PENGEMBANGAN APLIKASI *BANK ACCOUNT FRAUD DETECTION* DENGAN MENGGUNAKAN ALGORITMA XGBOOST

Arfian Haris Prayoga, Apriade Voutama
Sistem Informasi, Universitas Singaperbangsa Karawang
Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat, Indonesia
arfianharis7@gmail.com

ABSTRAK

Penelitian ini menyelidiki pengembangan aplikasi "*Bank account fraud detection*" yang menggunakan algoritma *XGBoost* untuk menemukan penipuan dalam transaksi rekening bank. Pendahuluan membahas bagaimana teknologi informasi dan komunikasi (TIK) sangat penting bagi kehidupan manusia, khususnya di sektor keuangan. Kejahatan siber, yang mencakup *skimming* dan *social engineering* dalam industri jasa keuangan, adalah masalah utama. Tujuan penelitian adalah menciptakan model yang paling efektif untuk mendeteksi penipuan transaksi keuangan yang dilakukan secara elektronik. Proses pengumpulan data, eksplorasi data, modeling dengan algoritma *XGBoost*, evaluasi model, dan *deployment* aplikasi adalah bagian dari metodologi penelitian. Hasil penelitian menunjukkan bahwa *XGBoost Classifier* memiliki akurasi mendeteksi penipuan tertinggi dengan 95.67%. Dengan bantuan aplikasi, manajer bank dapat dengan cepat menemukan transaksi mencurigakan dan mengambil tindakan pencegahan, meningkatkan keamanan rekening bank dan melindungi data nasabah.

Kata kunci : *Machine learning, Bank account fraud detection,XGBoost,Cyber crime, Artificial intelligence, Big data*

1. PENDAHULUAN

Buku ini membahas perkembangan TIK yang semakin mendominasi kehidupan manusia, termasuk dalam IJK. Pemahaman masyarakat terhadap teknologi keuangan beragam, dari yang sudah mahir hingga yang baru mengenalnya. Penerapan TIK di bidang keuangan terbukti memberikan kemudahan dan kenyamanan, namun juga memiliki konsekuensi negatif. Penyedia jasa teknologi keuangan (*cyber security*) harus mampu memitigasi konsekuensi negatif ini untuk melindungi konsumen, baik Lembaga Jasa Keuangan (LJK) maupun konsumen[1].

Layanan perbankan terus berkembang pesat untuk memudahkan akses bagi nasabah. Salah satu contohnya adalah penggunaan Anjungan Tunai Mandiri (ATM) yang menggantikan fungsi kasir konvensional. ATM memungkinkan nasabah untuk melakukan berbagai kegiatan transaksi seperti tarik tunai maupun fungsi kasir lainnya[2].

Perbankan elektronik merupakan terobosan baru di bidang perbankan yang disukai banyak orang. Hal ini didorong oleh meningkatnya pengguna internet di Indonesia, yang menurut Kominfo mencapai 202,6 juta pengguna pada tahun 2021, meningkat 11% dari tahun sebelumnya[2].

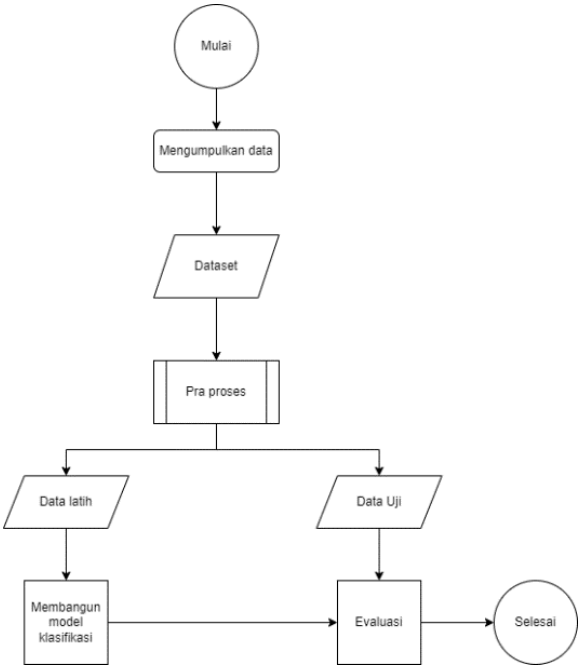
Skimming dan *social engineering* adalah contoh tindak kejahatan siber di sektor jasa keuangan dan perbankan. Teknik penipuan, penggelapan, dan pencurian digunakan untuk melakukan kedua jenis kejahatan siber ini. Hakim sering menggunakan pasal-pasal dalam KUHP, seperti Pasal 64 tentang perbuatan berulang, karena pelaku menjalankan aksinya berulang kali, Pasal 362 tentang Pencurian, Pasal 363 tentang pencurian yang dilakukan satu atau lebih dari satu orang, dan Pasal 378 tentang penggelapan, saat menerapkan hukum pidana terhadap pelaku kejahatan siber di sektor jasa keuangan dan perbankan[3],[15].

Kejahatan siber (*cyber crime*) biasanya terdiri dari pencurian identitas, spionase siber, pemerasan siber, pencurian data perusahaan, dan carding. Ini dilakukan oleh individu yang sangat mahir dalam peretasan dan menggunakan komputer sebagai sarana untuk melakukan tindak kriminal. Namun, menurut UU No. 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik, kejahatan siber adalah tindak pidana yang berkaitan dengan aktivitas ilegal, gangguan (interferens), menyediakan perbuatan yang tidak diperbolehkan, dan pemalsuan informasi atau dokumen elektronik[3].

Tujuan penelitian ini adalah untuk menemukan model yang paling cocok untuk kemudian diterapkan ke dalam sistem pembayaran perbankan dengan tujuan mendeteksi penipuan transaksi keuangan secara elektronik dan mencegah penipuan yang akan merugikan perbankan dan nasabah[4].

2. TINJAUAN PUSTAKA

Secara umum, empat langkah digunakan dalam proses penelitian ini: pengumpulan data, pra-proses data, pembuatan model klasifikasi, dan evaluasi model (gambar 1). Processor Intel® CoreTM i5-8250OU CPU @ 1.60 GHz yang dilengkapi dengan 12 GB RAM akan menjalankan proses secara keseluruhan.

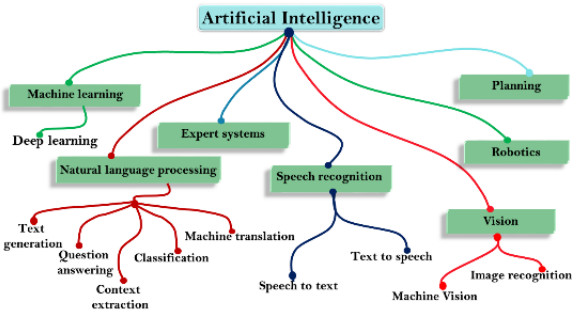


Gambar 1. Alur proses penelitian

2.1. Artificial Intelligence

Artificial Intelligence (juga dikenal sebagai AI atau hanya AI) adalah kecerdasan buatan yang ditambahkan ke sistem yang dapat diatur secara ilmiah.Kecerdasan buatan adalah "kemampuan system untuk menafsirkan data eksternal dengan benar, untuk belajar dari data tersebut, dan menggunakan pembelajaran tersebut untuk mencapai tujuan dan tugas tertentu melalui adaptasi yang fleksibel", menurut definisi Andreas Kaplan dan Michael Haenlein.Sistem seperti ini biasanya disebut komputer.Kecerdasan adalah kemampuan mesin atau komputer untuk melakukan tugas yang sama dengan yang dapat dilakukan oleh manusia.Sistem pakar, permainan komputer (games), logika fuzzy, jaringan saraf tiruan, dan robotika adalah beberapa bidang yang menggunakan kecerdasan buatan [5],[13].

Mesin pintar atau cerdas akan secara bertahap menggantikan dan meningkatkan kemampuan manusia di berbagai bidang. Kecerdasan yang ditunjukkan oleh mesin bagian dari ilmu komputer yang dikenal sebagai kecerdasan buatan (AI)[6].



Gambar 2. Artificial intelligence

2.2. Data Mining

Data mining adalah proses mendapatkan pola penting, keterkaitan, dan kecenderungan dalam sekumpulan data besar yang disimpan dalam penyimpanan dengan memeriksanya dengan memakai metode pengenalan pola metode statistik dan matematika (2005). Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, seperti deskripsi, estimasi, prediksi, klasifikasi, pengklusteran, dan asosiasi [7],[14].

2.3. Big Data

Istilah "big data", yang terus berkembang, mengacu pada jumlah besar data terstruktur, semi terstruktur, dan tidak terstruktur yang berpotensi ditambang agar mendapatkan informasi. Dari definisi ini, dapat disimpulkan bahwa data yang dikumpulkan harus bervolume besar dan dapat digunakan untuk menghasilkan sebuah atau beberapa informasi di masa depan. Sebenarnya, pengolahan lebih lanjut dapat menghasilkan keluaran sebagai pendukung keputusan[7].

2.4. Clustering

Pembelajaran mesin menggunakan metode clustering untuk mengumpulkan data menjadi beberapa kumpulan atau klaster yang sesuai. Tujuan dari clustering adalah untuk menemukan struktur yang tidak diketahui sebelumnya dalam data. Terdapat banyak algoritma clustering yang tersedia, termasuk K-Means, Hierarchical Clustering, dan DBSCAN. K-Means merupakan algoritma yang paling popular dan sederhana, yang mengumpulkan data ke dalam jumlah klaster yang ditentukan sebelumnya. Hierarchical Clustering mengumpulkan data ke dalam hierarki klaster, sedangkan DBSCAN mengumpulkan data berdasarkan jarak spasial[16].

2.5. Supervised Learning

Algoritma pembelajaran yang diawasi adalah algoritma yang bergantung pada data input berlabel untuk mempelajari fungsi dan menghasilkan output yang sesuai ketika diberi data baru tanpa label. Algoritma XGBoost percaya bahwa objek serupa ada di dekatnya[8].

2.6. XGBoost

XGBoost adalah algoritma yang ditingkatkan yang didasarkan pada pilihan pohon peningkatan gradient dan memiliki kemampuan untuk membangun pohon peningkatan yang efektif dan berfungsi secara bersamaan. Pada dasarnya, XGBoost adalah metode ensemble yang didasarkan pada pohon peningkatan gradient, dan merupakan salah satu teknik pembelajaran mesin yang dirancang untuk menangani masalah regresi dan klasifikasi. Berdasarkan Gradient Boosting Decision Tree (GBDT), XGBoost adalah solusi terbaik. Di dalam pohon regresi, node bagian dalam menunjukkan nilai nilai tes atribut, dan node cabang dengan skor menunjukkan keputusan[10].

2.7. Jupyter Notebook

Jupyter adalah aplikasi web gratis yang memungkinkan Anda membuat dan membagikan dokumen dengan teks, kode, hasil hitungan, dan visualisasi. Jupyter adalah singkatan dari tiga bahasa pemrograman Julia (Ju), Python (Py), dan R, yang sangat penting bagi seorang ilmuwan data. Jupyter membantu membuat narasi komputasi yang menjelaskan makna data dan memberikan insight[9],[12].

3. METODE PENELITIAN

Pada penelitian ini penulis menggunakan metode XGBoost. Untuk metode ini, fungsi objektif yang berguna diperlukan untuk menilai seberapa baik model yang dihasilkan sesuai dengan data latihan. Fungsi objektif terdiri dari dua komponen, yaitu nilai regularisasi dan nilai pelatihan yang hilang, seperti yang ditunjukkan dalam persamaan (1) berikut

$$obj(\theta) = L(\theta) + \Omega(\theta) \quad (1)$$

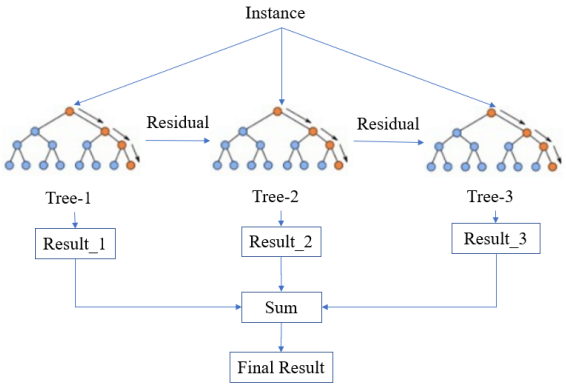
Dimana L adalah fungsi pelatihan yang hilang, dan Ω adalah fungsi regularisasi, dan θ adalah parameter model terkait. L dapat ditulis seperti pada persamaan (2) sebagai berikut.

$$L(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) \quad (2)$$

Di sini, n adalah jumlah nilai iterasi model, *yi* adalah nilai data sebenarnya yang dianggap benar dan *yi^*adalah hasil nilai prediksi model[11].

Metodologi *XGBoost* (*eXtreme Gradient Boosting*) adalah metode penelitian yang digunakan penulis. Metodologi *XGBoost* terdiri dari beberapa tahapan utama yaitu :

a. Inisialisasi
b. Iterasi
c. Regularisasi
d. Ensemble
e. Optimasi
f. Prediksi



Gambar 3. Proses metode XGBoost

Pada penelitian ini terdiri dari beberapa tahapan yaitu :

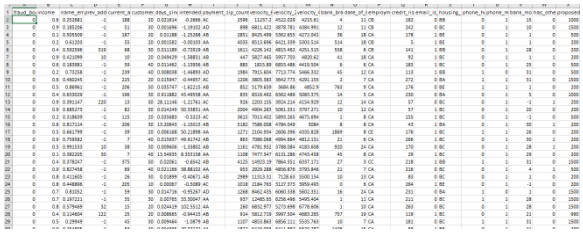
a. Data Acquisition
b. Data Exploration
c. Modeling
d. Evaluation
e. Deployment

4. HASIL DAN PEMBAHASAN

Penelitian ini berfokus pada menyelesaikan permasalahan ini dengan menyediakan alat yang dapat mendeteksi transaksi penipuan secara otomatis dan akurat. Dengan menggunakan model Machine Learning dan algoritma XGBoost, aplikasi ini memungkinkan bank untuk mengidentifikasi transaksi yang mencurigakan dan mengambil tindakan pencegahan dengan cepat. Dengan demikian, proyek ini bertujuan untuk meningkatkan keamanan rekening bank, melindungi data pribadi nasabah, dan meminimalkan risiko penipuan, menjadikannya solusi yang sangat relevan dalam industri perbankan yang terus berkembang.

4.1. Data Acquisition

Dalam proyek pembuatan aplikasi/dashboard "Bank Account Fraud Detection," penulis menggunakan jenis data supervised learning. Terdapat sekitar 1.000.000 baris data yang terdiri dari transaksi perbankan dengan kategori fraud dan non-fraud.

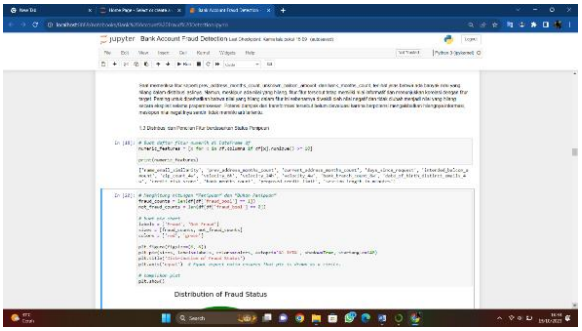


Gambar 4. Dataset bank account fraud detection

Sumber data berasal dari Kaggle, sebuah platform sumber terbuka yang menyediakan berbagai dataset untuk analisis dan proyek data science. Data ini sangat berharga dalam pengembangan model Machine Learning untuk mendeteksi penipuan rekening bank. Proses pengambilan data melibatkan pengunduhan dataset dari Kaggle, diikuti dengan proses pemrosesan, eksplorasi, dan pengolahan sebelum digunakan dalam pembuatan model. Kualitas dan jumlah data yang besar sangat penting untuk memastikan akurasi dan kinerja model deteksi penipuan.

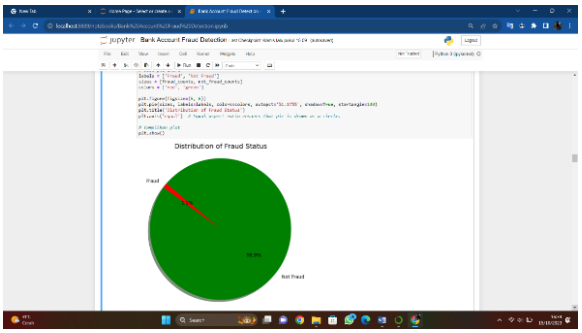
4.2. Data Exploration

Tahap data Exploration membantu dalam melihat ringkasan statistik, seperti mean, standar deviasi, nilai minimum, dan maksimum, yang merupakan informasi penting untuk pemodelan dan analisis. Pada gambar 3 ditampilkan tentang distribusi dan pencilan fitur berdasarkan status penipuan.



Gambar 5. Distribusi dan Pencilan Fitur

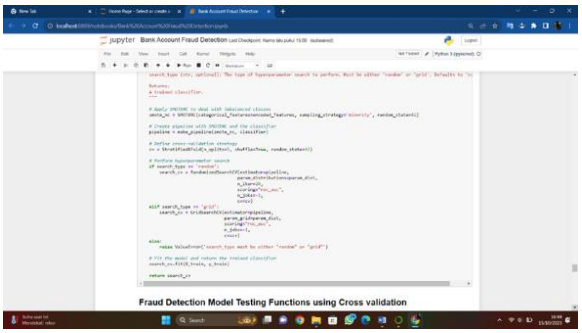
Data ini memberikan gambaran tentang perbandingan antara jumlah transaksi penipuan dan non-penipuan. Ini membantu dalam memahami apakah ada pola khusus dalam keberadaan nilai yang hilang di antara transaksi penipuan dan non-penipuan. Distribusi dan proporsi status penipuan dan non-penipuan diilustrasikan dalam bentuk diagram pie. Diagram ini memvisualisasikan seberapa besar proporsi transaksi yang merupakan penipuan dan seberapa besar yang bukan penipuan.



Gambar 6. Visualisasi Pie Chart Fraud dan Not Fraud

Setelah data tersebut divisualisasikan maka proses berlanjut pada tahapan modeling dengan menggunakan model training functions menggunakan teknik SMOTE (Synthetic Minority Over-sampling Technique) dan Grid Search CV (Cross-Validation).

4.3. Modeling



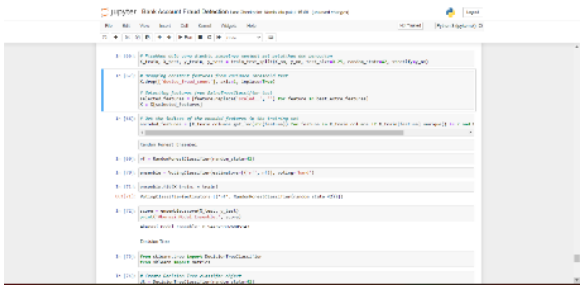
Gambar 7. Model Training Using SMOTE and GRID Search CV

Dalam pengembangan aplikasi "Bank Account Fraud Detection," beberapa algoritma dan pendekatan digunakan untuk mendeteksi penipuan dalam transaksi rekening bank. Salah satu pendekatan yang digunakan adalah model training functions menggunakan teknik Synthetic Minority Over sampling Technique dan Grid Search CrossValidation.

SMOTE digunakan untuk mengatasi ketidakseimbangan dataset, di mana jumlah transaksi penipuan jauh lebih sedikit dibandingkan dengan transaksi bukan penipuan, dengan membuat dataset sintetis yang menambah sampel-sampel minoritas (penipuan) berdasarkan sampel-sampel yang ada. Grid Search CV digunakan untuk mencari hyperparameter terbaik untuk algoritma yang digunakan, memungkinkan penentuan parameter-parameter yang mengoptimalkan kinerja model. Beberapa algoritma yang diuji mencakup Random Forest Ensemble, Decision Tree, XGB Classifier, dan LightGBM Classifier. Setelah tahapan modeling selesai maka proses berlanjut pada tahapan evaluasi algoritma dan model yang menampilkan hasil akurasi dari setiap model.

4.4. Evaluation

Dalam evaluasi algoritma dan model yang digunakan untuk mendeteksi penipuan dalam aplikasi "Bank Account Fraud Detection", beberapa metrik penting digunakan untuk mengukur kinerja masing-masing algoritma. Metrik yang digunakan termasuk akurasi dan metrik spesifik untuk masing-masing algoritma, seperti ensemble.score dan metrics.accuracy_score.

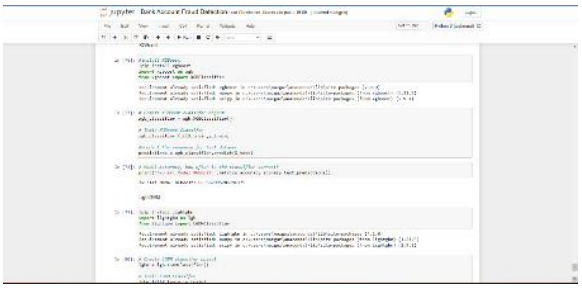


Gambar 8. Pengujian akurasi random forest ensemble

Algoritma Random Forest Ensemble memiliki hasil akurasi sekitar 0.9479 dengan menggunakan ensemble.score.

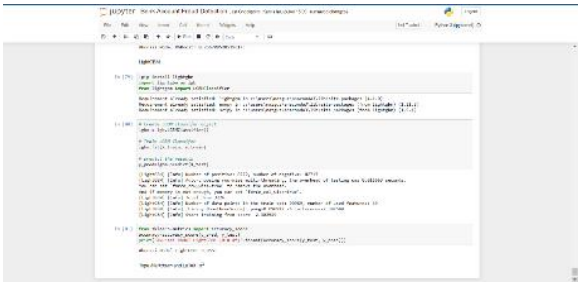
Gambar 9. Pengujian akurasi decision tree

Algoritma Decision Tree mencapai akurasi sekitar 0.9155 dengan menggunakan metrics.accuracy_score.



Gambar 10. Pengujian akurasi model XGBoost

Algoritma XGBoost Classifier memberikan akurasi tertinggi di antara algoritma yang diuji, mencapai 0.9567, dengan metrik accuracy_score.



Gambar 11. Pengujian akurasi model LightGBM

LightGBM juga menghasilkan akurasi yang baik, yaitu 0.9551 dengan metrik accuracy_score. Namun, XGBoost Classifier menonjol dengan akurasi yang lebih tinggi.

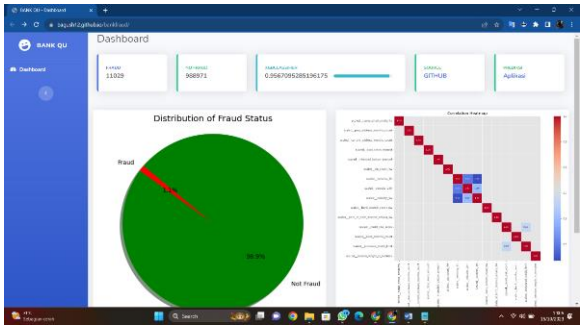
Hasil evaluasi akurasi model menunjukkan bahwa algoritma XGBoost Classifier adalah yang paling tepat untuk digunakan dalam aplikasi "Bank Account Fraud Detection" yang memiliki hasil akurasi 0.9567. Maka dari itu, algoritma XGBoost Classifier dipilih sebagai algoritma utama dalam aplikasi ini karena mampu memberikan prediksi yang paling akurat dalam mendeteksi penipuan dalam transaksi rekening bank.

4.5. Deployment

Tujuan dari deployment dalam konteks aplikasi "Bank Account Fraud Detection" adalah memungkinkan manajer bank mengakses dan menggunakan alat deteksi penipuan dengan mudah dan efisien. Proses deployment bertujuan untuk membuat aplikasi dan dashboard ini fungsional dan dapat diakses oleh pengguna akhir. Proses deployment aplikasi "Bank Account Fraud Detection" terbagi menjadi dua tahap. Pertama, deployment dashboard menggunakan Bootstrap, kerangka kerja desain web yang membangun antarmuka pengguna interaktif dan responsif.

Pada tahap ini, dashboard yang berisi informasi penting seperti jumlah transaksi fraud dan non-fraud, akurasi model, sumber kode GitHub, dan aplikasi prediksi penipuan diintegrasikan ke dalam antarmuka web. Ini memungkinkan manajer bank memantau situasi penipuan dalam satu tampilan yang mudah dinavigasi. Kedua, deployment aplikasi prediksi penipuan menggunakan Streamlit, kerangka kerja Python yang memungkinkan pengembang membuat aplikasi data interaktif tanpa pengetahuan mendalam dalam desain antarmuka.

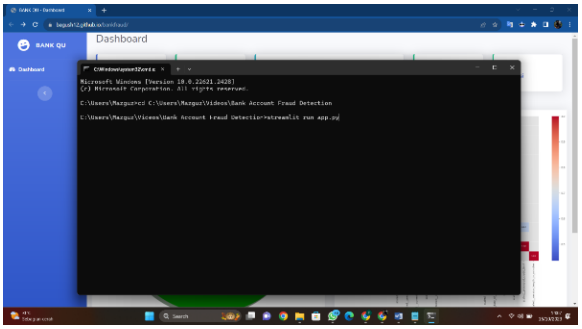
Aplikasi ini memungkinkan manajer bank memasukkan nilai-nilai atribut tertentu, seperti pendapatan, kemiripan nama-email, jumlah bulan alamat sebelumnya, dan jumlah bulan alamat saat ini, untuk memprediksi apakah suatu transaksi mungkin merupakan penipuan. Keseluruhan proses deployment bertujuan untuk memberikan kemudahan akses dan penggunaan alat deteksi penipuan kepada manajer bank. Dengan dashboard dan aplikasi yang telah dideploy, mereka dapat dengan cepat dan efisien mengidentifikasi serta mengatasi transaksi penipuan, menjaga keamanan rekening bank, dan melindungi data pribadi nasabah.



Gambar 12. Dashboard Bank Account Fraud Detection

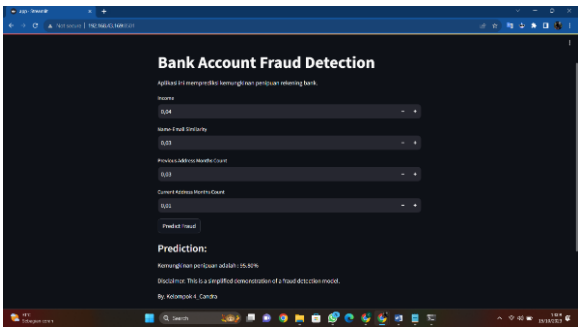
Pada dashboard ini, terdapat sejumlah card yang memberikan informasi penting terkait dengan penanganan penipuan rekening bank. Card pertama menunjukkan jumlah kasus penipuan, yang mencapai 11,029 transaksi. Di sisi lain, card kedua menampilkan jumlah transaksi yang sah, yang mencapai 988,971. Hal ini memperlihatkan bahwa jumlah transaksi yang sah jauh lebih tinggi dibandingkan dengan penipuan. Card ketiga menyajikan metrik akurasi dari model XGBoost Classifier yang digunakan untuk mendeteksi penipuan. Akurasi model ini mencapai 95.67%, yang menandakan bahwa model mampu memprediksi dengan sangat baik apakah sebuah transaksi adalah penipuan atau bukan. Card keempat adalah sumber model, dan Anda dapat mengakses detail model dan proses permodelan di GitHub melalui tautan yang disediakan. Ini memungkinkan untuk memahami bagaimana model dibangun dan melihat dokumentasi terkaitnya. Terakhir, card kelima merinci penggunaan model. Aplikasi ini memungkinkan pengguna untuk melakukan prediksi penipuan dengan model yang telah dibangun. Ini adalah langkah penting dalam memastikan keamanan rekening bank dan melindungi data pribadi nasabah.

Selain itu, terdapat juga diagram pie yang mengilustrasikan proporsi transaksi fraud dan not fraud. Dari diagram pie, kita dapat melihat bahwa transaksi yang merupakan penipuan hanya sekitar 1.1% dari total transaksi, sementara transaksi yang sah mencapai 98.9%. Terakhir, ada correlation heatmap yang memvisualisasikan hubungan antara variabel dalam dataset. Ini membantu dalam pemahaman faktor-faktor yang berkaitan dengan penipuan dan dapat digunakan untuk perbaikan lebih lanjut dalam mendeteksi penipuan.



Gambar 13. Run Streamlit

Pada Gambar 13 Untuk menjalankan aplikasi prediksi fraud dan membuat localhost terbuka, langkah pertama yang perlu dilakukan adalah menjalankan perintah "streamlit run app.py". Dengan perintah ini, aplikasi akan dijalankan menggunakan framework Streamlit, yang memungkinkan Anda untuk membuat antarmuka pengguna yang interaktif untuk memprediksi penipuan rekening bank. Setelah perintah dijalankan, Anda dapat membuka peramban web dan mengakses localhost untuk mengakses aplikasi ini. Dengan begitu, Anda dapat dengan mudah melakukan prediksi terhadap transaksi yang masuk dan memastikan keamanan rekening bank dengan cepat dan efisien. Ini adalah langkah penting dalam mendeteksi potensi penipuan dan menjaga keamanan data pribadi nasabah.



Gambar 14. Aplikasi Prediksi Fraud

Pada gambar 14 Dalam aplikasi prediksi fraud, langkah pertama adalah klik aplikasi pada dashboard untuk menuju ke aplikasi prediksi penipuan. Setelah itu, pengguna diminta untuk memasukkan nilai dari beberapa atribut yang diperlukan, termasuk Income (Pendapatan), Name-Email Similarity (Kemiripan Nama-Email), Previous Address Months Count (Jumlah Bulan Alamat Sebelumnya), dan Current Address Months Count (Jumlah Bulan Alamat Sekarang). Sebagai contoh, jika pengguna memasukkan nilai pendapatan sebesar 0.04, nilai kemiripan nama-email sebesar 0.03, jumlah bulan alamat sebelumnya sebesar 0.03, dan jumlah bulan alamat sekarang sebesar 0.01, maka hasil kemungkinan penipuan akan ditampilkan sebagai 95.80%. Dengan aplikasi ini, pengguna dapat dengan mudah memasukkan nilai-nilai ini dan mendapatkan prediksi terkait dengan apakah suatu transaksi dapat dianggap sebagai penipuan atau tidak. Ini merupakan alat yang sangat berguna untuk mengidentifikasi potensi penipuan dengan cepat dan efisien.

5. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian ini adalah berhasilnya pengembangan aplikasi "Bank Account Fraud Detection" menggunakan algoritma XGBoost dalam mendeteksi penipuan dalam transaksi rekening bank. Dengan hasil akurasi mencapai 0,9567 pada algoritma XGBoost. Aplikasi ini memberikan solusi efektif dalam meningkatkan keamanan rekening bank, melindungi data pribadi nasabah, dan meminimalkan risiko penipuan. Saran untuk penelitian selanjutnya adalah memperluas cakupan data yang digunakan untuk melatih model agar lebih representatif, serta mengintegrasikan teknologi keamanan yang lebih canggih untuk menghadapi tantangan kejahatan siber yang semakin kompleks.

DAFTAR PUSTAKA

[1] R. Akyuwen, Lebih Mengenal Digital Banking. 2020. [Online]. Available: http://repository.upstegal.ac.id/3051/

[2] C. H. Ratulangi, D. A. S. Wahongan, and F. R. Mewengkang, “Tindak Pidana Cyber Crime Dalam Kegiatan Perbankan,” Lex Priv., vol. IX, no. 5, pp. 179–187, 2021.

[3] W. P. W. Widayanti, “TINDAK PIDANA PENCURIAN DATA NASABAH DALAM BIDANG PERBANKAN SEBAGAI CYBER CRIME Putri,” no. 8.5.2017, pp. 2003–2005, 2022.

[4] F. Zamachsari and N. Puspitasari, “Penerapan Deep Learning dalam Deteksi Penipuan Transaksi Keuangan Secara Elektronik,” J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 5, no. 2, pp. 203–212, 2021, doi: 10.29207/resti.v5i2.2952.

[5] M. Siahaan, C. Harsana Jasa, K. Anderson, M. V. Rosiana, S. Lim, and W. Yudianto, “Penerapan Artificial Intelligence (AI) Terhadap Seorang Penyandang Disabilitas Tunanetra,” 2020.

[6] A. Roihan, P. A. Sunarya, and A. S. Rafika, “Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper,” IJCIT (Indonesian Journal on Computer and Information

Technology), vol. 5, no. 1, May 2020, doi: 10.31294/ijcit.v5i1.7951.

[7] A. O. P. Dewi, "Big Data di Perpustakaan dengan Memanfaatkan Data Mining," Anuva: Jurnal Kajian Budaya, Perpustakaan, dan Informasi, vol. 4, no. 2, pp. 223–230, Jun. 2020, doi: 10.14710/anuva.4.2.223-230.

[8] K. Kristiawan, D. D. Somali, T. A. Linggan jaya, and A. Widjaja, "Deteksi Buah Menggunakan Supervised Learning dan Ekstraksi Fitur untuk Pemeriksa Harga," Jurnal Teknik Informatika dan Sistem Informasi, vol. 6, no. 3, Dec. 2020, doi: 10.28932/jutisi.v6i3.3029.

[9] S. Junaidi, M. Devegi, and H. Kurniawan, "Pelatihan Pengolahan dan Visualisasi Data Penduduk menggunakan Python," ADMA : Jurnal Pengabdian dan Pemberdayaan Masyarakat, vol. 4, no. 1, pp. 151–162, Jul. 2023, doi: 10.30812/adma.v4i1.2963.

[10] I. M. Karo Karo, "Implementasi Metode XGBoost dan Feature Important untuk Klasifikasi pada Kebakaran Hutan dan Lahan," Journal of Software Engineering, Information and Communication Technology (SEICT), vol. 1, no. 1, pp. 11–18, Mar. 2022, doi: 10.17509/seict.v1i1.29347.

[11] S. E. Herni Yulianti, Oni Soesanto, and Yuana Sukmawaty, "Penerapan Metode Extreme Gradient Boosting (XGBOOST) pada Klasifikasi Nasabah Kartu Kredit," Journal of Mathematics: Theory and Applications, pp. 21–26, Aug. 2022, doi: 10.31605/jomta.v4i1.1792.

[12] I. A. Ashari, A. Wirasto, D. Nugroho Triwibowo, and P. Purwono, "Implementasi Market Basket Analysis dengan Algoritma Apriori untuk Analisis Pendapatan Usaha Retail," MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput., vol. 21, no. 3, pp. 701–709, 2022, doi: 10.30812/matrik.v21i3.1439.

[13] A. R. Bayu Setia, "Penerapan Logika Fuzzy pada Sistem Cerdas," J. Sist. Cerdas, vol. 02, pp. 61–66, 2019.

[14] K. Erwansyah, B. Andika, and R. Gunawan, "Implementasi Data Mining Menggunakan Asosiasi Dengan Algoritma Apriori Untuk Mendapatkan Pola Rekomendasi Belanja Produk Pada Toko Avis Mobile," J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD), vol. 4, no. 1, p. 148, 2021, doi: 10.53513/jsk.v4i1.2628.

[15] T. Rompi and H. S. Muaja, "Tindak Kejahatan Siber Di Sektor Jasa Keuangan Dan Perbankan," Lex Priv., vol. 9, no. 4, pp. 183–192, 2021, [Online]. Available: https://ejournal.unsrat.ac.id/index.php/lexprivatum/article/view/33358

[16] M. Rafi Nahjan, Nono Heryana, and Apriade Voutama, "Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell," JATI (Jurnal Mhs. Tek. Inform., vol. 7, no. 1, pp. 101–104, 2023, doi: 10.36040/jati.v7i1.6094.