, A., Gupta, N., & Goyal, A. (2020). ScienceDirect Credit Card Fraud Detection ICITETM2020 using Pipeling and Ensemble Learning Credit Card Fraud Detection using Ensemble a Pipeling and Goyal c Learning. *Procedia Computer Science*, 173(2019), 104–112. <https://doi.org/10.1016/j.procs.2020.06.014>
- Brunton, S. L., Noack, B. R., & Koumoutsakos, P. (2020). Machine Learning for Fluid Mechanics. *Annual Review of Fluid Mechanics*, 52, 477–508. <https://doi.org/10.1146/annurev-fluid-010719-060214>
- Cabanac, G., Labbé, C., & Magazinov, A. (2021). *Tortured phrases: A dubious writing style emerging in science. Evidence of critical issues affecting established journals.* <http://arxiv.org/abs/2107.06751>
- Charbuty, B., & Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, 2(01), 20–28. <https://doi.org/10.38094/jastt20165>
- Dairu, X., & Shilong, Z. (2021). Machine Learning Model for Sales Forecasting by Using XGBoost. *2021 IEEE International Conference on Consumer Electronics and Computer Engineering, ICCECE 2021, Iccece*, 480–483. <https://doi.org/10.1109/ICCECE51280.2021.9342304>
- Dong, W., Huang, Y., Lehane, B., & Ma, G. (2020). XGBoost algorithm-based prediction of concrete electrical resistivity for structural health monitoring. *Automation in Construction*, 114(January), 103155. <https://doi.org/10.1016/j.autcon.2020.103155>
- Fernández, C., & Provost, F. (2019). Causal Classification: Treatment Effect vs. Outcome Prediction. *SSRN Electronic Journal*, 1–25. <https://doi.org/10.2139/ssrn.3408524>
- Hairani, H., & Bumigora, U. (2020). *K-means-SMOTE untuk menangani ketidakseimbangan kelas dalam klasifikasi penyakit diabetes dengan C4 . 5 , SVM , dan naive Bayes K-means-SMOTE for handling class imbalance in the classification of diabetes. April.* <https://doi.org/10.14710/jtsiskom.8.2.2020.89-93>
- Jiang, Y., Tong, G., Yin, H., & Xiong, N. (2019). A Pedestrian Detection Method Based on Genetic Algorithm for Optimize XGBoost Training Parameters. *IEEE Access*, 7, 118310–118321. <https://doi.org/10.1109/ACCESS.2019.2936454>
- Kalid, S. N., Ng, K. H., Tong, G. K., & Khor, K. C. (2020). A Multiple Classifiers System for Anomaly Detection in Credit Card Data with Unbalanced and Overlapped Classes. *IEEE Access*, 8, 28210–28221. <https://doi.org/10.1109/ACCESS.2020.2972009>

- Kurniawan, A., & Yulianingsih, Y. (2021). Pendugaan Fraud Detection pada kartu kredit dengan Machine Learning. *Kilat*, 10(2), 320–325. <https://doi.org/10.33322/kilat.v10i2.1482>
- Lee, K., Joo, S., Baik, H., Han, S., & In, J. (2020). Unbalanced data, type II error, and nonlinearity in predicting M&A failure. *Journal of Business Research*, 109(November 2019), 271–287. <https://doi.org/10.1016/j.jbusres.2019.11.083>
- Marshall, I. J., & Wallace, B. C. (2019). Toward systematic review automation: A practical guide to using machine learning tools in research synthesis. *Systematic Reviews*, 8(1), 1–10. <https://doi.org/10.1186/s13643-019-1074-9>
- Melo, L. S., Sampaio, R. F., Leão, R. P. S., Barroso, G. C., & Bezerra, J. R. (2019). Python-based multi-agent platform for application on power grids. *International Transactions on Electrical Energy Systems*, 29(6), 1–14. <https://doi.org/10.1002/2050-7038.12012>
- Mittal, S., & Tyagi, S. (2019). Performance evaluation of machine learning algorithms for credit card fraud detection. *Proceedings of the 9th International Conference On Cloud Computing, Data Science and Engineering, Confluence 2019*, 320–324. <https://doi.org/10.1109/CONFLUENCE.2019.8776925>
- Muslim, I., & Karo, K. (2020). Implementasi Metode XGBoost dan Feature Importance untuk Klasifikasi pada Kebakaran Hutan dan Lahan. *Journal of Software Engineering, Information and Communication Technology*, 1(1), 10–16.
- Nath, V. (2020). ScienceDirect Credit Card Fraud Detection using Machine Learning Algorithms Credit Card Fraud Detection using Machine Learning Algorithms. *Procedia Computer Science*, 165, 631–641. <https://doi.org/10.1016/j.procs.2020.01.057>
- Neural, A. (2020). *Prediksi Hipertensi Menggunakan Decision Tree , Naïve Bayes dan Artificial Neural Network pada Software KNIME*. 19(4), 353–363.
- Perwara, L. D., & Bachtiar, F. A. (2020). *Penerapan Algoritma Decision Tree C4 . 5 Untuk Deteksi Fraud Pada Kartu Kredit dengan Oversampling Synthetic Minority Technique (SMOTE)*. 4(8), 2664–2669.
- Pesantez-Narvaez, J., Guillen, M., & Alcañiz, M. (2019). Predicting motor insurance claims using telematics data—XGboost versus logistic regression. *Risks*, 7(2). <https://doi.org/10.3390/risks7020070>
- Prajapati, D., Tripathi, A., Mehta, J., Jhaveri, K., & Kelkar, V. (2021). Credit Card Fraud Detection Using Machine Learning. *2021 7th IEEE International Conference on Advances in Computing, Communication and Control, ICAC3 2021*, 8592(1), 16–19.

- <https://doi.org/10.1109/ICAC353642.2021.9697227>
- Ramdhani, A. (2019). Studi Algoritma Linear Support Vector Machine pada Deteksi Ujaran Kebencian Berbahasa Indonesia. *Prosiding Seminar Nasional Teknoka*, 3(2502), 42. <https://doi.org/10.22236/teknoka.v3i0.2899>
- Randhawa, K., Loo, C. K., Seera, M., Lim, C. P., & Nandi, A. K. (2018). Credit Card Fraud Detection Using AdaBoost and Majority Voting. *IEEE Access*, 6, 14277–14284. <https://doi.org/10.1109/ACCESS.2018.2806420>
- Roscher, R., Bohn, B., Duarte, M. F., & Garcke, J. (2020). Explainable Machine Learning for Scientific Insights and Discoveries. *IEEE Access*, 8, 42200–42216. <https://doi.org/10.1109/ACCESS.2020.2976199>
- Sistem, R. (2021). *JURNAL RESTI Analisis Perbandingan Algoritma Optimasi pada Random Forest untuk*. 1(10), 187–192.
- Thabtah, F., Abdelhamid, N., & Peebles, D. (2019). A machine learning autism classification based on logistic regression analysis. *Health Information Science and Systems*, 7(1), 1–11. <https://doi.org/10.1007/s13755-019-0073-5>
- Tingfei, H., Guangquan, C., & Kuihua, H. (2020). Using Variational Auto Encoding in Credit Card Fraud Detection. *IEEE Access*, 8, 149841–149853. <https://doi.org/10.1109/ACCESS.2020.3015600>
- Trivedi, N. K., Simaiya, S., Lilhore, U. K., & Sharma, S. K. (2020). An efficient credit card fraud detection model based on machine learning methods. *International Journal of Advanced Science and Technology*, 29(5), 3414–3424.
- Tyralis, H., & Papacharalampous, G. (2019). Scientists and Practitioners and Their Recent History. *Water*.
- Vallat, R. (2018). Pingouin: statistics in Python. *Journal of Open Source Software*, 3(31), 1026. <https://doi.org/10.21105/joss.01026>
- Veeramani, V. (2022). *Comparison of Frequent Fraud Activity Detection System by using the Comprehensive Analysis Method*. October 2021.
- Yuanika, M., Ayu, N., & Nurul, D. (2021). Analisis perbandingan algoritma Naïve Bayes , k-Nearest Neighbor dan Neural Network untuk permasalahan class-imbalanced data pada kasus credit card fraud dataset Comparative analysis of the Naïve Bayes , k-Nearest Neighbor and Neural Network algorithms fo. 11(2), 69–73.
- Yun, H. (2021). Prediction model of algal blooms using logistic regression and confusion matrix. *International Journal of Electrical and Computer Engineering*, 11(3), 2407–

2413. <https://doi.org/10.11591/ijece.v11i3.pp2407-2413>

Zamachsari, F., & Puspitasari, N. (2021). Penerapan Deep Learning dalam Deteksi Penipuan Transaksi Keuangan Secara Elektronik. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(2), 203–212. <https://doi.org/10.29207/resti.v5i2.2952>

Zhu, L., Zhou, X., & Zhang, C. (2021). Rapid identification of high-quality marine shale gas reservoirs based on the oversampling method and random forest algorithm. *Artificial Intelligence in Geosciences*, 2(July), 76–81. <https://doi.org/10.1016/j.aiig.2021.12.001>

LAMPIRAN

The screenshot shows a web browser window with the URL check-plagiarism.com/id/. The page displays a plagiarism check result for a document. At the top, there are two boxes: '100% Konten unik' (100% Unique Content) and '0% Konten yang dijiplak' (0% Copied Content). A blue progress bar indicates 'COMPLETED 100%'. Below this, there are tabs for 'Kalimat hasil bijak' (Selected sentence results) and 'URL yang Cocok' (Matching URLs). The 'Kalimat hasil bijak' tab is active, showing a list of 10 unique sentences. To the right of the list is a button labeled 'Buat Laporan Plagiarisme' (Create Plagiarism Report). Below the list is a section titled 'Kata Kunci Kata Kepadatan' (Keyword Word Density) with three options: '1- word', '2- words', and '3- words'. The '2- words' option is selected. A 'Feedback' button is visible on the left side of the page. The Windows taskbar at the bottom shows the time as 6:35 AM on 7/7/2022.

100% Konten unik

0% Konten yang dijiplak

✓ COMPLETED 100%

Kalimat hasil bijak URL yang Cocok

Buat Laporan Plagiarisme

unik	
unik	Peningkatan penggunaan kartu kredit terjadi secara signifikan seiring dengan perkembangan...
unik	Namun peningkatan penggunaan tersebut tak terlepas dari kenaikan terjadinya penipuan...
unik	Hal ini tentunya sangat merugikan perusahaan pengelola keuangan setiap tahun.
unik	Dengan perkembangan teknologi yang ada maka kasus penipuan transaksi pada kartu kredit...
unik	Pada penelitian ini penulis telah melakukan analisa pada...
unik	penggunaan kartu kredit untuk meminimalisir adanya penipuan pada kartu kredit.
unik	Selain itu juga pada penelitian ini digunakan beberapa metode yang digunakan untuk...
unik	yang baik, sehingga dapat diketahui metode mana yang paling baik digunakan untuk men...
unik	kredit.
unik	Dengan menggunakan empat metodologi penelitian yaitu Decision tree, Random forest,
unik	Maka dari hasil penelitian ini dapat disimpulkan bahwa metode Xgboost memiliki nilai...
unik	kredit.

Kata Kunci Kata Kepadatan

1- word 2- words 3- words

Feedback

Type here to search

6:35 AM 7/7/2022

SURAT PERNYATAAN
TIDAK MELAKUKAN PLAGIARISME

Yang bertanda tangan dibawah ini:

Nama : Dinda Anik Masruro

NPM : 1184003

Program Studi : D4 Teknik Informatika

Judul : Review Perbandingan Performansi Metode Decision Tree, Random Forest, Logistic Regression, dan Xgboost pada Dataset Penipuan Kartu Kredit

Menyatakan bahwa:

1. Program Internship II saya ini adalah asli dan belum pernah diajukan untuk memenuhi kelulusan matakuliah Internship pada Program Studi D4 Teknik Informatika baik di Politeknik Pos Indonesia maupun di Perguruan Tinggi lainnya.
2. Program Internship II ini adalah murni gagasan, rumusan, dan penelitian saya sendiri tanpa bantuan pihak lain, kecuali arahan pembimbing.
3. Dalam Program Internship II ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan-penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya ini, serta sanksi lainnya sesuai dengan norma yang berlaku diperguruan tinggi lain.

Bandung, 7 Juli 2022
Yang Membuat Pernyataan,
Materai
6000

Dinda Anik Masruro.....
NPM:1184003