

BAB II KERANGKA TEORITIS

A. Landasan Teori

Penelitian ini berlandaskan pada konsep teknologi One-Time Password (OTP) dan aplikasi Machine Learning (ML) dalam sistem keamanan. OTP, sebuah mekanisme autentikasi yang menghasilkan kata sandi sekali pakai, memainkan peran penting dalam menjamin keamanan informasi pengguna. Namun, penipuan OTP menjadi masalah yang masih harus diatasi. Dalam upaya untuk memecahkan masalah ini, penelitian ini memanfaatkan prinsip Machine Learning, khususnya algoritma Random Forest, untuk mendeteksi pola penipuan dalam data transaksi OTP.

1. One-Time Password (OTP)

OTP dihasilkan oleh sistem dan dikirimkan kepada pengguna untuk setiap transaksi atau sesi login. Kode ini hanya berlaku untuk satu kali penggunaan dan berubah setiap kali transaksi atau sesi login baru dilakukan. Menurut (Yeluri & Castro-Leon, 2014, p. 155) OTP adalah pengenal tunggal yang tidak bisa ditebak atau diulang oleh penyerang. Mekanisme ini menambah lapisan keamanan tambahan, dengan memastikan bahwa hanya pengguna yang sah yang menerima dan menggunakan kode OTP.

2. Fraud atau Penipuan

Definisi singkat penipuan diuraikan dalam *Black's Law Dictionary* menjelaskan bahwa tindakan penipuan atau ke tidak jujur yang disengaja yang dilakukan oleh satu atau lebih individu, umumnya untuk keuntungan finansial. Adapun penipuan pada perbankan menurut Peraturan Otoritas Jasa Keuangan Republik Indonesia Nomor 39 /POJK.03/2019 tentang Penerapan Strategi Anti *Fraud* Bagi Bank Umum Bab 1 pasal 1, *fraud* adalah tindakan penyimpangan atau pembiaran yang sengaja dilakukan untuk mengelabui, menipu, atau memanipulasi Bank, nasabah, atau pihak lain, yang terjadi di lingkungan Bank dan/atau menggunakan sarana Bank sehingga mengakibatkan Bank, nasabah, atau pihak lain menderita kerugian dan/atau pelaku *fraud* memperoleh keuntungan keuangan baik secara langsung maupun tidak langsung. Dari berbagai pengertian yang ada, jadi dapat dikatakan bahwa Penipuan adalah tindakan penyimpangan yang dilakukan untuk memperoleh keuntungan dari tindak kecurangan yang merugikan pihak lain.

3. Machine Learning

Machine learning merupakan cabang dari ilmu komputer yang dikhususkan untuk pengembangan algoritma yang mampu belajar dari dan membuat prediksi berdasarkan data. Teknik ini melibatkan pemahaman dan penerapan algoritma yang belajar dari kumpulan contoh fenomena (Andriy Burkov, 2019, p. 3). Contoh-

contoh ini bisa berasal dari alam, dibuat secara manual oleh manusia, atau dihasilkan oleh algoritma lain.

Machine learning dalam implementasinya memiliki berbagai jenis pendekatan. Menurut (Murphy, 2012, p. 2) machine learning memiliki tiga jenis pendekatan, yaitu:

- (a) supervised learning adalah pendekatan untuk mempelajari pemetaan dari input (x) ke output (y);
- (b) unsupervised learning adalah pendekatan untuk menemukan pola menarik dalam data;
- (c) reinforcement learning adalah pendekatan untuk belajar bagaimana bertindak atau berperilaku ketika diberikan sinyal penghargaan atau hukuman yang sesekali.

Pemahaman terhadap ketiga pendekatan dalam machine learning ini sangat penting dalam menentukan strategi yang paling efektif dalam menangani masalah.

4. Random Forest

Random Forest adalah metode dalam machine learning yang merupakan bentuk lanjutan dari pohon keputusan (decision trees). Menurut (Andriy Burkov, 2019, p. 9) Random Forest menerapkan dua paradigma utama dari ensemble learning, yaitu:

- (a) bagging : proses membuat beberapa sub set dari data set asli dengan prosedur bootstrap, yaitu dengan mengambil sampel secara acak dengan penggantian;
- (b) boosting : proses melatih model secara berurutan yang setiap model berikutnya mencoba untuk memperbaiki kesalahan yang dibuat oleh model sebelumnya, dengan memberikan bobot lebih kepada contoh-contoh yang diprediksi salah oleh model sebelumnya.

Dengan demikian, Random Forest menerapkan konsep bagging dan boosting untuk membentuk serangkaian "pohon" keputusan yang bekerja bersama-sama untuk menghasilkan prediksi yang lebih akurat dan stabil.

Random Forest memiliki batasan dalam pemrosesan data. Secara khusus, Random Forest hanya dapat memproses data numerik. Jika terdapat data dalam bentuk non-numerik, seperti kategori atau label, diperlukan sebuah teknik yang disebut One-Hot Encoding. One-Hot Encoding adalah teknik yang digunakan untuk mengubah variabel kategori atau label menjadi bentuk numerik. Pada dasarnya, setiap nilai kategori diubah menjadi vektor biner di mana setiap nilai kategori memiliki kolom sendiri dan diwakili oleh nilai 0 atau 1. Oleh karena itu, penerapan One-Hot Encoding menjadi langkah krusial dalam pengolahan data

non-numerik untuk memastikan keefektifan model Random Forest. Dengan demikian, transformasi data ini memungkinkan Random Forest untuk menangani beragam jenis data, meningkatkan fleksibilitas dan kegunaannya dalam berbagai aplikasi machine learning.

5. Unified Model Language (UML)

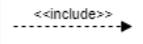
Dalam perancangan sistem, dibutuhkan suatu pemodelan yang dapat menggambarkan sistem. UML merupakan bahasa visual yang digunakan untuk komunikasi dan pemodelan sebuah sistem dengan menggunakan diagram dan teks-teks pendukung (M. Shalahuddin & Rosa A.S, 2015, p. 137). Kategori UML yang digunakan dalam perancangan sistem deteksi penipuan OTP terdiri dari activity diagram, usecase diagram, class diagram dan sequence diagram.

a. Use Case Diagram

Use case diagram merupakan pemodelan untuk tingkah laku sistem informasi yang akan dibuat (M. Shalahuddin & Rosa A.S, 2015, p. 155). Use case diagram mendeskripsikan interaksi antara satu aktor atau lebih dengan sistem informasi yang akan dibuat. Dalam pemodelan usecase diagram terdapat simbol-simbol, setiap simbol yang digunakan ini memiliki deskripsi. Berikut ini adalah simbol-simbol yang ada pada usecase diagram :

Tabel 2. Simbol use case diagram

Simbol	Nama	Keterangan
	Actor	Menggambarkan seseorang yang berinteraksi dengan sistem
	Use Case	Menggambarkan apa yang dikerjakan sistem, biasanya menggunakan kata kerja
	Association	Digunakan untuk menghubungkan antara aktor dengan usecase
	Generalization	Digunakan untuk hubungan generalisasi atau spesialisasi antara dua buah use case

Simbol	Nama	Keterangan
	<i>Include</i>	Digunakan untuk menghubungkan <i>usecase</i> dengan <i>usecase</i> yang prosesnya diharuskan atau menjadi syarat yang harus dipenuhi
	<i>Extend</i>	Digunakan sebagai perluasan dari <i>usecase</i> lain jika kondisi terpenuhi

Sumber : (M. Shalahuddin & Rosa A.S, 2015)

b. Class Diagram

Class diagram menggambarkan struktur sistem dari pendefinisian kelas-kelas yang akan dibuat untuk membangun sistem (M. Shalahuddin & Rosa A.S, 2015, p. 141). Dalam pemodelan *class* diagram terdapat simbol-simbol, setiap simbol yang digunakan ini memiliki deskripsi. Berikut ini adalah simbol-simbol pada *class* diagram:

Tabel 3. Simbol class diagram

Simbol	Nama	Keterangan
	<i>Class</i>	Kelas pada struktur sistem.
	<i>Interface</i>	Elemen yang diperlukan elemen lain agar berfungsi dengan benar
	<i>Association</i>	Merupakan deskripsi koneksi antara <i>instance</i> kelas
	<i>Dependency</i>	Digunakan untuk menghubungkan antara dua elemen model
	<i>Realization</i>	Digunakan untuk menghubungkan antara spesifikasi dan implementasinya
	<i>Generalization</i>	Digunakan untuk menghubungkan antara deskripsi yang lebih spesifik dan lebih umum, digunakan untuk pewarisan deklarasi tipe polimorfik

Sumber: (M. Shalahuddin & Rosa A.S, 2015)

c. Activity Diagram

Activity diagram menggambarkan aliran kerja atau aktivitas dari sebuah sistem atau proses bisnis yang ada pada perangkat lunak (M. Shalahuddin & Rosa A.S, 2015, p. 161). Simbol pada *activity* diagram membantu visualisasi dan pengertian lebih mendalam tentang struktur dan alur proses yang sedang dianalisis. Berikut adalah simbol-simbol yang ada pada *activity* diagram :

Tabel 4. Simbol activity diagram

Simbol	Nama	Keterangan
	Start Point	Menunjukkan awal dari suatu proses
	Activities	Menunjukkan langkah aktivitas sistem
Simbol	Nama	Keterangan
	Fork	Digunakan untuk menunjukkan dua atau lebih langkah dalam aliran kerja berjalan secara serentak
	Join	Digunakan untuk menunjukkan akhir dari proses yang berjalan bersamaan
	Decision	Digunakan untuk menunjukkan sebuah flow masuk dan dua atau lebih flow keluar lalu menunjukkan aktivitas yang dapat dipilih
	Swimlane	Digunakan untuk mengelompokkan organisasi bisnis yang bertanggung jawab dengan aktivitas
	Final State	Digunakan untuk menunjukan status akhir yang dilakukan sistem

Sumber: (M. Shalahuddin & Rosa A.S, 2015)

d. Sequence Diagram

Sequence diagram menggambarkan tingkah laku objek pada use case dengan mendeskripsikan pesan yang dikirimkan dan diterima oleh objek (M. Shalahuddin & Rosa A.S, 2015, p. 165). Oleh karena itu untuk menggambarkan sequence diagram maka harus diketahui objek-objek yang terlibat dalam sebuah use case, dengan metode-metode yang dimiliki. Berikut ini adalah simbol-simbol yang ada pada sequence diagram:

Tabel 5. Simbol sequence diagram

Simbol	Nama	Keterangan
	Actor	Menggambarkan orang yang sedang melakukan interaksi dengan sistem

Simbol	Nama	Keterangan
	Boundary	Menggambarkan sebuah gambaran dari form
	Control	Menggambarkan penghubung antaraboundary dengan tabel
	Entity	Menggambarkan hubungan kegiatan yang akan dilakukan
	Life Line	Menggambarkan tempat mulai dengan berakhirnya sebuah message
	Message	Menggambarkan pengiriman pesan

Sumber: (M. Shalahuddin & Rosa A.S, 2015)

6. Business Process Model and Notation (BPMN)

Dalam menggambarkan alur kerja, diperlukan notasi untuk memudahkan dalam perancangan sistem maupun proses bisnis yang akan dilaksanakan. Menurut (Matjaz B. Juric & Kapil Pant, 2008, p. 111) *Business Process Model and Notation* (BPMN) merupakan notasi standar yang dirancang untuk memetakan sebagian besar jenis proses yang digunakan dalam suatu organisasi; sebagian besar organisasi menggunakan BPMN untuk membuat suatu proses repositori untuk semua prosesnya. Oleh karena itu, dengan BPMN berbagai proses dipetakan dan dikategorikan sesuai dengan penggunaan bisnis. Berikut adalah element inti pada BPMN:

Tabel 6. Simbol BPMN

Simbol	Nama	Keterangan
	Events	Menentukan bagaimana suatu organisasi akan merespons situasi melalui proses bisnisnya
	Activities	Menunjukkan serangkaian kegiatan yang dilakukan untuk mencapai tujuan yang memiliki titik awal dan titik akhir yang jelas
	Sub-Process	Menunjukkan bahwa aktivitas yang diberikan dijabarkan lebih lanjut dalam diagram terpisah

Simbol	Nama	Keterangan
	Sub-Process	Menunjukkan bahwa aktivitas yang diberikan dijabarkan lebih lanjut dalam diagram terpisah
	Association	Digunakan untuk menghubungkan antara informasi dan objek dengan menampilkan informasi teks tambahan pada suatu aktivitas
	Pool	Digunakan untuk mengelompokkan berbagai aktivitas yang dilakukan oleh partisipan yang berbeda dalam suatu proses
	Lane	digunakan untuk menunjukkan kategori lebih lanjut dengan membagi <i>pool</i> dalam beberapa bagian
	Gateway	Digunakan untuk mengontrol bagaimana alur interaksi saat bertemu dan menyimpang di dalam proses
	Sequence Flow	Digunakan untuk menunjukkan urutan aktivitas yang dilakukandalam proses
	Message Flow	Digunakan untuk menunjukkan aliran pesan yang menghubungkan dua objek yang dilakukan dalam proses, biasanya <i>message flow</i> dapat menghubungkan antar <i>pool</i> atau <i>swimlanes</i>

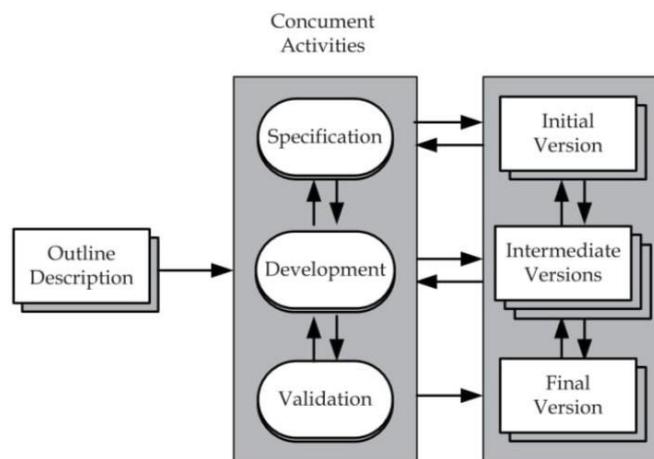
Sumber: (Matjaz B. Juric & Kapil Pant, 2008)

7. SDLC - Throwaway Prototyping

Menurut (David C. Yen, 2019, p. 3) System Development Life Cycle (SDLC) adalah serangkaian langkah yang menjadi dasar bagi sebagian besar metodologi analisis dan desain sistem. Langkah-langkah ini melibatkan berbagai tahapan mulai dari perencanaan, analisis, desain, hingga pengujian dan pemeliharaan. Dalam konteks ini, SDLC menjadi fondasi penting dalam pengembangan sistem yang efektif dan efisien. Oleh karena itu, pemahaman yang mendalam terhadap SDLC sangat penting untuk membangun sistem yang efisien dan efektif.

Menurut (Bidgoli Hossein, 2004, p. 136) *throwaway prototyping* dibuat untuk menunjukkan kemampuan antarmuka pengguna yang disimulasikan tetapi perangkat lunak tersebut tidak dimaksudkan untuk digunakan dalam sistem akhir. Prototipe jenis ini berfungsi sebagai alat bantu visualisasi awal bagi pengguna dan pengembang untuk memahami fitur dan fungsionalitas yang diinginkan dalam sistem. Meski demikian, perlu diingat bahwa prototipe ini tidak sepenuhnya mencerminkan kualitas dan performa sistem akhir.

Menurut (Supiandi, 2020, p. 16) prototyping merupakan model pengembangan evolusioner berdasarkan gagasan untuk mengembangkan implementasi awal, menunjukkannya kepada pengguna untuk komentar, dan secara bertahap meningkatkan hingga memiliki sistem yang memenuhi persyaratan; throwaway prototype merupakan salah satu jenis model pengembangan prototype dengan tujuan untuk memahami apa yang dibutuhkan pelanggan dan menentukan kebutuhan yang lebih baik untuk sistem; berikut adalah model prototyping;



Gambar 3 Model Prototyping

Sumber : (Supiandi, 2020)

Penelitian ini menggunakan SDLC throwaway prototyping karena memungkinkan pengujian fungsionalitas sistem secara lebih awal, sehingga potensi kesalahan atau kegagalan bisa terdeteksi dan diperbaiki lebih cepat. Akhirnya, throwaway prototyping juga dapat membantu mengurangi risiko dan ketidakpastian dalam proyek pengembangan sistem, karena memungkinkan validasi konsep dan desain sebelum pengembangan sistem yang sebenarnya dimulai.

8. Bahasa Pemrograman

Menurut (Dios Kurniawan, 2020, p. 30) python adalah *interpreted language*, yang artinya *source code* yang ditulis akan dijalankan per-baris oleh *python interpreter*; python adalah bahasa yang populer digunakan untuk *machine learning* dan *artificial intelligence*; python adalah bahasa *scripting* yang artinya bisa dieksekusitanpa perlu *compiler*; python memiliki banyak *library* tambahan untuk manipulasi dan visualisasi data. Python digunakan pada penelitian ini karena kemampuannya untuk memanipulasi data dan pengolahan data untuk analisis.

9. Web Server

Menurut (Rudyanto Arief, 2011), *web server* adalah program aplikasi yang memiliki kemampuan untuk menyimpan dokumen web. Semua dokumen web

yang ditulis menggunakan skrip sisi klien atau skrip sisi server disimpan di direktori root serverweb (root dokumen); contoh web server adalah Apache; server web ini biasanya tersedia dalam satu paket dengan server database di dalam aplikasi yang membantu proses pengembangan program.

10. Intranet

Menurut (Samuel Prakoso et al., 2007, p. 119) Intranet adalah sebuah kumpulan jaringan komputer lokal yang menggunakan perangkat lunak internet dan protokol TCP/IP atau HTTP. Oleh karena itu, sebuah jaringan intranet memiliki semua fasilitas yang dimiliki oleh internet seperti e-mail, File Transfer Protocol (FTP), dan lain sebagainya. Jaringan intranet merupakan jaringan internet yang hanya dimiliki oleh perusahaan dan tidak dapat diakses dari luar.

B. Tinjauan Pustaka

Tinjauan pustaka ini bertujuan untuk memberikan landasan teoritis dan empiris bagi penelitian ini, yang berfokus pada optimasi keamanan autentikasi OTP berbasis Machine Learning menggunakan algoritma Random Forest. Melalui penelaahan berbagai sumber literatur yang ada, diharapkan dapat memperoleh pemahaman yang lebih mendalam tentang permasalahan dan solusi yang ditawarkan, serta membangun kerangka penelitian yang kuat untuk studi ini. Berikut ini adalah beberapa penelitian terkait yang telah dilakukan oleh para ahli dan menjadi acuan dalam penelitian ini:

1. (Muñoz et al., 2019) **Adjusting Emergent Herbaceous Wetland Elevation with Object-Based Image Analysis, Random Forest and NLCD** Penelitian ini mengembangkan metodologi untuk memperbarui dan mengoreksi elevasi hutan basah herba emergen menggunakan analisis citra berbasis objek dan teknik hutan acak dengan citra reflektansi permukaan Landsat. Metode ini meningkatkan akurasi pemetaan hutan basah dan mengembangkan alat 'koreksi-DEM' di ArcGIS untuk mengotomatiskan koreksi elevasi rawa asin. Alat ini divalidasi dengan data elevasi kinematik waktu nyata dan membantu mengoreksi elevasi rawa asin yang terlalu tinggi, sehingga dapat digunakan untuk pemetaan banjir yang lebih akurat di sistem estuaria.
2. (Bahrawi, 2019) **Sentiment Analysis Using Random Forest Algorithm-Online Social Media Based** Setiap hari miliaran data dalam bentuk teks membanjiri internet, baik dari forum, blog, media sosial, atau situs ulasan. Analisis sentimen dapat mengubah data yang sebelumnya tidak terstruktur menjadi lebih terstruktur dan menjadikan data tersebut informasi yang penting. Analisis sentimen merupakan salah satu bidang dari Pengolahan Bahasa Alami (NLP) yang membangun sistem untuk mengenali dan mengekstrak opini dalam bentuk teks.

Penelitian ini melakukan analisis sentimen dengan sumber data dari Twitter menggunakan pendekatan algoritma Random Forest. Hasil evaluasi menunjukkan akurasi sekitar 75%, yang cukup baik namun penelitian lebih lanjut dengan algoritma lain disarankan.

3. (Soogun et al., 2022) **Identifying Potential Factors Associated with High HIV Viral Load in South Africa Using Random Forest Analysis** Penelitian ini menganalisis faktor epidemiologis dan biologis yang terkait dengan beban virus HIV tinggi (HHVL) di KwaZulu-Natal, Afrika Selatan, menggunakan analisis korespondensi berganda (MCA) dan analisis hutan acak (RFA). Hasilnya menunjukkan bahwa lebih dari 40% peserta memiliki HHVL, dan faktor-faktor yang terkait dengan HHVL termasuk pengetahuan tentang status HIV, penggunaan ART, dosis ARV, hitungan sel CD4 saat ini, risiko yang dirasakan terkait dengan HIV, jumlah tes HIV seumur hidup, jumlah pasangan seks seumur hidup, dan pernah didiagnosis dengan TB. Kesimpulannya adalah bahwa proporsi individu dengan HHVL tinggi menunjukkan bahwa target pengobatan 95-95-95 UNAIDS untuk penekanan viral HIV kurang mungkin dicapai.
4. (Cattani, 2023) **Random Forest for Selecting Optimal Locations of Solar PV Plants** Studi ini menginvestigasi pengaruh variabel lingkungan pada efisiensi panel surya PV. Analisis Data Envelopment (DEA) digunakan untuk mengestimasi efisiensi 91 panel surya PV di Australia selama periode 2010-2020. Pengaruh variabel lingkungan pada efisiensi yang diestimasi diukur menggunakan model regresi terpotong. Hutan acak kemudian digunakan untuk memprediksi efisiensi panel surya PV di setiap kota di Australia. Hasilnya memungkinkan untuk menentukan lokasi dan wilayah yang paling sesuai untuk produksi energi surya PV di Australia. Studi ini menyediakan alat yang menarik dan mudah diinterpretasikan untuk pembuat kebijakan.
5. (A. Kurniawan & Yulianingsih, 2021) **Pendugaan Fraud Detection pada Kartu Kredit dengan Machine Learning** Menjelaskan serangkaian pengujian terhadap model yang telah dikembangkan menggunakan algoritma *Random Forest Classification* diperoleh kesimpulan bahwa *feature* yang paling relevan adalah V17, melalui visualisasi akurasi dengan *confusion matrix*s dan perhitungan akurasi diperoleh nilai akurasi yang baik. Dengan demikian dapat disimpulkan bahwa algoritma *Random Forest Classifier* dengan model yang dikembangkan mampu melakukan klasifikasi untuk melakukan analisis tindak pidana penipuan kartu kredit.
6. (Zamachsari & Puspitasari, 2021) **Penerapan Deep Learning dalam Deteksi Penipuan Transaksi Keuangan Secara Elektronik** Perbandingan berbagai jenis

algoritma untuk mendeteksi penipuan transaksi keuangan secara elektronik. Penelitian menunjukkan nilai akurasi *Extra Trees Classification* yang dihasilkan mendapatkan hasil paling baik yaitu sebesar 99.835%. Selanjutnya ada algoritma random forest yang mendapatkan nilai akurasi 99.72%.

7. (Sumarni & Rustam, 2023) **Analisa Bonus Demografi Dengan Algoritma Machine Learning Di Kabupaten Gorontalo Utara** Penelitian ini membahas dampak bonus demografi di provinsi Gorontalo, Indonesia, di tengah pandemi COVID-19 dan tantangan yang datang dengan era digital. Pemerintah telah berupaya mengatasi dan mengendalikan pertumbuhan penduduk, yang telah mempengaruhi struktur demografi penduduk. Penelitian ini bertujuan untuk menghasilkan deskripsi produk analisis pola pengetahuan data penduduk di Gorontalo Utara dengan ekstraksi data menggunakan algoritma machine learning.
8. (Benedictus Simarmata & Dwi, 2022) **Analisa Rekomendasi Fitur Persetujuan Pinjaman Perusahaan Financial Technology Menggunakan Metode Random Forest** Penelitian melakukan semua tahapan mulai dari *Data understanding*, *Feature extraction*, *Data Pre-processing*, Exploratory Data Analysis (EDA), *Modelling*, hingga *Insight*. Kesimpulan pertama yang didapat, algoritma *Random forest* digunakan dalam tahap *modelling* menghasilkan *recall* score sebesar 0.97, tergolong tinggi.
9. (Prasojo & Haryatmi, 2021) **Analisa Prediksi Kelayakan Pemberian Kredit Pinjaman dengan Metode Random Forest** Penelitian ini bertujuan untuk mengetahui penerapan metode klasifikasi dengan algoritma random forest dalam analisa pemberian kredit oleh bank untuk mengetahui kelayakan kredit pinjaman nasabah. Penelitian ini menggunakan data kreditur dan analisis dilakukan pada variabel V1 hingga V20 menggunakan perangkat lunak R. Hasil dari penelitian ini menunjukkan bahwa algoritma random forest memiliki tingkat akurasi sebesar 77%, yang termasuk dalam kategori klasifikasi model yang cukup baik.
10. (Reka, 2022) **Penerapan Data Mining Analisa Data Penjualan Obat Menggunakan Metode Random Forest** Penelitian ini membahas penggunaan metode Data Mining, khususnya metode Random Forest, untuk menganalisis data penjualan di apotek. Karena transaksi penjualan terjadi setiap hari, data penjualan yang tersimpan semakin besar seiring waktu. Metode Random Forest memungkinkan pengambilan data dan atribut secara acak untuk menghasilkan model pohon keputusan, yang memudahkan analisis data besar dan membantu memberikan informasi pada data penjualan yang diolah.

Tabel 7. Tinjauan Pustaka

No	:	1
Nama Peneliti	:	David F. Muñoz 1, Jordan R. Cissell, Hamed Moftakhari
Judul	:	Adjusting Emergent Herbaceous Wetland Elevation with Object-Based Image Analysis, Random Forest and NLCD
Permasalahan	:	ketidakakuratan model elevasi digital (DEM) saat ini dalam menggambarkan elevasi hutan basah herba emergen dan rawa asin, yang berdampak pada pemetaan banjir di sistem estuaria.
Link	:	https://com-mendeley-prod-publicsharing-pdfstore.s3.eu-west-1.amazonaws.com/3cad-CC-BY-2/10.3390/rs11202346.pdf
Kontribusi	:	Keefektifan algoritma Random Forest dalam melakukan optimasi pemetaan banjir yang lebih akurat dalam sistem esturia
No	:	2
Nama Peneliti	:	Bahrawi N
Judul	:	Sentiment Analysis Using Random Forest Algorithm-Online Social Media Based
Permasalahan	:	Sulitnya mengolah data teks yang besar dan tidak terstruktur dari media sosial seperti Twitter untuk mengenali dan mengekstrak opini masyarakat.
Link	:	https://doi.org/10.30818/jitu.2.2.2695
Kontribusi	:	Random forest dapat digunakan untuk mengolah data besar dan tidak terstruktur
No	:	3
Nama Peneliti	:	Adenike O. Soogun, Ayesha B. M. Kharsany, Temesgen Zewotir, Delia North, Ropo Ebenezer Ogunsakin
Judul	:	Identifying Potential Factors Associated with High HIV Viral Load in South Africa using Random Forest Analysis
Permasalahan	:	Proporsi yang tinggi dari individu dengan beban virus HIV tinggi (HHVL), yang mengindikasikan bahwa target pengobatan 95-95-95 UNAIDS untuk penekanan viral HIV kurang mungkin dicapai.
Link	:	https://link.springer.com/content/pdf/10.1186/s12874-022-01625-6.pdf
Kontribusi	:	Penelitian ini memberikan pemahaman mendalam mengenai implementasi algoritma random forest dalam mengolah data

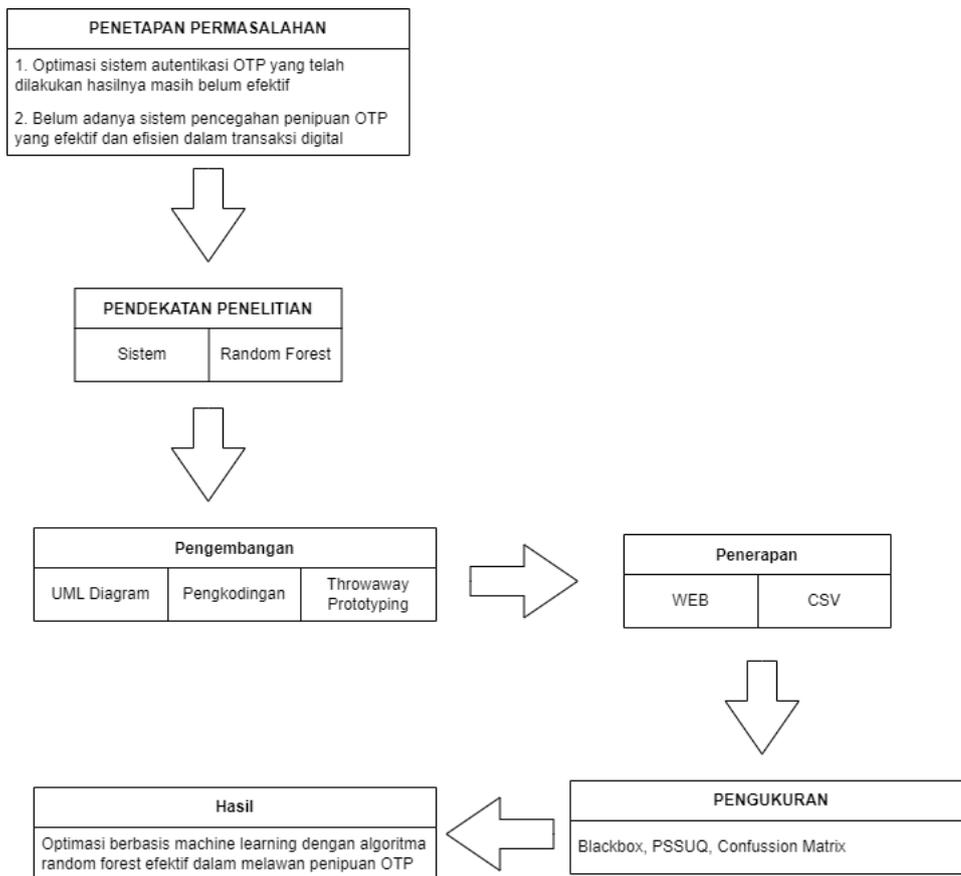
		yang besar dan bersifat kompleks
No	:	4
Nama Peneliti	:	Gilles Cattani
Judul	:	Random Forest for selecting optimal locations of solar PV plants
Permasalahan	:	Perlunya efisiensi dalam implementasi panel surya PV
Link	:	https://www.sciencedirect.com/science/article/pii/S2666546822000684
Kontribusi	:	Pemanfaatan banyak fitur data sehingga dapat menghasilkan model yang akurat dalam implementasi random forest
No	:	5
Nama Peneliti	:	Kurniawan, Yulianingsih
Judul	:	Pengukuran efektifitas algoritma random forest dalam menganalisa tindak pidana penipuan kartu kredit
Permasalahan	:	Pengukuran efektifitas metode ADASYN dalam pengolahan data untuk digunakan pada algoritma random forest
Link	:	https://jurnal.itpln.ac.id/kilat/article/view/1482/888
Kontribusi	:	Eksperimen dengan algoritma Random Forest Classification menunjukkan efektivitasnya dalam analisis tindak pidana penipuan kartu kredit, namun penting untuk mencatat bahwa hasil ini terbatas pada fitur dan model tertentu dan mungkin berbeda dalam skenario data yang berbeda atau dengan set fitur yang berbeda
No	:	6
Nama Peneliti	:	Sumarni S, Rustam S
Judul	:	Analisa Bonus Demografi Dengan Algoritma Machine Learning Di Kabupaten Gorontalo Utara
Permasalahan	:	Dampak bonus demografi di provinsi Gorontalo, seperti pemulihan ekonomi yang tidak merata, pengangguran, stunting, dan gizi buruk, di tengah pandemi COVID-19 dan kemajuan teknologi informasi
Link	:	https://dx.doi.org/10.56327/jtksi.v6i1.1391
Kontribusi	:	Penggunaan algoritma machine learning, seperti random forest

	untuk menganalisis data demografis
No	: 7
Nama Peneliti	: Zachmansari, Puspitasari
Judul	: Penerapan Deep Learning dalam Deteksi Penipuan Transaksi Keuangan Secara Elektronik
Permasalahan	: Mengukur efektifitas berbagai algoritma machine learning yang berbasiskan tree
Link	: https://jurnal.iaii.or.id/index.php/RESTI/article/view/2952/391
Kontribusi	: Penelitian ini memperlihatkan keefektifan algoritma Extra Trees Classification dan Random Forest dalam mendeteksi penipuan transaksi keuangan, namun diperlukan optimasi lebih lanjut untuk meningkatkan kinerja algoritma Random Forest
No	: 8
Nama Peneliti	: Reka C
Judul	: Penerapan Data Mining Analisa Data Penjualan Obat Menggunakan Metode Random Forest
Permasalahan	: Cara menganalisis data penjualan yang sangat besar dan terus meningkat setiap harinya di apotek untuk mendapatkan informasi yang berguna
Link	: https://www.mendeley.com/catalogue/3598a632-1609-3709-a4c2-74990d3a14e7/
Kontribusi	: Penerapan metode Random Forest dalam Data Mining untuk menganalisis data yang besar
No	: 9
Nama Peneliti	: Budi Prasojoa, Emy Haryatmi
Judul	: Analisa Prediksi Kelayakan Pemberian Kredit Pinjaman dengan Metode Random Forest
Permasalahan	: Bagaimana cara memprediksi kelayakan kredit pinjaman nasabah secara dini untuk membantu bank dalam mengambil keputusan atau kebijakan.
Link	: https://www.mendeley.com/catalogue/api/fulltext-resolver/e65e9408-bd13-3643-bf46-60cb6541e4aa/?t=1694228963514

Kontribusi	:	penggunaan algoritma random forest untuk memprediksi kejadian berdasarkan data historis
No	:	10
Nama Peneliti	:	Benerdictus, Dwi K
Judul	:	Analisa Rekomendasi Fitur Persetujuan Pinjaman Perusahaan Financial Technology Menggunakan Metode Random Forest
Permasalahan	:	https://www.researchgate.net/publication/363589003
Link	:	Pengujian akurasi random forest untuk persetujuan pinjaman perusahaan
Kontribusi	:	Penelitian ini menunjukkan bahwa melalui tahapan mulai dari pemahaman data, ekstraksi fitur, pra-pemrosesan data, analisis data eksploratoris, pemodelan, hingga wawasan, algoritma Random Forest berhasil mencapai skor recall sebesar 0.97, yang merupakan nilai tinggi.

Penelitian ini memberikan kontribusi baru dan penting dalam bidang deteksi penipuan, khususnya penipuan OTP, dengan fokus pada implementasi algoritma Random Forest. Meskipun studi sebelumnya telah menunjukkan efektivitas algoritma Random Forest dalam mendeteksi penipuan, penelitian ini berfokus pada pengembangan model yang spesifik untuk penipuan OTP dalam transaksi digital. Selain itu, penelitian ini mengintegrasikan model tersebut ke dalam API yang dapat digunakan oleh berbagai platform transaksi digital. Berbeda dari penelitian sebelumnya yang telah menjelajahi metode machine learning, penelitian ini secara khusus memanfaatkan algoritma Random Forest. Algoritma ini dipilih berdasarkan keefektifannya dalam mendeteksi penipuan dalam berbagai konteks, termasuk dalam data yang tidak seimbang. Dengan demikian, penelitian ini menghasilkan pendekatan baru dan lebih spesifik untuk mendeteksi dan mencegah penipuan OTP, yang berpotensi untuk meningkatkan keamanan transaksi digital dan memberikan perlindungan yang lebih baik bagi pengguna.

C. Kerangka Pemikiran



Gambar 4. Kerangka Pemikiran

Berdasarkan gambar 4, maka kerangka pemikiran dapat di deskripsikan sebagai berikut:

1. Penetapan Masalah: Autentikasi One-Time Password (OTP) memegang peran penting dalam memastikan keamanan transaksi digital. Namun, optimasi sistem autentikasi OTP yang telah dilakukan masih menunjukkan kinerja yang kurang optimal. Selain itu, belum ada sistem pencegahan penipuan OTP yang efektif dan efisien, sehingga menimbulkan kerentanan terhadap penyalahgunaan kode OTP.
2. Pendekatan Penelitian: Untuk mengatasi permasalahan tersebut, penelitian ini merancang pembuatan sistem menggunakan algoritma Random Forest. Algoritma ini dipilih berdasarkan keakuratannya dalam mengklasifikasikan data dan potensinya dalam mendeteksi penipuan.
3. Pengembangan: Penelitian ini merancang model sistem menggunakan Unified Modeling Language (UML), sebelum melanjutkan ke proses pengkodean dengan bahasa pemrograman Python. Metode Software Development Life Cycle (SDLC) tipe throwaway prototyping digunakan dalam pengembangan sistem, memungkinkan proses iteratif dan peningkatan berkelanjutan.

4. Penerapan: Sistem yang dikembangkan ini nantinya akan diimplementasikan pada platform web dan juga pada data CSV. Diharapkan dengan penerapan ini, sistem mampu beroperasi secara optimal dan melakukan deteksi penipuan OTP.
5. Pengukuran: Evaluasi kinerja sistem dilakukan menggunakan teknik pengujian Black Box, metode Post-Study System Usability Questionnaire (PSSUQ) untuk menilai kepuasan pengguna, serta Confusion Matrix untuk melihat performa klasifikasi algoritma.
6. Hasil: Melalui pendekatan dan proses yang telah dilakukan, diharapkan sistem berbasis machine learning dengan algoritma Random Forest yang dikembangkan ini dapat berperan efektif dalam mencegah dan mendeteksi penipuan OTP, sehingga meningkatkan keamanan dan kepercayaan pengguna dalam transaksi digital.

D. Hipotesis Penelitian

Algoritma Random Forest merupakan metode pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. Mengacu pada hasil penelitian sebelumnya, metode Random Forest mampu melakukan deteksi penipuan dengan presisi yang tinggi. Dalam konteks penelitian ini, terdapat kebutuhan untuk meningkatkan efektivitas sistem deteksi penipuan OTP dalam transaksi digital. Oleh karena itu, penelitian ini akan mengusulkan penggunaan algoritma Random Forest untuk meningkatkan presisi dan akurasi dalam mendeteksi penipuan OTP. Algoritma ini diduga mampu memberikan hasil yang lebih baik dalam mendeteksi penipuan OTP, dan dengan demikian dapat memberikan informasi yang lebih akurat dan tepat waktu untuk pihak-pihak terkait dalam mengambil tindakan pencegahan dan penanganan penipuan OTP.