

PROJECT AKHIR SEMESTER

PENERAPAN METODE *RADIAL BASIS FUNCTION* DENGAN *GENETIC ALGORITHM* UNTUK MENDETEKSI PENIPUAN KARTU KREDIT



Disusun oleh:

1. Reyner Ongkowijoyo (181221042)
2. Muh. Fath Rajihan N. (181221084)

Dosen Pengampu:

Auli Damayanti, S.Si., M.Si.

PROGRAM STUDI S-1 MATEMATIKA

DEPARTEMEN MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS AIRLANGGA

2024

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas limpahan rahmat dan anugrah dari-Nya kami dapat menyelesaikan project akhir semester dalam mata kuliah Jaringan Syaraf Tiruan tentang **“Penerapan Metode Radial Basis Function dengan Genetic Algorithm untuk Mendeteksi Penipuan Kartu Kredit”** ini. Dalam makalah ini kami membahas mengenai penerapan salah satu metode pada Jaringan Syaraf Tiruan, yaitu *Radial Basis Function* untuk mendeteksi *fraud* pada kartu kredit.

Makalah ini disusun untuk memenuhi penilaian ujian akhir semester dalam mata kuliah jaringan syaraf tiruan. Di samping itu, kami ucapkan terimakasih kepada Auli Damayanti, S.Si., M.Si. selaku dosen pengampu mata kuliah jaringan syaraf tiruan yang telah membimbing kami sehingga ilmu yang diberikan dapat kami manfaatkan untuk mengimplementasikan permasalahan secara nyata yang ada pada saat ini.

Penulis menyadari bahwa penyusunan makalah ini masih jauh dari kata sempurna. Oleh karena itu, kritik dan saran pembaca sangat diharapkan untuk perbaikan tugas penulis selanjutnya. Penulis mohon maaf apabila dalam tugas akhir ini terdapat kesalahan, baik dalam penulisan maupun isi makalah ini. Terima kasih.

Surabaya, 1 Desember 2024

Penulis

DAFTAR ISI

KATA PENGANTAR.....	i
DAFTAR ISI.....	ii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian.....	2
1.4 Manfaat Penelitian.....	2
1.5 Batasan Masalah.....	3
BAB II TINJAUAN PUSTAKA	4
2.1 Kartu Kredit.....	4
2.2 <i>Fraud</i> pada perbankan.....	4
2.3 Jaringan Syaraf Tiruan	5
2.4 Radial Basis Function.....	6
2.5 Algoritma Genetika	7
2.6 Evaluasi	8
2.7 Python.....	9
BAB III METODOLOGI PENELITIAN	10
3.1 Data penelitian.....	10
3.2 Alur penelitian.....	10
3.3 <i>Data Preprocessing</i>	11
3.4 <i>Modelling</i>	12
BAB IV HASIL DAN PEMBAHASAN	14
4.1 Parameter Algoritma Genetika.....	14
4.2 Analisa Sistem (Manual).....	14
4.3 Analisa Sistem (Program)	19
BAB V PENUTUP.....	20
5.1 Kesimpulan.....	20
5.2 Saran.....	20
DAFTAR PUSTAKA.....	21

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di era digitalisasi dan Revolusi Industri 4.0, transformasi digital di sektor perbankan menjadi kebutuhan yang mendesak untuk menghadapi perubahan perekonomian global [1]. Transformasi ini memungkinkan institusi perbankan untuk meningkatkan efisiensi operasional, mempercepat inovasi layanan, dan memperkuat daya saing di tengah dinamika pasar yang semakin kompleks. Dalam menghadapi persaingan yang semakin ketat, adopsi teknologi digital membantu bank tetap relevan sekaligus memenuhi harapan nasabah yang terus berkembang, khususnya dalam hal kenyamanan, aksesibilitas, personalisasi layanan, dan tentunya keamanan [2].

Seiring dengan semakin populernya transaksi dan metode pembayaran digital, risiko penipuan keuangan juga mengalami peningkatan signifikan, menimbulkan kerugian ekonomi yang besar bagi pelanggan maupun lembaga keuangan [3]. Transformasi digital di sektor keuangan, meskipun memberikan banyak kemudahan dan efisiensi, juga membuka peluang bagi pelaku kejahatan siber untuk mengembangkan teknik penipuan yang semakin canggih. Jenis penipuan seperti phishing, *malware infections*, dan penggunaan situs web palsu (*ghost websites*) kini menjadi ancaman yang semakin sulit untuk dideteksi oleh pendekatan tradisional berbasis aturan [4].

Penipuan kartu kredit adalah salah satu tantangan global yang terus berkembang di era digital. Dengan semakin luasnya adopsi transaksi non-tunai, ancaman penipuan kartu kredit semakin meningkat, mengancam keamanan data pengguna dan kepercayaan terhadap sistem keuangan digital. Fenomena ini menunjukkan pentingnya pengembangan metode deteksi dan pencegahan penipuan yang lebih canggih untuk mengantisipasi potensi kerugian finansial [5]

Seperti yang dilaporkan dalam studi ACFE (Association of Certified Fraud Examiners) tahun 2022, penyelewengan aset merupakan penyebab 86% kasus penipuan yang dilaporkan, termasuk penyalahgunaan otoritas dan pencurian data dalam ekosistem keuangan. Dalam konteks penipuan kartu kredit, bentuk-bentuk yang sering terjadi meliputi penggunaan kartu secara ilegal, pencurian data kartu melalui metode *phishing* atau *scamming*, hingga transaksi tidak sah akibat kerentanan sistem keamanan.

Peningkatan ancaman ini menuntut perusahaan keuangan, penyedia layanan pembayaran, dan pengguna untuk lebih waspada. Implementasi teknologi berbasis kecerdasan buatan dan jaringan syaraf buatan telah menjadi solusi utama dalam mendeteksi pola anomali pada transaksi kartu kredit. Algoritma modern dapat menganalisis ribuan data transaksi dalam waktu singkat untuk mengenali perilaku

mencurigakan, seperti transaksi dalam jumlah besar yang tidak biasa atau penggunaan kartu di lokasi geografis yang berbeda dalam waktu berdekatan.

Untuk mengatasi tantangan tersebut, Algoritma Genetika yang terinspirasi oleh proses evolusi biologis, dapat digunakan sebagai metode optimasi hiperparameter [6]. Algoritma ini secara iteratif mengeksplorasi ruang hiperparameter dan memilih kromosom atau individu terbaik yang mewakili sebuah hiperparameter berdasarkan fungsi kebugaran (*fitness function*), yang dalam hal ini adalah skor metrik berdasarkan dataset pengujian. Dengan cara ini, Algoritma Genetika dapat membantu mengidentifikasi kombinasi hiperparameter yang optimal untuk meningkatkan skor metrik dalam klasifikasi deteksi penipuan kartu kredit menggunakan model Radial Basis Function.

Penelitian ini berfokus pada pembangunan model jaringan saraf Radial Basis Function untuk mendeteksi penipuan kartu kredit, mengoptimalkan hiperparameternya menggunakan Algoritma Genetika, dan mengevaluasi kinerjanya setelah proses optimasi. Dengan mengintegrasikan metode ini, penelitian ini bertujuan untuk berkontribusi dalam pengembangan sistem deteksi penipuan yang lebih akurat serta mengeksplorasi potensi Algoritma Genetika dalam penyetelan *hyperparameters*.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian, rumusan masalah dalam penelitian ini adalah sebagai berikut.

1. Bagaimana cara menerapkan model Radial Basis Function yang dioptimasi dengan Algoritma Genetika untuk mendeteksi penipuan pada kartu kredit ?
2. Bagaimana performa model Radial Basis Function yang dioptimasi dengan Algoritma Genetika untuk mendeteksi penipuan pada kartu kredit ?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah penelitian yang telah dibuat, maka tujuan penelitian ini adalah sebagai berikut.

1. Menerapkan Algoritma Genetika untuk mengoptimalkan hyperparameter pada model RBF dalam deteksi penipuan kartu kredit.
2. Mengevaluasi kinerja model RBF-GA dalam mendeteksi penipuan kartu kredit melalui analisis akurasi dengan F1-Score dan efisiensi model.

1.4 Manfaat Penelitian

Berdasarkan rumusan tujuan penelitian yang telah dibuat, maka manfaat penelitian ini adalah sebagai berikut.

1. Menghasilkan model deteksi penipuan yang lebih akurat dan efisien dengan mengurangi kesalahan deteksi. Diharapkan dapat meminimalkan kerugian akibat penipuan kartu kredit.
2. Memperdalam wawasan tentang efektivitas Algoritma Genetika dalam mengoptimasi model RBF. Hal ini relevan dalam menangani masalah anomali kompleks seperti penipuan.
3. Metode yang dikembangkan dapat diterapkan pada berbagai bidang keamanan finansial. Solusi ini mendukung sistem keamanan data keuangan yang lebih luas dan terintegrasi

1.5 Batasan Masalah

Penelitian ini memiliki beberapa batasan yang perlu diperhatikan. Dataset yang digunakan merupakan bagian dari *The Bank Account Fraud (BAF) Suite* yang dipublikasikan pada NeurIPS 2022. Dataset ini telah dilakukan proses *downsampling* dari total sekitar 1 juta data menjadi 2.000 data secara seimbang (*balanced*). Dari 32 fitur yang tersedia, hanya 6 fitur yang dipilih berdasarkan pengetahuan domain untuk menyederhanakan proses komputasi. Model RBF yang digunakan dalam penelitian ini hanya memanfaatkan fungsi basis radial Gaussian sebagai fungsi aktivasi pada lapisan tersembunyi. Selain itu, penggunaan Algoritma Genetika terbatas hanya untuk optimasi parameter pada model RBF, tidak mencakup pemilihan fitur atau pemrosesan data lainnya.

BAB II

TINJAUAN PUSTAKA

2.1 Kartu Kredit

Kartu kredit adalah alat pembayaran yang berfungsi sebagai pengganti uang tunai dan dapat digunakan oleh konsumen kapan saja untuk membeli barang atau jasa di tempat yang menerima kartu kredit (*merchant*). Selain itu, kartu kredit juga memungkinkan konsumen untuk mencairkan uang tunai melalui bank penerbit atau jaringan yang terhubung (*cash advance*) [7]. Kartu kredit merupakan salah satu bentuk Alat Pembayaran Menggunakan Kartu (APMK) yang menawarkan dua fungsi utama bagi konsumen, yaitu sebagai alat pembayaran dan sebagai sumber kredit [8].

Sementara itu, Bank Indonesia (2020) mendefinisikan kartu kredit sebagai salah satu bentuk APMK yang dapat digunakan untuk melakukan pembayaran kewajiban yang muncul dari aktivitas ekonomi, termasuk transaksi belanja maupun penarikan tunai. Dalam penggunaannya, kewajiban pembayaran pemegang kartu akan terlebih dahulu dipenuhi oleh penerbit atau *acquirer*. Pemegang kartu kemudian berkewajiban melunasi pembayaran sesuai dengan kesepakatan, baik dengan pelunasan sekaligus (*charge card*) maupun secara angsuran.

Kartu kredit selain digunakan sebagai alat pembayaran juga bermanfaat untuk pinjaman murah, untuk berbelanja online, dalam transaksi online dapat dibatalkan dengan mudah, pengeluaran tercatat, banyak potongan ketika berbelanja, traveling menjadi lebih murah, akomodasi lebih hemat, bisa dijadikan mata uang internasional, praktis dan aman, bisa digunakan untuk mengajukan kredit tanpa agunan [9].

2.2 Fraud pada perbankan

Penipuan perbankan adalah kejahatan keuangan yang melibatkan praktik ilegal untuk memperoleh keuntungan secara tidak sah. Penipuan ini mencakup berbagai aktivitas seperti penyetoran, penarikan, transfer, pengajuan pinjaman, dan investasi, serta dapat dilakukan oleh individu maupun organisasi. Bentuk penipuan yang umum meliputi pencurian identitas, phishing, dan pengajuan pinjaman palsu. Penipu memanfaatkan celah sistem, menggunakan dokumen palsu, atau mencuri informasi pribadi untuk melakukan tindakan ilegal.

Menurut [10], beberapa jenis penipuan perbankan meliputi:

1. **Penipuan akuntansi:** Manipulasi laporan keuangan untuk menutupi kondisi sebenarnya.
2. **Penipuan diskon tagihan:** Penyalahgunaan tagihan tanpa agunan memadai.
3. **Cheque kiting:** Memanfaatkan jeda waktu pencairan cek.
4. **Pemalsuan dokumen:** Penggunaan dokumen palsu untuk keuntungan ilegal.
5. **Pemalsuan cek:** Mengubah informasi cek untuk tujuan penipuan.
6. **Pencurian identitas:** Penggunaan informasi pribadi untuk aktivitas ilegal.
7. **Pencucian uang:** Menyembunyikan sumber uang ilegal melalui sistem keuangan.

8. **Penipuan kartu pembayaran:** Penggunaan kartu kredit/debit secara ilegal.
9. **Phishing:** Teknik manipulasi untuk mencuri informasi sensitif.
10. **Penipuan transfer kawat:** Transfer uang secara ilegal melalui layanan transfer.

Kerugian akibat penipuan dapat berdampak besar pada bank dan nasabah. Oleh karena itu, penting untuk meningkatkan kewaspadaan dengan memeriksa aktivitas rekening, melindungi informasi pribadi, dan melaporkan aktivitas mencurigakan. Penegakan hukum yang ketat juga diperlukan untuk mencegah kasus ini.

2.3 Jaringan Syaraf Tiruan

Jaringan Syaraf Tiruan (JST) adalah metode pemrosesan informasi yang terinspirasi dari cara kerja sistem syaraf biologis, khususnya otak, dalam memproses data. Konsep ini berfokus pada pembentukan struktur komputasi yang terdiri dari banyak unit pemrosesan kecil yang saling terhubung, disebut neuron. Neuron-neuron ini bekerja secara kolaboratif untuk menyelesaikan masalah tertentu, seperti pengenalan pola, klasifikasi, atau prediksi. Dengan memanfaatkan kemampuan adaptasi melalui pembelajaran, JST mampu meniru perilaku kecerdasan alami, menjadikannya solusi yang sangat fleksibel untuk berbagai aplikasi teknologi modern [11]. Secara umum, arsitektur Jaringan Syaraf Tiruan (JST) terdiri dari tiga lapisan utama:

1. **Input Layer:** Lapisan ini menerima data awal yang dimasukkan ke jaringan. Setiap neuron dalam input layer mewakili fitur dari data yang digunakan sebagai masukan.
2. **Hidden Layer:** Lapisan tersembunyi bertugas memproses data dari input layer. Di sinilah perhitungan kompleks berlangsung melalui bobot, bias, dan fungsi aktivasi, sehingga jaringan dapat mengenali pola atau hubungan yang tersembunyi dalam data.
3. **Output Layer:** Lapisan ini menghasilkan hasil akhir dari JST, seperti klasifikasi, prediksi, atau keluaran lain sesuai tujuan jaringan. Neuron-neuron di lapisan ini mewakili kategori atau nilai hasil.

Setiap lapisan bekerja secara berurutan dan saling mendukung dalam menghasilkan output berdasarkan data masukan. Pada penelitian [12] proses pembelajaran Jaringan Syaraf Tiruan mendukung beberapa metode pembelajaran, yaitu:

1. **Supervised Learning:** Metode pembelajaran yang melibatkan pelatihan model *machine learning* atau jaringan syaraf tiruan dengan dataset berlabel, yang terdiri dari input dan output yang diketahui. Selama pelatihan, model belajar untuk membuat pemetaan antara input dan output melalui iterasi, dengan bantuan pengawas yang mengatur parameter model. Model yang terbaik dalam mencocokkan pemetaan ini akan dihasilkan dan kemudian dapat digunakan untuk memprediksi output yang diharapkan ketika diberikan input baru [12].
2. **Unsupervised Learning:** Metode pembelajaran yang berfokus pada penemuan pola tersembunyi dan struktur dalam data yang tidak berlabel. Model ini mengelompokkan data berdasarkan kesamaan antar data tersebut tanpa adanya

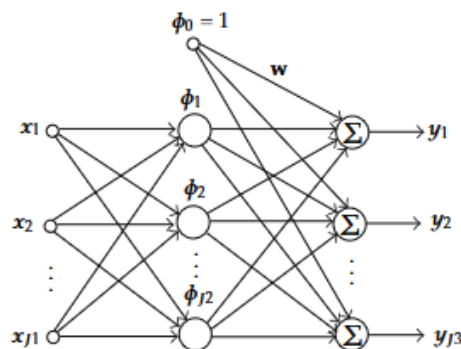
supervisor atau output yang terkait dengan input. Berbeda dengan pembelajaran terawasi, di mana output sudah diketahui, pembelajaran tanpa pengawasan menggunakan fitur yang telah dipelajari untuk mengenali kelas data baru. Masalah dalam pembelajaran tanpa pengawasan meliputi klustering, reduksi dimensi, dan deteksi anomaly [12].

3. **Semi-Supervised Learning:** Sebuah pendekatan yang menggabungkan elemen dari *supervised learning* dan *unsupervised learning*. Dalam metode ini, sebagian besar data tidak memiliki label, tetapi sejumlah kecil data berlabel tersedia. Model belajar dari data berlabel tersebut untuk kemudian memperkirakan atau mengklasifikasikan data tak berlabel dengan menggunakan pola yang ditemukan. Pembelajaran semi-terawasi berguna dalam situasi di mana memperoleh data berlabel sangat mahal atau sulit, namun masih ingin mendapatkan manfaat dari jumlah data tak berlabel yang besar..
4. **Reinforcement Learning:** suatu kerangka kerja di mana agen atau pengendali mengoptimalkan perilakunya dengan berinteraksi dengan lingkungan. Dalam RL, agen tidak diberitahu sebelumnya tindakan apa yang harus dilakukan, melainkan belajar dari pengalaman melalui percakapan (*trial and error*). Tujuan utama agen adalah untuk memaksimalkan penghargaan kumulatif (*reward*) yang diterima dengan melakukan aksi yang membawa hasil terbaik. Proses ini melibatkan pemetaan keadaan (*state*) ke tindakan (*action*), dengan agen berusaha menemukan pola yang memberikan penghargaan tertinggi. Keunikan RL terletak pada ketergantungannya pada aksi dan umpan balik yang diberikan oleh lingkungan.

Setiap metode pembelajaran memiliki karakteristik yang sesuai dengan kebutuhan aplikasi yang berbeda, seperti pengenalan pola, klasifikasi, atau optimasi. Dengan kemampuannya untuk belajar dari data dan beradaptasi melalui proses iteratif, Jaringan Syaraf Tiruan terus menjadi fondasi utama dalam pengembangan teknologi cerdas yang dapat diterapkan di berbagai bidang, mulai dari pengolahan gambar, pemrosesan bahasa alami, hingga pengambilan keputusan otomatis.

2.4 Radial Basis Function

Radial Basis Function (RBF) merupakan jaringan syaraf tiruan dengan tiga lapisan, yang terdiri dari *input layer*, *hidden layer*, dan *output layer*, seperti yang ditunjukkan pada Gambar 1 [13].



Gambar 1 Arsitektur Jaringan RBF. $\phi_0(\vec{x}) = \mathbf{1}$ berhubungan dengan bias dalam lapisan output, sedangkan $\phi_i(\vec{x})$ menunjukkan non-linieritas pada *hidden nodes*.

Setiap node tersembunyi menerapkan Radial Basis Function, yang disebut $\phi(r)$, sebagai fungsi aktivasi nonlinear. Lapisan tersembunyi melakukan transformasi nonlinier dari input, sementara lapisan output bertindak sebagai penggabung linear, mengubah nonlinieritas ke ruang baru. Biasanya, semua node RBF berbagi fungsi aktivasi yang sama. Nonlinieritas pada setiap node RBF direpresentasikan sebagai $\phi_i(\vec{x} - \vec{c}_i)$, dimana \vec{c}_i menunjukkan pusat atau prototipe dari node ke- i , dan $\phi_i(\vec{x})$ menunjukkan Radial Basis Function. Selain itu, bias untuk neuron-neuron di lapisan output dapat dimasukkan melalui neuron tambahan di lapisan tersembunyi [13].

Salah satu fungsi yang dapat digunakan sebagai RBF adalah fungsi Gaussian [14], yang didefinisikan sebagai:

$$\phi(r) = e^{-r^2/2\sigma^2} \quad (1)$$

dimana $r > 0$ menunjukkan nilai jarak dari titik data \vec{x} terhadap pusat or *centroid* \vec{c} , and σ yang kadang disebut sebagai spread digunakan untuk mengontrol kelancaran fungsi interpolasi. Selain lapisan tersembunyi, lapisan output pada jaringan RBF adalah penggabung linear yang didefinisikan sebagai

$$y_i(\vec{x}) = \sum_{i=1}^J w_i \phi_i(\vec{x}) \quad (2)$$

dimana w_i adalah bobot yang menghubungkan lapisan tersembunyi ke lapisan output, J adalah jumlah node RBF di lapisan tersembunyi, dan $y_i(\vec{x})$ adalah *output* dari node ke- i pada jaringan untuk *input* \vec{x} yang diberikan.

Salah satu metode pelatihan untuk optimasi bobot adalah regresi linear yang meminimalkan kesalahan output. Dalam praktiknya, metode pseudo-inverse atau metode kuadrat terkecil dapat digunakan untuk menghitung bobot-bobot ini:

$$W = \Phi^+ Y \quad (3)$$

Dimana Φ adalah matriks yang dihasilkan pada *hidden layer* (*Gaussian activations*), Y merupakan *target output matrix*, dan Φ^+ adalah fungsi pseudo-inverse dari matriks Φ [13].

2.5 Algoritma Genetika

Algoritma genetika adalah sebuah metode optimasi dalam bidang komputasi yang berbasis pada prinsip-prinsip evolusi genetika dan seleksi alam. Sehingga, Istilah-istilah dalam algoritma genetika mengadopsi istilah-istilah pada makhluk hidup, yaitu [14]:

1. **Kromosom:** Kromosom, atau sering disebut individu, merepresentasikan satu solusi dari masalah optimasi tertentu. Sebuah kromosom dapat terdiri dari beberapa gen (variabel).
2. **Induk:** Induk merujuk pada individu (solusi) yang dipilih untuk dikawinkan atau digabungkan dengan individu lain sehingga menghasilkan solusi baru yang disebut keturunan (offspring).
3. **Seleksi Induk:** Seleksi induk adalah proses memilih individu dari populasi untuk dijadikan induk.
4. **Pindah Silang:** Pindah silang merupakan proses kombinasi sifat atau nilai dari dua individu induk. Hasilnya adalah satu atau dua individu baru (offspring) yang menggabungkan karakteristik dari kedua induk.
5. **Mutasi:** Mutasi adalah proses modifikasi sifat atau nilai dalam suatu individu, menghasilkan individu baru. Mutasi berfungsi untuk mengeksplorasi ruang solusi yang lebih luas. Tingkat mutasi yang tinggi memungkinkan eksplorasi lebih banyak, tetapi berisiko merusak solusi yang sudah baik. Oleh karena itu, probabilitas mutasi harus diatur secara optimal.
6. **Fitness:** Fitness menggambarkan tingkat kelayakan atau ketahanan suatu individu dalam populasi. Individu dengan nilai fitness tinggi memiliki peluang lebih besar untuk bertahan hingga generasi berikutnya. Namun, dalam kasus tertentu individu dengan nilai fitness rendah memiliki peluang lebih besar untuk bertahan hingga generasi berikutnya.

Algoritma ini dikembangkan oleh John Holland pada tahun 1975 dan dipopulerkan oleh David Goldberg pada tahun 1989. Salah satu keuntungan utama dari algoritma genetika dibandingkan dengan metode konvensional adalah bahwa algoritma genetika tidak memerlukan evaluasi gradien fungsi tujuan [15]. Sifat utama dari algoritma genetika adalah mencari beberapa kemungkinan dari solusi awal untuk mendapatkan sebuah solusi yang optimal dalam menyelesaikan permasalahan komputasi

2.6 Evaluasi

Data pelatihan yang telah disiapkan akan diuji melalui beberapa metrik evaluasi, yaitu *F1 Score*, *F1 Score* adalah metrik yang menggambarkan keseimbangan antara *precision* dan *recall*, pada [17] menunjukkan seberapa baik model dalam mengidentifikasi data yang relevan. Agar mudahnya dapat dilihat pada gambar 2.

		Nilai Aktual	
		Positive	Negative
Nilai Prediksi	Positive	TP (True Positive)	FP (False Positive)
	Negative	FN (False Negative)	TN (True Negative)

Gambar 2 *Confussion Matrix.*

2.6.1 Akurasi

Akurasi merupakan rasio antara jumlah prediksi yang benar (baik untuk kelas positif maupun negatif) dibandingkan dengan total keseluruhan data. Akurasi digunakan untuk menggambarkan sejauh mana model klasifikasi dapat memprediksi data dengan benar. Nilai akurasi dihitung dengan persamaan berikut:

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

2.6.2 Precision

Precision adalah rasio antara jumlah prediksi positif yang benar (*true positives*) dibandingkan dengan keseluruhan hasil yang diprediksi sebagai positif (*true positives* ditambah *false positives*). *Precision* mengukur seberapa akurat model dalam memprediksi kelas positif, atau dengan kata lain, berapa banyak prediksi positif yang benar-benar relevan. Nilai *precision* dihitung dengan persamaan berikut:

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

2.6.3 Recall

Recall mengukur rasio prediksi positif yang benar (*true positives*) terhadap seluruh data yang benar-benar positif. *Recall* menggambarkan seberapa baik model dalam menemukan seluruh data yang relevan atau mengidentifikasi kembali informasi yang ada. Nilai *recall* dihitung dengan persamaan berikut:

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

2.6.4 F1 Score

F1 Score adalah metrik yang menggabungkan nilai *precision* dan *recall* dalam satu angka untuk memberikan gambaran tentang kinerja model secara keseluruhan. *F1 Score* memberikan keseimbangan antara *precision* dan *recall*, dengan memberikan bobot yang lebih sama antara kedua metrik tersebut. Nilai *F1 Score* dihitung dengan persamaan berikut:

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

2.7 Python

Menurut pengertian dari Python Software Foundation (2016), Python adalah bahasa pemrograman interpretatif, berorientasi objek dan semantik yang dinamis. Python memiliki high-level struktur data, dynamic typing dan dynamic binding. Python memiliki sintaks sederhana dan mudah dipelajari untuk penekanan pada kemudahan membaca dan mengurangi biaya perbaikan program. Selain itu, bahasa pemrograman Python memiliki bermacam-macam Library yang memiliki kegunaannya masing-masing dan dapat digunakan oleh siapa saja di berbagai sistem operasi, atau dengan kata lain bersifat open source.

BAB III METODOLOGI PENELITIAN

3.1 Data penelitian

Dataset yang digunakan dalam penelitian ini merupakan data sekunder yang tersedia secara publik dan diperoleh dari platform Kaggle. Karena adanya keterbatasan komputasi, peneliti memilih subset data untuk diproses lebih lanjut dengan melakukan down sampling dari total sekitar 1 juta data menjadi 1.000 data yang seimbang untuk keperluan klasifikasi biner, khususnya antara transaksi penipuan (fraud) dan transaksi yang sah (non-fraud).

Selain itu, dari 32 fitur yang tersedia, dipilih 6 fitur dengan 1 target berdasarkan pengetahuan domain untuk menyederhanakan proses komputasi. Data hasil down sampling kemudian dibagi secara acak menjadi set pelatihan dan validasi, dengan alokasi 75% untuk pelatihan dan 25% untuk validasi. Gambaran umum dataset, termasuk fitur-fitur yang dipilih dan data hasil *down sampling*, disajikan dalam Gambar 3.

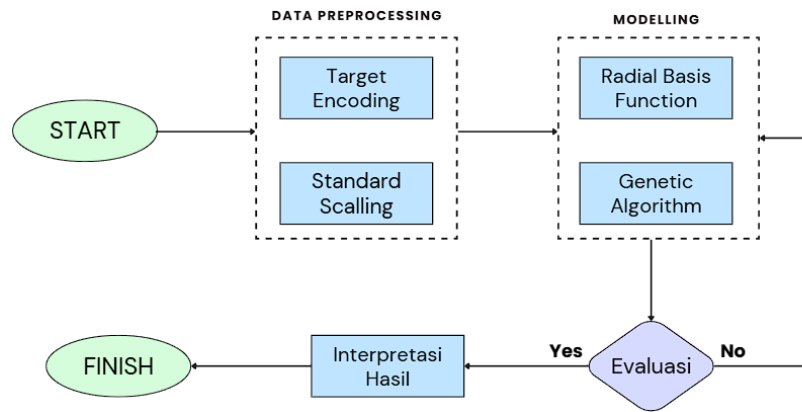
	income	credit_risk_score	customer_age	payment_type	employment_status	current_address_months_count	fraud_bool
0	0.8	295	40	AC	CA	73	1
1	0.6	97	40	AB	CA	136	1
2	0.2	58	30	AA	CD	70	1
3	0.1	224	40	AB	CA	131	1
4	0.6	268	50	AB	CA	149	1
...
745	0.8	141	20	AD	CA	30	0
746	0.6	95	40	AB	CA	245	0
747	0.1	196	20	AB	CA	8	0
748	0.2	61	30	AB	CB	383	0
749	0.1	143	20	AB	CB	28	0

Gambar 3 Data pelatihan

Berdasarkan gambar 3, atribut-atribut dalam dataset yang terkait dengan profil pemohon kredit. Data ini mencakup informasi pendapatan, skor risiko kredit, usia, jenis pembayaran, status pekerjaan, lama tinggal di alamat saat ini, dan label biner yang menunjukkan penipuan atau tidaknya aplikasi tersebut.

3.2 Alur penelitian

Pada penelitian ini, jenis penelitian yang digunakan adalah penelitian kuantitatif. Penelitian ini bertujuan untuk mendeteksi penipuan pada kartu kredit dengan memanfaatkan teknik gabungan antara Radial Basis Function sebagai model utama dengan algoritma genetika sebagai optimasi inisialisasi parameter yang terdapat pada radial basis function, sehingga alur penelitian ini dijelaskan pada gambar 4.



Gambar 4 Alur Penelitian

Berdasarkan Gambar 4, diagram alir ini menggambarkan tahapan proses penelitian yang dimulai dari tahap pra-pemrosesan data hingga interpretasi hasil. Tahap awal dimulai dengan *data preprocessing* yang mencakup dua metode utama, yaitu *Target Encoding* dan *Standard Scalling*. *Target Encoding*, kemudian setelah itu, data yang telah disiapkan akan digunakan untuk membangun model (*modelling*). Pada tahap modelling penelitian ini menggunakan Teknik gabungan antara Algoritma Genetika dengan Radial Basis Function. Selanjutnya setelah model di latih oleh data pelatihan, maka model akan di evaluasi dengan metrik evaluasi hingga hasil evaluasinya memuaskan sebelum diterapkan pada data validasi.

3.3 *Data Preprocessing*

Data preprocessing merupakan tahap awal yang sangat krusial dalam pengolahan data, di mana tujuan utamanya adalah mempersiapkan data mentah agar lebih bersih, konsisten, dan siap digunakan dalam tahap analisis atau pemodelan. Pada tahap ini, data yang akan diolah melalui serangkaian proses untuk mengatasi berbagai masalah kualitas data, seperti adanya *noise* (gangguan), data yang tidak konsisten, data yang hilang (*missing values*), serta data duplikat [17].

Tahap ini berperan penting dalam meningkatkan performa model analisis atau pemodelan. Data yang telah melalui tahap ini akan lebih berkualitas, lebih akurat, dan lebih relevan dengan tujuan penelitian, sehingga dapat menghasilkan model yang lebih baik dan handal. Dengan demikian, *data preprocessing* menjadi pondasi utama dalam memastikan keberhasilan suatu penelitian atau analisis berbasis data. Pada penelitian ini *data preprocessing* yang dilakukan oleh peneliti ada dua metode utama, yaitu *Target Encoding* dan *Standard Scalling*. *Target Encoding*

3.3.1 *Target Encoding*

Target encoding dihitung dengan mengambil rata-rata target untuk setiap kategori dalam data pelatihan, sehingga setiap kategori memiliki nilai numerik yang merepresentasikan kemungkinan kejadiannya dalam kaitannya dengan target. Rumus perhitungan mean target untuk kategori adalah sebagai berikut:

$$Enc(k) = \frac{\sum_{i=0}^{N_k} y_i}{N_k} \quad (8)$$

Dimana:

N_k : Jumlah total data dalam kategori k

y_i : Nilai target (1 *Fraud* and 0 *non-Fraud*)

$Enc(k)$: Nilai rata-rata target untuk kategori k

3.3.2 Standard Scaling

Standard Scaling bertujuan untuk mengubah distribusi fitur menjadi memiliki rata-rata 0 dan standar deviasi 1, agar skala nilai antar-fitur menjadi seragam dan meningkatkan stabilitas algoritma pada tahap pemodelan. Rumus Standard Scaling untuk setiap nilai x pada kolom fitur adalah sebagai berikut:

$$x' = \frac{x - \mu}{\sigma} \quad (9)$$

Dimana:

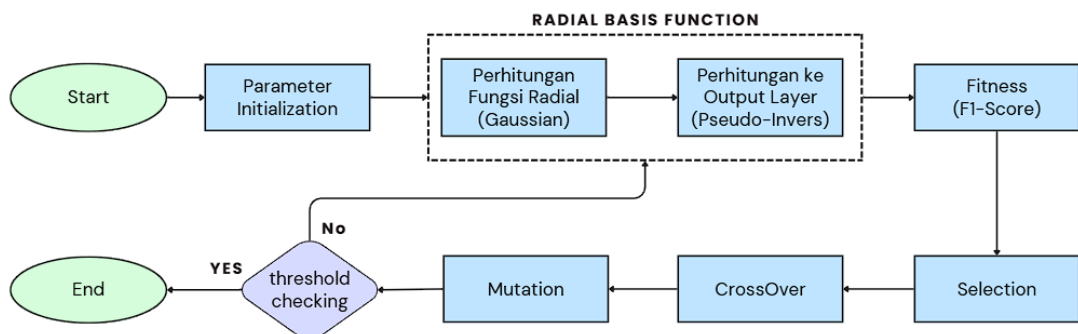
x' : Nilai dari fitur/kolom input

μ : Rata-rata dari fitur/kolom input

σ : Standard deviasi dari fitur/kolom input

3.4 Modelling

Setelah dataset diproses sehingga menjadi data yang siap pakai, maka Langkah berikutnya merupakan tahap pemodelan, proses ini bertujuan untuk menemukan pola atau hubungan dalam data yang dapat digunakan untuk melakukan prediksi atau klasifikasi pada data yang telah diproses. Langkah-langkah dalam tahap ini mencakup pemilihan algoritma, pelatihan model, evaluasi performa, dan penyempurnaan model berdasarkan hasil evaluasi. Pada tahap pemodelan penelitian ini menerapkan algoritma genetik dengan radial basis function untuk melatih model berdasarkan data yang telah diproses sebelumnya. Algoritma gabungan antara Radial Basis Function dengan Algoritma Genetika ditunjukkan pada gambar 5.



Gambar 5 Algoritma Hybrid RBF-GA

Implementasi antara algoritma genetika dengan radial basis function berdasarkan gambar 5, yaitu:

1. Inisialisasi nilai parameter: jumlah populasi (*population_size*), iterasi maksimum (*max_iteration*), probabilitas mutasi centroid (*mutation_rate*), probabilitas mutasi kromosom (*exposure_rate*), jumlah parent yang dipertahankan dan dikawinkan (*num_selection*).
2. Menghitung Radial Basis Function untuk melatih data pelatihan dengan cara menghitung fungsi gaussian dahulu untuk terhubung ke *hidden layers*, setelah itu menghitung pseudo-invers untuk terhubung ke *output layers* dan kemudian nilai dari *output layers* diklasifikasi dengan menggunakan fungsi aktivasi.
3. Menghitung nilai fitness dengan menggunakan *fl score* dari setiap individu yang telah mendapatkan nilai output dari Radial Basis Function dalam populasi.
4. Seleksi semua individu dalam populasi menjadi subpopulasi induk sesuai dengan nilai parameter.
5. Melakukan pindah silang pada subpopulasi induk dengan pasangan induk dipilih secara acak untuk menghasilkan subpopulasi individu anak (*offspring*).
6. Melakukan proses mutasi terhadap individu anak dan juga pada centroid secara acak berdasarkan probabilitas mutasi.
7. Cek kondisi akhir apakah sudah terpenuhi, jika tidak ada maka kembali ke prosedur 2 dengan generasi gabungan antara subpopulasi induk dengan subpopulasi anak. Kriteria terpenuhinya kondisi akhir adalah ada satu individu yang mencapai batasan dalam perolehan nilai fitness.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Parameter Algoritma Genetika

Pada penelitian ini, algoritma genetika digunakan untuk mengestimasi serta mengoptimalkan parameter pada model radial basis function. Untuk itu parameter algoritma genetika yang digunakan dalam penelitian ini terdapat pada tabel 1.

Tabel 1 Parameter algoritma genetika

No	Parameters	Value
1	Population	500
2	Iteration	10
3	Mutation Rate	0.4
4	Exposure Rate	0.25
5	Number Parent	250

Berdasarkan tabel 3, parameter dalam algoritma genetika ini dipilih berdasarkan eksperimen awal dengan beberapa konfigurasi untuk menemukan pengaturan yang memberikan hasil terbaik.

4.2 Analisa Sistem (Manual)

Untuk memahami implementasi jaringan saraf tiruan Radial Basis Function secara manual, pada subbab ini akan dijelaskan perhitungan menggunakan studi kasus sederhana. Studi kasus ini bertujuan untuk menunjukkan langkah-langkah perhitungan RBF.

Misalkan terdapat 4 data pelatihan secara manual dengan 6 fitur atau kolom yang digunakan untuk mengklasifikasikan *output* dari kolom *fraud*. Data pelatihan dapat dilihat pada gambar 6.

income	credit_risk_score	customer_age	payment_type	employment_status	current_address_months_count	fraud_bool
0.7	-1.168150	-0.600278	0.170667	0.762667	-0.397838	1
0.8	0.517794	0.178629	0.350667	0.064000	0.841240	1
0.1	-0.860496	-1.379185	0.372000	0.762667	-1.071723	0
0.7	-0.171351	-0.600278	0.372000	0.126667	-0.408708	0

Gambar 6 Data pelatihan manual

Langkah 1, Menginisialisasi parameter Radial Basis Function dengan 2 *centroid* dan 1 *spread*, sehingga pada *hidden layers* memiliki 2 node, yaitu:

$$Centroid_1 = [0.75, -0.325178, -0.21082, 0.260667, 0.4133335, 0.221701]$$

$$Centroid_2 = [0.4, -0.515924, -0.98973, 0.372, 0.444667, -0.7402155]$$

$$Spread = 0.5$$

Langkah 2, Memperbarui bobot pada *hidden layers* dengan fungsi gaussian (1).

$$G(\|x - \mu_i\|) = \exp\left(-\frac{\|x - \mu_i\|^2}{2\sigma^2}\right)$$

Fungsi tersebut diimplementasikan pada data pelatihan dengan inisialisasi centroidnya, sehingga diperoleh

$$\begin{aligned} G_{11} &= \exp\left(-\frac{(0.7-0.75)^2+(-1.168+0.325)^2+(-0.6+0.21)^2+(0.17-0.26)^2+(0.762-0.413)^2+(-0.397+0.221)^2}{2(0.5)^2}\right) \\ &= 0.063452 \\ G_{21} &= \exp\left(-\frac{(0.8-0.75)^2+(0.517+0.325)^2+(0.178+0.21)^2+(0.35-0.26)^2+(0.064-0.413)^2+(0.841+0.221)^2}{2(0.5)^2}\right) \\ &= 0.063452 \\ G_{31} &= \exp\left(-\frac{(0.1-0.75)^2+(-0.86+0.325)^2+(-1.379+0.21)^2+(0.372-0.26)^2+(0.762-0.413)^2+(-1.071+0.221)^2}{2(0.5)^2}\right) \\ &= 0.000425 \\ G_{41} &= \exp\left(-\frac{(0.7-0.75)^2+(-0.171+0.325)^2+(-0.6+0.21)^2+(0.372-0.26)^2+(0.1266-0.413)^2+(-0.408+0.221)^2}{2(0.5)^2}\right) \\ &= 0.261937 \\ G_{12} &= \exp\left(-\frac{(0.7-0.4)^2+(-1.168+0.515)^2+(-0.6+0.989)^2+(0.17-0.372)^2+(0.762-0.444)^2+(-0.397+0.74)^2}{2(0.5)^2}\right) \\ &= 0.1569372 \\ G_{22} &= \exp\left(-\frac{(0.8-0.4)^2+(0.517+0.515)^2+(0.178+0.989)^2+(0.35-0.372)^2+(0.064-0.444)^2+(0.841+0.74)^2}{2(0.5)^2}\right) \\ &= 0.000028 \\ G_{32} &= \exp\left(-\frac{(0.1-0.4)^2+(-0.86+0.515)^2+(-1.379+0.989)^2+(0.372-0.372)^2+(0.762-0.444)^2+(-1.071+0.74)^2}{2(0.5)^2}\right) \\ &= 0.318908 \\ G_{42} &= \exp\left(-\frac{(0.7-0.4)^2+(-0.171+0.515)^2+(-0.6+0.989)^2+(0.372-0.372)^2+(1.266-0.444)^2+(-0.408+0.74)^2}{2(0.5)^2}\right) \\ &= 0.318908 \end{aligned}$$

Sehingga diperoleh matriks gaussiannya, sebagai berikut:

	<i>Centroid₁</i>	<i>Centroid₂</i>	<i>Bias</i>
$G =$	0.063452	0.156937	1
	0.063452	0.000028	1
	0.000425	0.318908	1
	0.261937	0.318908	1

Langkah 3, Menghitung pseudo-invers matriks gaussian pada *hidden layers* (2).

$$(G^T G) \rightarrow (G^T G)^{-1}$$

Dari matriks gaussian pada langkah 2, maka diperoleh

$$G^T = \begin{pmatrix} 0.063452 & 0.063452 & 0.000425 & 0.261937 \\ 0.156937 & 0.000028 & 0.318909 & 0.318909 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Sehingga dapat diperoleh matriks $G^T G$, sebagai berikut

$$G^T G = \begin{pmatrix} 0.063452 & 0.063452 & 0.000425 & 0.261937 \\ 0.156937 & 0.000028 & 0.318909 & 0.318909 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} Centroid_1 & Centroid_2 & Bias \\ 0.063452 & 0.156937 & 1 \\ 0.063452 & 0.000028 & 1 \\ 0.000425 & 0.318908 & 1 \\ 0.261937 & 0.318908 & 1 \end{pmatrix}$$

$$G^T G = \begin{pmatrix} 0.076664 & 0.093629 & 0.389266 \\ 0.093629 & 0.228035 & 0.794782 \\ 0.389266 & 0.794782 & 4 \end{pmatrix}$$

Dan juga dapat diperoleh maktris $(G^T G)^{-1}$, sebagai berikut

$$(G^T G)^{-1} = \begin{pmatrix} 28.5716 & -6.6356 & -1.4620 \\ -6.6356 & 15.8034 & -2.4943 \\ -1.4620 & -2.4943 & 0.8878 \end{pmatrix}$$

Langkah 4, Memperbarui bobot yang terhubung pada *output layer* (3)

$$W = (G^T G)^{-1} (G^T d)$$

Dari Langkah 1-3, dapat dihitung nilai bobot pada *output layer* sehingga diperoleh,

$$G^T d = \begin{pmatrix} 0.063452 & 0.063452 & 0.000425 & 0.261937 \\ 0.156937 & 0.000028 & 0.318909 & 0.318909 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$G^T d = \begin{pmatrix} 0.1269 \\ 0.15697 \\ 2 \end{pmatrix}$$

Maka diperoleh,

$$W = \begin{pmatrix} 28.5716 & -6.6356 & -1.4620 \\ -6.6356 & 15.8034 & -2.4943 \\ -1.4620 & -2.4943 & 0.8878 \end{pmatrix} \begin{pmatrix} 0.1269 \\ 0.15697 \\ 2 \end{pmatrix} = \begin{pmatrix} -0.33976 \\ -3.35012 \\ 1.19872 \end{pmatrix} \rightarrow \begin{matrix} w_1 \\ w_2 \\ bias \end{matrix}$$

Langkah 5, Menghitung *output* jaringan Radial Basis Function dengan rumus sebagai berikut:

$$Y_i = w_1 G_{i1} + w_2 G_{i2} + bias$$

Sehingga diperoleh nilai *output* jaringan RBFnya,

$$Y_1 = (-0.33976)(0.063452) + (-3.35012)(0.156972) + 1.19872 = 0.6514$$

$$Y_2 = (-0.33976)(0.063452) + (-3.35012)(0.000028) + 1.19872 = 1.177$$

$$Y_3 = (-0.33976)(0.000425) + (-3.35012)(0.318908) + 1.19872 = 0.1301$$

$$Y_4 = (-0.33976)(0.261937) + (-3.35012)(0.318908) + 1.19872 = 0.0414$$

Langkah 6, Menghitung fungsi aktivasi pada *output* jaringan Radial Basis Function, dengan fungsi aktivasi sebagai berikut,

$$f(Y_i) = \begin{cases} 1, & Y_i \geq 0.5 \\ 0, & Y_i < 0.5 \end{cases}$$

Sehingga, hasil dari langkah 1-5 ada pada tabel 2.

Tabel 2 Hasil Pelatihan Manual RBF

No	Y_i	$f(Y_i)$	Fraud_bool
1	0.6514	1	1
2	1.1771	1	1
3	0.1301	0	0
4	0.0414	0	0

Berdasarkan tabel 2, hasil pelatihan data secara manual menggunakan jaringan Radial Basis Function menghasilkan tingkat akurasi **100%**. Selanjutnya dari bobot yang diperoleh pada pelatihan data secara manual tersebut akan di uji pada data validasi yang ada pada gambar 7.

X1	X2	X3	X4	X5	X6	fraud_bool
0,200000	1,526899	-0,600278	0,372000	0,762667	2,047711	1
0,400000	0,370120	0,178629	0,350667	0,762667	0,330392	1
0,600000	-0,860496	-1,242234	0,356000	0,768000	-0,439216	0
0,200000	-0,614373	-0,600278	0,372000	0,762667	-0,332624	0

Gambar 7 Data validasi manual

Langkah 1, Menggunakan parameter jaringan Radial Basis Function yang sama dengan parameter pada pelatihan, yaitu

$$Centroid_1 = [0.75, -0.325178, -0.21082, 0.260667, 0.4133335, 0.221701]$$

$$Centroid_2 = [0.4, -0.515924, -0.98973, 0.372, 0.444667, -0.7402155]$$

$$Spread = 0.5$$

Langkah 2, Memperbarui bobot pada *hidden layers* dengan fungsi gaussian.

$$G_{11} = \exp\left(-\frac{(0.2-0.75)^2+(1.526+0.325)^2+(-0.6+0.21)^2+(0.372-0.26)^2+(0.762-0.413)^2+(2.047+0.221)^2}{2(0.5)^2}\right)$$

$$= 0.00000$$

$$G_{21} = \exp\left(-\frac{(0.4-0.75)^2+(0.3701+0.325)^2+(0.178+0.21)^2+(0.35-0.26)^2+(0.762-0.413)^2+(0.3303+0.221)^2}{2(0.5)^2}\right)$$

$$= 0.165444$$

$$G_{31} = \exp\left(-\frac{(0.6-0.75)^2+(-0.86+0.325)^2+(-1.2422+0.21)^2+(0.356-0.26)^2+(0.768-0.413)^2+(-0.439+0.221)^2}{2(0.5)^2}\right)$$

$$\begin{aligned}
&= 0.020463 \\
G_{41} &= \exp\left(-\frac{(0.2-0.75)^2+(-0.614+0.325)^2+(-0.6+0.21)^2+(0.372-0.26)^2+(0.762-0.413)^2+(-0.332+0.221)^2}{2(0.5)^2}\right) \\
&= 0.140996 \\
G_{12} &= \exp\left(-\frac{(0.2-0.4)^2+(1.526+0.515)^2+(-0.6+0.989)^2+(0.372-0.372)^2+(0.762-0.444)^2+(2.047+0.74)^2}{2(0.5)^2}\right) \\
&= 0.00000 \\
G_{22} &= \exp\left(-\frac{(0.4-0.4)^2+(0.37+0.515)^2+(0.178+0.989)^2+(0.35-0.372)^2+(0.762-0.444)^2+(-0.330+0.74)^2}{2(0.5)^2}\right) \\
&= 0.001118 \\
G_{32} &= \exp\left(-\frac{(0.6-0.4)^2+(-0.86+0.515)^2+(-1.242+0.989)^2+(0.356-0.372)^2+(0.768-0.444)^2+(-0.439+0.74)^2}{2(0.5)^2}\right) \\
&= 0.433637 \\
G_{42} &= \exp\left(-\frac{(0.2-0.4)^2+(-0.614+0.515)^2+(-0.6+0.989)^2+(0.372-0.372)^2+(0.762-0.444)^2+(-0.332+0.74)^2}{2(0.5)^2}\right) \\
&= 0.391707
\end{aligned}$$

Sehingga diperoleh matriks gaussiannya, sebagai berikut:

	<i>Centroid₁</i>	<i>Centroid₂</i>
	0	0
$G =$	0.16544	0.00111
	0.02046	0.43353
	0.14099	0.39170

Langkah 3, Menghitung *output* jaringan Radial Basis Function dengan rumus sebagai berikut:

$$Y_i = w_1 G_{i1} + w_2 G_{i2} + bias$$

Sehingga diperoleh nilai *output* jaringan RBFnya,

$$Y_1 = (-0.33976)(0.000000) + (-3.35012)(0.000000) + 1.19872 = 1.1987$$

$$Y_2 = (-0.33976)(0.165444) + (-3.35012)(0.001118) + 1.19872 = 1.1387$$

$$Y_3 = (-0.33976)(0.020463) + (-3.35012)(0.433537) + 1.19872 = -0.2606$$

$$Y_4 = (-0.33976)(0.140996) + (-3.35012)(0.391707) + 1.19872 = -0.1614$$

Langkah 4, Menghitung fungsi aktivasi pada *output* jaringan Radial Basis Function, dengan fungsi aktivasi sebagai berikut,

$$f(Y_i) = \begin{cases} 1, & Y_i \geq 0.5 \\ 0, & Y_i < 0.5 \end{cases}$$

Sehingga, hasil dari pengujian ada pada tabel 3.

Tabel 3 Hasil Pengujian Manual RBF

No	Y_i	$f(Y_i)$	Fraud_bool
1	1.19871	1	1
2	1.1387	1	1
3	-0.26063	0	0
4	-0.16125	0	0

Berdasarkan tabel 3, hasil pengujian data secara manual menggunakan jaringan Radial Basis Function dengan centroid, spread dan bobot dari hasil pelatihan data secara manual menghasilkan tingkat akurasi **100%**.

4.3 Analisa Sistem (Program)

Berdasarkan sistem yang dibuat untuk mengklasifikasikan penipuan kartu kredit menggunakan metode Radial Basis Function yang dioptimasi oleh Algoritma Genetika pada 1000 data dengan 75% data pelatihan dan 25% data pengujian, hasil dari keseluruhan evaluasi model Radial Basis Function ditunjukkan pada tabel 4.

Tabel 4 Hasil Evaluasi RBF menggunakan *F1 Score*

Iterasi	Pelatihan		Pengujian	
	<i>f1 score</i>	<i>Accuracy</i>	<i>f1 score</i>	<i>Accuracy</i>
1	0.7836	0.7707	0.7983	0.8040
2	0.8430	0.8387	0.8083	0.8160
3	0.9270	0.9253	0.8898	0.8960
4	0.9308	0.9293	0.8870	0.8920
5	0.9322	0.9307	0.8823	0.8786
6	0.9352	0.9333	0.8786	0.8464
7	0.9076	0.8973	0.8464	0.8520
8	1.0000	0.9760	0.9333	0.9360
9	1.0000	0.9587	0.9121	0.9160
10	0.9986	0.9440	0.9243	0.9280

Berdasarkan tabel 4. Hasil evaluasi model Radial Basis Function dengan parameter yang dioptimasi oleh algoritma genetika menunjukkan hasil yang cukup baik pada setiap iterasinya mengalami kenaikan dan penurunan pada *f1 score* namun model tetap menemukan parameter terbaiknya, artinya performa terbaik model tercapai ketika iterasi ke-8 dengan *f1 score* dan *accuracy* pada data pelatihan sebesar 1.00 dan 0.9760 sedangkan *f1 score* dan *accuracy* pada data pengujian sebesar 0.9333 dan 0.9360.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil dari penerapan metode Radial Basis Function dengan parameter yang dioptimasi oleh Algoritma Genetika, dapat disimpulkan bahwa:

1. Penerapan model Radial Basis Function (RBF) yang dioptimasi dengan Algoritma Genetika dilakukan melalui inisialisasi parameter, pelatihan RBF dengan menghitung fungsi Gaussian dan pseudo-invers, serta klasifikasi menggunakan fungsi aktivasi. Evaluasi performa dilakukan menggunakan skor F1 sebagai metrik utama untuk menilai kualitas deteksi penipuan. Algoritma Genetika digunakan untuk mengoptimasi parameter model melalui proses seleksi, crossover, dan mutasi pada parameter centroid dan spread, hingga tercapai nilai fitness yang memenuhi kriteria penghentian.
2. Model Radial Basis Function (RBF) yang dioptimasi dengan Algoritma Genetika mampu mendeteksi penipuan kartu kredit secara efektif. Proses optimasi menggunakan Algoritma Genetika menghasilkan peningkatan *f1 score* pada setiap iterasi, dengan performa terbaik dicapai pada iterasi ke-8. Pada tahap ini, model memperoleh *f1 score* dan *accuracy* sebesar 1.00 dan 0.9760 pada data pelatihan, sedangkan *f1 score* dan *accuracy* sebesar 0.9333 dan 0.9360 pada data pengujian.

Model Radial Basis Function yang dioptimasi dengan Algoritma Genetika dapat diterapkan secara efektif untuk mendeteksi penipuan kartu kredit. Model ini menawarkan pendekatan yang kuat untuk menangani pola data yang kompleks dan menghasilkan performa deteksi yang baik berdasarkan *f1 score*, yang menjadi metrik evaluasi utama.

5.2 Saran

Berdasarkan hasil penelitian dan kesimpulan yang telah diuraikan, salah satu saran untuk pengembangan lebih lanjut adalah mengeksplorasi metode optimasi alternatif, seperti *Particle Swarm Optimization* (PSO) atau *Differential Evolution* (DE). Penerapan metode ini dapat menjadi langkah strategis untuk membandingkan efektivitasnya dengan Algoritma Genetika dalam mengoptimasi parameter model Radial Basis Function. Dengan membandingkan kinerja berbagai pendekatan, penelitian lanjutan dapat memberikan wawasan lebih mendalam tentang metode optimasi yang paling efisien dan akurat, sehingga dapat meningkatkan performa model dalam mendeteksi penipuan kartu kredit.

DAFTAR PUSTAKA

- [1] Ardianto R, Ramdhani RF, Apriliana Dewi LO, Prabowo A, Saputri YW, Lestari AS, Hadi N. Transformasi Digital dan Antisipasi Perubahan Ekonomi Global dalam Dunia Perbankan. MARAS: Jurnal Penelitian Multidisiplin. 2024;2(1):80–88. doi:10.60126/maras.v2i1.114
- [2] Mijwil MM, Aljanabi M, ChatGPT. Towards Artificial Intelligence-Based Cybersecurity: The Practices and ChatGPT Generated Ways to Combat Cybercrime. Iraqi Journal for Computer Science and Mathematics. 2023;4(1):65–70. doi:10.52866/ijcsm.2023.01.01.0019
- [3] Ghai, V., Kang, S.S. (2021), Role of Machine Learning in Credit Card Fraud Detection. In: 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N). IEEE. p939-943.
- [4] Minastireanu, E.A., Mesnita, G. (2019), An analysis of the most used machine learning algorithms for online fraud detection. Informatica Economica, 23(1), 5-16.
- [5] Widhiastuti, N. L. P., and Kumalasari, P. D. 2020. Kemampuan Mahasiswa Dalam Mendeteksi Fraud. Jurnal Riset Akuntansi Dan Bisnis Airlangga. Vol. 5 No. 1. Pp.762-783.
- [6] Raji ID, Bello-Salau H, Umoh IJ, Onumanyi AJ, Adegboye MA, Salawudeen AT. Simple Deterministic Selection-Based Genetic Algorithm for Hyperparameter Tuning of Machine Learning Models. Applied Sciences (Switzerland). 2022;12(3). doi:10.3390/app12031186
- [7] Dhirima, Silvia Anisa; Tuti Susilawati dan Mahipal. 2019. “Kajian Penerapan Prinsip Kehati-hatian Dalam Penerbitan Kartu Kredit Pada Bank Syariah”. Jurnal Cendekia Ihya2(1), 30-42.
- [8] Lestari, Bunga Ayu; Budi Suharjo dan Istiqlaliyah Muflikhati. 2017. “Minat Kepemilikan Kartu Kredit (Studi Kasus Kota Bogor)”. Jurnal Aplikasi Bisnis dan Manajemen3 (1), 143-151.
- [9] Tokopedia. 2019. 10 Manfaat Kartu Kredit yang Memudahkan Urusan Finansial. Diakses dari: <https://www.tokopedia.com/blog/fin-manfaat-kartu-kredit-keuntungan/>, tanggal 3 Desember 2024.
- [10] Fraud.com. (n.d.). *Bank fraud - Types and examples*. Diakses dari: <https://www.fraud.com/post/bank-fraud>, tanggal 3 Desember 2024
- [11] Cynthia, Eka Pandu, and Edi Ismanto. "Jaringan Syaraf Tiruan Algoritma Backpropagation Dalam Memprediksi Ketersediaan Komoditi Pangan Provinsi Riau." Rabit: Jurnal Teknologi dan Sistem Informasi Univrab 2, no. 2 (2017): 83-98.
- [12] Yazici, İ., Shaye, I., & Din, J. (2023). A survey of applications of artificial intelligence and machine learning in future mobile networks-enabled systems. In *Engineering Science and Technology, an International Journal* (Vol. 44). Elsevier B.V. <https://doi.org/10.1016/j.jestch.2023.101455>.
- [13] Wu Y, Wang H, Zhang B, Du K-L. Using Radial Basis Function Networks for Function Approximation and Classification. ISRN Applied Mathematics. 2012; 2012:1–34. doi:10.5402/2012/324194
- [14] Poggio T, Girosi F. Networks for approximation and learning. Proceedings of the IEEE. 1990;78(9):1481–1497. doi:10.1109/5.58326
- [15] Windarto. (2015). *Modul Perkuliahan Simulasi Pengantar Algoritma Genetika*.
- [16] D.E. Goldberg, 1989, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley Publishing Company, Inc.
- [17] Clara, S., Laksmi Prianto, D., al Habsi, R., Friscila Lumbantobing, E., Chamidah, N., Kom, S., Kom, M., Informatika, J., Ilmu Komputer, F., Pembangunan Nasional Veteran Jakarta Jl Fatmawati Raya, U. R., Labu, P., Cilandak, K., & Depok, K. (2021).

Implementasi Seleksi Fitur Pada Algoritma Klasifikasi Machine Learning Untuk Prediksi Penghasilan Pada Adult Income Dataset. In *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta-Indonesia*.

- [18] Alghifari, F., & Juardi, D. (2021). *Fauzan Alghifari Penerapan Data Mining Pada Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes*.