

Optimalisasi Metode Naive Bayes Classifier Untuk Prediksi Persetujuan Kredit**Achmad Syakur¹, Rendri Purwandi Putra², Christina Juliane³**ac.syakur@stikesmi.ac.id, rendripurwandi@stikesmi.ac.id, christina.juliane@likmi.ac.id^{1,2,3} STMIK LIKMI Bandung^{1,2} Sekolah Tinggi Ilmu Kesehatan Sukabumi

Informasi Artikel	Abstrak
Diterima : (kosongkan) Direview : (kosongkan) Disetujui : (kosongkan)	Kredit adalah bentuk pembiayaan yang banyak orang ajukan ke bank atau perusahaan penyedia kredit. Dalam proses pengajuan kredit, dilakukan analisis untuk menentukan apakah kredit yang diajukan layak atau tidak. Penelitian ini bertujuan untuk membantu bank atau perusahaan penyedia kredit dalam melakukan persetujuan kredit dengan efektif dan akurat dalam menentukan status pengajuan. Penelitian ini menggunakan teknik data mining dan kumpulan dataset yang berasal dari kaggle.com. Terdapat 12 atribut dan 2 kelas yang digunakan dalam penelitian ini. Dalam penelitian ini, metode klasifikasi Naive Bayes dan optimasi kelompok partikel (PSO) digunakan. Prediksi persetujuan kredit dengan metode naïve bayes classifier menghasilkan nilai akurasi sebesar 80,00% dengan nilai AUC 0,884. Sebaliknya, prediksi persetujuan kredit dengan metode particle swarm optimization (PSO) menghasilkan nilai akurasi sebesar 96,67% dengan nilai AUC 0,69.
Kata Kunci Persetujuan Kredit, Data Mining, Klasifikasi Naïve Bayes	

Keywords	Abstract
Credit Approval, Data Mining, Naïve Bayes Classifier	<i>Credit is a form of financing that many people apply to banks or credit provider companies. In the credit application process, an analysis is carried out to determine whether the proposed credit is feasible or not. This research aims to assist banks or credit provider companies in approving credit effectively and accurately in determining the status of the application. This research uses data mining techniques and a collection of datasets from kaggle.com. This study employed two classes and twelve attributes. In this study, Particle Swarm Optimization (PSO) and the Naive Bayes Classifier approach are applied. The accuracy value of 80.00% and AUC value of 0.884 are obtained when credit approval is predicted using the naïve bayes classifier approach. In the meantime, an AUC value of 0.969 and an excellent accuracy of 96.67% are produced when the naïve bayes classifier with Particle Swarm Optimization (PSO) optimization is used to forecast credit approval.</i>

A. Pendahuluan

Kredit adalah bentuk pembiayaan yang banyak digunakan dalam kegiatan ekonomi [10]. Permintaan pinjaman saat ini meningkat pesat karena kebutuhan keuangan masyarakat yang semakin tinggi, terutama di negara-negara berkembang salahsatunya negara Indonesia. Pemeriksaan kredit diperlukan untuk pembiayaan yang tepat dan aman [6]. Jika nasabah tidak melunasi pinjamannya dengan lancar maka merugikan perusahaan [8].

Menentukan kelayakan kredit perusahaan dan risiko kredit merupakan topik penting yang telah memicu perdebatan akademis baik secara teoritis maupun empiris, dan sangat relevan dengan industri dan komunitas bisnis [4].

Saat ini, pemberian kredit kepada calon kreditur merupakan hal yang umum dilakukan untuk membantu memenuhi kebutuhan. Namun, dalam menentukan kelayakan kredit, terdapat banyak pertimbangan yang harus dipertimbangkan sebelum calon kreditur atau nasabah dapat dinyatakan layak dalam mengambil kredit [5].

Peminjaman merupakan proses yang sangat penting bagi manajer kredit. Salah memilih penerima pinjaman akan mengakibatkan tagihan pinjaman tidak terbayar [11]. untuk menghindari itu, diperlukan suatu analisis data calon kreditu untuk menentukan apakah pinjaman layak atau tidak dan mengklasifikasikannya [7].

Banyak penelitian telah dilakukan untuk menganalisis kelayakan pemberian kredit kepada calon kreditur. Metode klasifikasi yang digunakan untuk memprediksi adalah metode klasifikasi naïve bayes classifier [5]. Pengklasifikasi Naive Bayes adalah algoritma probabilistik yang secara langsung didasarkan pada teorema Bayes dan mengasumsikan independensi antar fitur. Algoritma ini digunakan untuk klasifikasi dan efektif untuk kumpulan data besar. Oleh karena itu, sangat cocok untuk memprediksi persetujuan pinjaman [9]. Metode Naive Bayes Classifier dipilih karena mudah diimplementasikan, cepat, dan memiliki akurasi yang tinggi. Namun, metode ini menggunakan seleksi fitur yang dapat menghasilkan berkurangnya nilai akurasi. Oleh karena itu, pemilihan fitur perlu dioptimalkan; Particle Swarm Optimization bertujuan untuk menghasilkan hasil pengklasifikasi Naive Bayes yang lebih akurat [3].

B. Metode Penelitian

1. Credit Approval

Persetujuan kredit adalah suatu proses yang dilakukan oleh bank atau perusahaan penyedia kredit. Proses ini didasarkan pada permintaan kredit dan proposal kredit dari peminjam. Persetujuan kredit sering kali sulit bagi bank atau perusahaan penyedia kredit karena jumlah permintaan dan klasifikasi harus dilakukan pada berbagai data yang diserahkan [1].

2. Data Mining

Data mining adalah sebuah proses pengolahan data dalam jumlah banyak dan diambil dalam sebuah database dan belum diketahui sebelumnya, akan tetapi data tersebut dapat dipahami dan bermanfaat untuk membuat keputusan bisnis yang sangat penting [14].

3. Naïve Bayes Classifier

Pengklasifikasi Naive Bayes telah banyak digunakan di banyak aplikasi. Untuk meningkatkan kinerja klasifikasi, pengklasifikasi Naïve Bayes telah dikembangkan [13]. Rumus dari metode Naïve Bayes Classifier sebagai berikut :

$$P(C_k|X) = \frac{P(X)P(X|C_k)}{P(C_k)}$$

Dimana C_k adalah kelas target, X adalah vektor fitur, $P(C_k|X)$ adalah probabilitas posteriori, $P(C_k)$ adalah probabilitas priori, dan $P(X|C_k)$ adalah likelihood.

4. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) adalah metode optimasi populasi yang mirip dengan genetic algorithm (GA), tetapi dasar dari metode ini adalah bahwa itu bekerja sama, bukan kompetitif. Beberapa metode untuk optimalisasi dataset dalam metode PSO termasuk menseleksi atribut (attribute selection), meningkatkan nilai atribut terhadap setiap atribut atau variabel yang digunakan, serta pemilihan atribut [12].

5. Confusion matrix

Confusion matrix merupakan suatu bagan yang memperlihatkan hasil klasifikasi pada suatu dataset. Beberapa rumus yang dapat dipakai untuk mengukur kinerja klasifikasi tersebut. Hasil dari akurasi, presisi, dan recall dapat disajikan dalam bentuk persentase [2].

Tabel 1. Tabel Evaluasi Confusion Matrix

	Class = Yes	Class = No
Class = Approved	(True Positive-TP)	(False Negative-FN)
Class = Rejected	(False Positive-FP)	(True Negative-TN)

a) Accuracy

Accuracy adalah jumlah proporsi prediksi yang benar. Accuracy dapat dihitung menggunakan formula berikut :

$$\frac{TP + TN}{TP + TN + FP + FN}$$

Dimana :

TP : umlah kasus positif yang teridentifikasi dengan benar sebagai positif

FP : jumlah kasus negatif yang salah diklasifikasikan sebagai positif.

TN : jumlah kasus negatif yang berhasil diklasifikasikan dengan benar sebagai negatif

FN : jumlah kasus positif yang salah diklasifikasikan sebagai negatif

b) Precision

Precision adalah perbandingan antara jumlah dokumen teks yang relevan dan terkendali dengan total dokumen teks yang dipilih oleh sistem.

c) Recall

Recall adalah proporsi antara jumlah dokumen teks yang terkendali dan seluruh dokumen teks relevan dalam kumpulan data.

C. Hasil dan Pembahasan**1. Instrumen Penelitian**

Berikut adalah beberapa elemen penting yang digunakan dalam penelitian ini:

a) Dataset Kaggle.com yang terdiri dari dua belas atribut dan dua kelas—diterima atau ditolak—digunakan untuk penelitian ini.

b) Informasi disajikan dalam bentuk tabel yang menggambarkan model dan variabel, dengan total 1003 data.

c) Perangkat lunak yang dipakai untuk melakukan analisis adalah Rapid Miner Studio versi 8.2.

d) Metode yang digunakan dalam prediksi persetujuan kredit adalah metode naïve bayes classifier dan metode naïve bayes classifier berbasis particle swarm optimization (PSO).

2. Evaluasi dan Validasi Model

Hasil dari pengujian model yang dilakukan yaitu prediksi credit approval dengan menggunakan metode Naive Bayes Classifier dan metode Naive Bayes Classifier dengan optimasi Particle Swarm Optimization (PSO) untuk menentukan nilai accuracy dan AUC.

a) Hasil Pengujian Menggunakan metode Naive Bayes Classifier

Dataset terdiri dari dua belas atribut prediksi untuk persetujuan kredit dan kelas. Kelas ini adalah hasil akhir prediksi yang digunakan dalam pembuatan model Naive Bayes. Data kemudian divalidasi untuk mempercepat proses prediksi dan memungkinkan pelatihan. Mencari probabilitas hipotesis untuk setiap Kelas $P(H)$ adalah langkah pertama dalam menerapkan model Naive Bayes. Dataset yang digunakan terdiri dari 1003 daftar, dengan 502 daftar kelas yang diterima dan 501 daftar kelas yang ditolak. Perhitungan probabilitas untuk dua kelas yaitu :

$$P(X|\text{Approved}) = 502 : 1003 = 0,5004985044865$$

$$P(X|\text{Rejected}) = 501 : 1063 = 0,4995014955143$$

Setelah mengetahui kemungkinan untuk setiap hipotesis, langkah berikutnya adalah menentukan kemungkinan kondisi tertentu (probabilitas X) berdasarkan kemungkinan masing-masing hipotesis (probabilitas H), yang juga dikenal sebagai probabilitas prior. Hasil penelitian adalah nilai akurasi dan nilai AUC (Area Under Curve).

1) Evaluasi model dengan Confusion Matrix

Untuk menghasilkan hasil yang diinginkan, model confusion matrix akan membentuk suatu tabel berukuran dua kali dua yang berisi true positif, atau tupel positif, dan true negatif, atau tupel negatif. Data ini kemudian dimasukkan ke dalam confusion matrix, yang menghasilkan hasil seperti yang ditunjukkan dalam tabel berikut:

Tabel 2. Confusion Matrix metode Naive Bayes Classifier

	True Approved	True Rejectes	Class Precision
Pred. Approved	32	2	94.12%
Pred. Rejected	18	48	72.73%
Class Recall	64.00%	96.00 %	

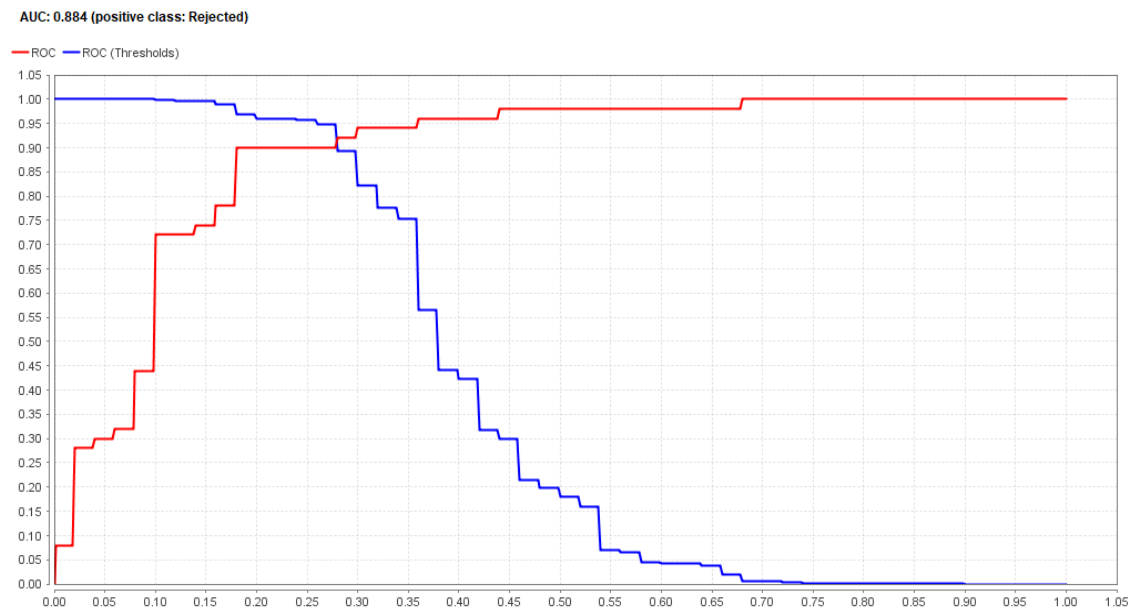
Berdasar tabel diatas prediksi credit approval menggunakan Naïve Bayes Classifier dengan optimasi PSO menghasilkan akurasi sebesar 80% dan terdapat rincian jumlah True Positive (TP) 32, False Negative (FN) 48, False Positive (FP) adalah 18 dan True Negative (TN) 2. Data tersebut dapat dihitung manual dengan perhitungan sebagai berikut :

$$\text{Accuracy} = \frac{32+48}{32+48+18+2} = 0.80$$

Berdasarkan perhitungan manual diatas, maka akurasi yang dihasilkan sebesar 80%.

2) Evaluasi Dengan ROC Curve

Hasil pengujian terhadap data testing untuk metode Naive Bayes Classifier terhadap nilai ROC diketahui pada gambar di bawah ini:

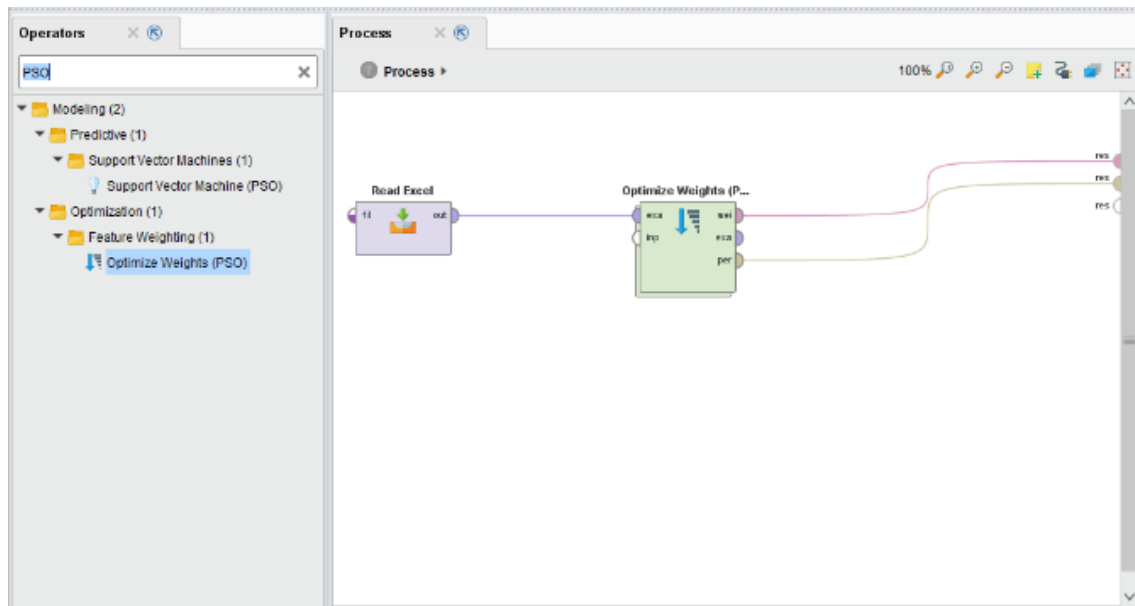
**Gambar 1.** Nilai AUC metode naive bayes dalam Grafik ROC

Nilai AUC sebesar 0,884 untuk evaluasi metode Naive Bayes Classifier menunjukan nilai Fair Classification.

b) Hasil Pengujian Menggunakan metode Naive Bayes Classifier dengan optimasi PSO

Pemodelan metode Naive Bayes dengan optimasi Particle swarm optimization (PSO) dilakukan sebagai berikut:

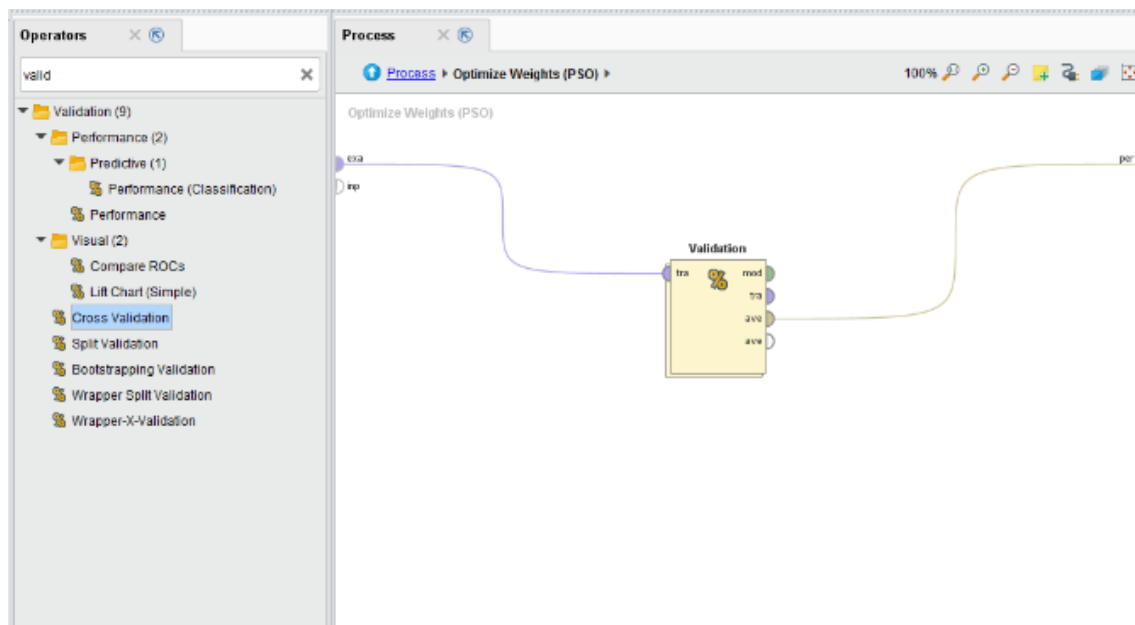
1) Berdasarkan dataset yang sama seperti di atas lakukan Langkah



Gambar 2. Pemilihan Optimize Selection

Pada Tab Operator, pilih Data Transformation lalu Attribute Set Reduction and Transformation kemudian Selection – Optimization – Optimize Weight (PSO).

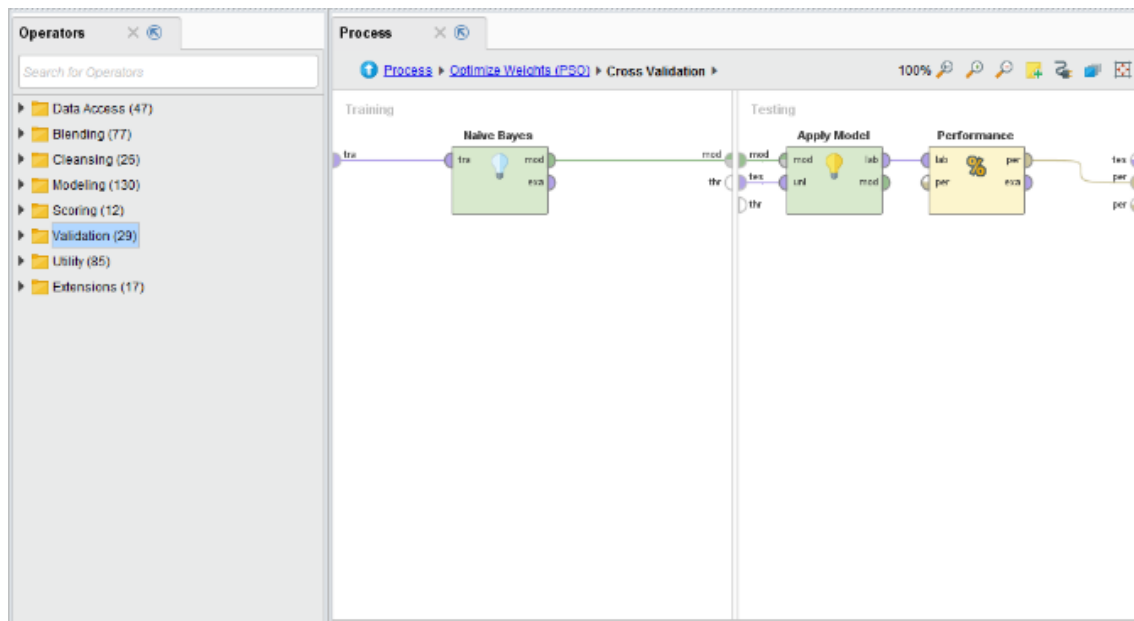
2) Pada tahap evaluation process diawali dari double klik Optimize Selection pada icon kotak warna biru kemudian pilih Evaluation - validation – Cross Validation simpan di main process dari Validation



Gambar 3. Pemilihan Validation

3) Lalu pilih model algoritma yang akan digunakan, pada penelitian ini adalah model naive bayes (Modelling – Predictive – Bayesian – Naive Bayes) lalu drag ke bagian Training. Sedangkan pada bagian Testing pilih Apply Model (Modeling –

Model Application – Confidences – Apply Model) dan Performance (Validation – Performance - Performance)



Gambar 4. Pilih Model Algoritma dan *Performance*

4) Evaluasi model dengan Confusion Matrix

Hasil dari uji coba ini mencakup perolehan nilai akurasi dan nilai AUC (Area Under Curve) dengan menerapkan metode Naive Bayes Classifier yang menggunakan optimisasi berbasis PSO sebagai berikut:

Tabel 3. Confusion Matrix metode Naive Bayes Classifier berbasis PSO

	True Approved	True Rejectes	Class Precision
Pred. Approved	14	0	100.00%
Pred. Rejected	1	15	93.75%
Class Recall	93.33%	100.00%	

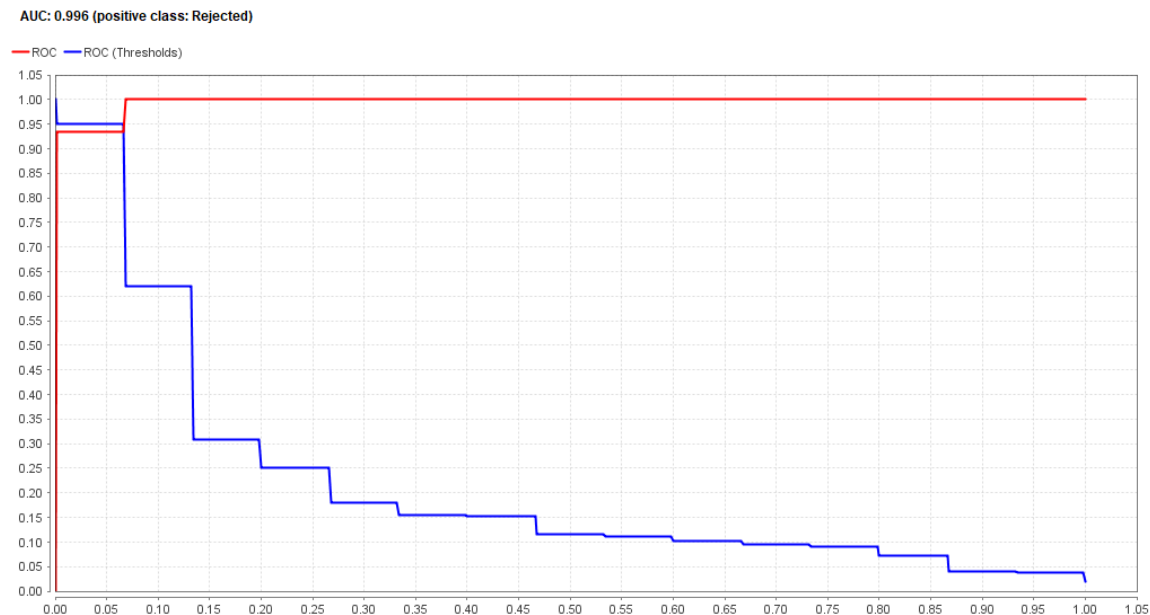
Berdasar tabel diatas prediksi credit approval menggunakan Naïve Bayes Classifier dengan optimasi PSO menghasilkan akurasi sebesar 96.67% dan terdapat rincian jumlah True Positive (TP) 14, False Negative (FN) 15, False Positive (FP) adalah 1 dan True Negative (TN) 0. Data tersebut dapat dihitung manual dengan perhitungan sebagai berikut :

$$\text{Accuracy} = \frac{14+15}{14+15+1+0} = 0.967$$

Berdasarkan perhitungan manual diatas, maka akurasi yang dihasilkan sebesar 96,67%

5) Evaluasi Dengan ROC Curve

Dari gambar yang tertera di bawah ini, ditemukan hasil pengujian terhadap data testing menggunakan metode Naive Bayes Classifier terkait nilai ROC.



Gambar 5. Nilai AUC metode naive bayes dalam Grafik ROC

Nilai AUC sebesar 0,996 untuk evaluasi metode Naive Bayes Classifier berbasis Particle Swarm Optimization (PSO) menunjukkan nilai Fair Classification.

D. Simpulan

Dari hasil akurasi diatas, metode Naïve bayes classifier dengan optimasi Particle Swarm Optimization (PSO) mampu menganalisa prediksi credit approval yang bertujuan agar dapat menganalisa nasabah yang baik dan nasabah yang buruk serta untuk pengembangan berbagai sistem yang mendukung dan memaksimalkan tingkat keberhasilan proses pemberian kredit kepada nasabah. Penelitian ini juga menghasilkan akurasi yang lebih besar dari peneltian sebelumnya yang dilakukan oleh (Irawan, dkk., 2020) menghasilkan akurasi sebesar 79,59%

E. Referensi

- [1] Agustian, A. A., & Bisri, A. (2019). Data Mining Optimization Using Sample Bootstrapping and Particle Swarm Optimization in the Credit Approval Classification. *Indonesian Journal of Artificial Intelligence and Data Mining*, 2(1), 18-27.
- [2] Andika, L. A., Azizah, P. A. N., & Respatiwan, R. (2019). Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier. *Indonesian Journal of Applied Statistics*, 2(1), 34-41.
- [3] Astuti, T., & Astuti, Y. (2022). Analisis Sentimen Review Produk Skincare Dengan Naïve Bayes Classifier Berbasis Particle Swarm Optimization (PSO). *Jurnal Media Informatika Budidarma*, 6(4), 1806-1815.
- [4] Bitetto, A., Cerchiello, P., Filomeni, S., Tanda, A., & Tarantino, B. (2023). Machine learning and credit risk: Empirical evidence from small-and mid-sized businesses. *Socio-Economic Planning Sciences*, 101746.

-
- [5] Fernando, S. D., Purnamasari, S. D., Syaputra, H., & Saputri, N. A. O. (2023). Prediksi Kredit Sepeda Motor pada Showroom Astra Motor Sako Palembang dengan Metode Naive Bayes. *DoubleClick: Journal of Computer and Information Technology*, 7(1), 17-23.
 - [6] Givari, M. R., Sulaeman, M. R., & Umaidah, Y. (2022). Perbandingan Algoritma SVM, Random Forest Dan XGBoost Untuk Penentuan Persetujuan Pengajuan Kredit. *Nuansa Informatika*, 16(1), 141-149.
 - [7] Habibulloh, W., & Topiq, S. (2021). Klasifikasi Kelayakan Kredit Menggunakan Algoritma Naive Bayes Pada Ksp Mekar Jaya Maleber. *Jurnal Responsif: Riset Sains dan Informatika*, 3(1), 92-99.
 - [8] Handayani, N., Wahyono, H., Trianto, J., & Permana, D. S. (2021). Prediksi Tingkat Risiko Kredit dengan Data Mining Menggunakan Algoritma Decision Tree C. 45. *JURIKOM (Jurnal Riset Komputer)*, 8(6), 198-204.
 - [9] Sano, A. V. D., Stefanus, A. A., Madyatmadja, E. D., Nindito, H., Purnomo, A., & Sianipar, C. P. (2023). Proposing a visualized comparative review analysis model on tourism domain using Naïve Bayes classifier. *Procedia Computer Science*, 227, 482-489.
 - [10] Supiyandi, S., Fuad, R. N., Hariyanto, E., & Larasati, S. (2020). Sistem pendukung keputusan pemberian kredit koperasi menggunakan metode weighted product. *Jurnal Media Informatika Budidarma*, 4(4), 1132-1139.
 - [11] Silvilestari, S. (2019). Penerapan Kombinasi Metode Simple Additive Weighting (SAW) dan Rank Order Centroid (ROC) dalam Keputusan Pemberian Kredit. *Jurnal Media Informatika Budidarma*, 3(4), 371-375.
 - [12] Utami, Lilyani Asri. (2017). Analisis Sentimen Opini Publik Berita Kebakaran Hutan Melalui Komparasi Algoritma Support Vector Machine dan K-nearest Neighbor berbasis Particle Swarm Optimization. Vol 13, No 1
 - [13] Wang, S., Ren, J., & Bai, R. (2023). A semi-supervised adaptive discriminative discretization method improving discrimination power of regularized naive Bayes. *Expert Systems with Applications*, 225, 120094.
 - [14] Watratan, A. F., & Moeis, D. (2020). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. *Journal of Applied Computer Science and Technology*, 1(1), 7-14.