https://jurnal.pustakagalerimandiri.co.id/index.php/pustakadata
DOI: https://doi.org/10.55382/jurnalpustakadata.v6i4.2276

Jurnal Pustaka Data

Pusat Akses Kajian Database, Analisa Teknologi, dan Arsitektur Komputer

Vol. 6 No. 4 (2026) 458 – 467

E ISSN: 2809-7483

Prediksi Keputusan Nasabah dalam Berlangganan Deposito Berjangka Menggunakan Algoritma Naive Bayes Classifier

Dewi Eka Putri[1], Eka Praja Wiyata Mandala[2]
[1]Teknik Informatika, Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK Padang, 081372255638
[2]Teknik Informatika, Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK Padang, 085213873216
[1]dewieka@upiyptk.ac.id. [2]ekaprajawm@upiyptk.ac.id

Abstract

The increasing competition in the banking industry requires banks to accurately identify customers who are likely to subscribe to term deposit products. Conventional customer targeting methods often result in ineffective marketing campaigns because not all customers share the same interest in term deposits. The objective of this study is to build a predictive model for customers' decisions to subscribe to time deposit products using the Naive Bayes Classifier algorithm. The dataset consists of 515 banking customer records that have been adapted to represent banking conditions in Padang, Indonesia. The variables used include savings balance, occupation, age, marital status, educational background, housing loan, personal loan, default status, call duration, and previous campaign outcome, with term deposit decision as the target variable. The research process consists of data collection, data selection, preprocessing, data transformation, data splitting, implementation of the Naive Bayes Classifier algorithm, and model evaluation. The dataset was divided into 80% training data and 20% testing data. The experimental results show that the Naive Bayes Classifier achieved an accuracy of 89.32% in predicting customers' decisions to subscribe to term deposits. These findings indicate that the proposed model can effectively support banks in identifying potential customers and improving the effectiveness of term deposit marketing strategies. Furthermore, the Naive Bayes Classifier offers a simple, efficient, and reliable approach for customer classification in the banking sector.

Keywords: naive bayes classifier, term deposit prediction, customer classification, data mining, machine learning

Abstrak

Persaingan industri perbankan yang semakin ketat menuntut bank untuk mampu mengidentifikasi nasabah yang berpotensi berlangganan deposito berjangka secara tepat. Proses penentuan target pemasaran yang masih dilakukan secara konvensional seringkali menyebabkan promosi kurang efektif karena tidak semua nasabah memiliki minat yang sama terhadap produk deposito. Tujuan penelitian ini adalah membangun model prediksi keputusan nasabah terhadap produk deposito berjangka menggunakan algoritma *Naive Bayes Classifier.* Dataset yang digunakan terdiri dari 515 data nasabah perbankan yang telah disesuaikan dengan kondisi perbankan di Kota Padang. Variabel yang digunakan meliputi saldo_tabungan, pekerjaan, usia, status_perkawinan, pendidikan_terakhir, kredit_rumah, kredit_pribadi, tunggakan, durasi_panggilan, dan hasil_kampanye_sebelumnya, dengan keputusan_deposito sebagai variabel target. Tahapan penelitian meliputi pengumpulan data, seleksi data, *preprocessing*, transformasi data, pembagian data, penerapan algoritma *Naive Bayes Classifier*, serta pengujian dan evaluasi model. Dataset dibagi menjadi data training sebesar 80% dan data testing sebesar 20%. Hasil pengujian menunjukkan bahwa model *Naive Bayes Classifier* mampu menghasilkan akurasi sebesar 89,32% dalam memprediksi keputusan nasabah berlangganan deposito berjangka. Hasil penelitian menunjukkan bahwa algoritma *Naive Bayes Classifier* dapat digunakan sebagai metode klasifikasi yang efektif untuk membantu pihak perbankan dalam menentukan target pemasaran deposito secara lebih tepat sasaran.

Submitted : 09-06-2026 | Reviewed : 17-06-2026 | Accepted : 05-08-2026

Kata kunci: naive bayes classifier, prediksi deposito, klasifikasi nasabah, data mining, machine learning

© 2026 Author
Creative Commons Attribution 4.0 International License

1. Pendahuluan

Perkembangan industri perbankan yang semakin kompetitif menuntut setiap bank untuk mampu merancang strategi pemasaran yang efektif dalam meningkatkan jumlah nasabah pada berbagai produk keuangan, termasuk deposito berjangka. Deposito berjangka merupakan salah satu sumber dana penting bagi bank karena dapat mendukung stabilitas likuiditas dan operasional perbankan. Namun demikian, tidak semua nasabah memiliki minat yang sama terhadap produk deposito [1]. Perbedaan karakteristik demografis, kondisi keuangan, serta riwayat interaksi nasabah dengan bank menyebabkan tingkat keberhasilan promosi deposito menjadi bervariasi. Akibatnya, pihak perbankan sering menghadapi kesulitan dalam menentukan nasabah yang berpotensi menerima penawaran deposito sehingga kampanye pemasaran yang dilakukan menjadi kurang efektif dan membutuhkan biaya operasional yang cukup besar. Kondisi tersebut mendorong perlunya penerapan teknologi analisis data yang mampu membantu bank dalam mengidentifikasi pola perilaku nasabah dan memprediksi keputusan mereka terhadap produk deposito berjangka [2].

Penggunaan *machine learning* di sektor perbankan semakin luas sebagai salah satu metode yang efektif dalam mendukung proses pengambilan keputusan berdasarkan analisis data. Dengan memanfaatkan data historis nasabah, model *machine learning* dapat digunakan untuk mengidentifikasi karakteristik pelanggan yang memiliki kecenderungan tinggi untuk menerima penawaran produk keuangan. Pendekatan ini memungkinkan bank untuk melakukan segmentasi dan penargetan promosi yang lebih tepat sehingga sumber daya pemasaran dapat digunakan secara lebih efisien. Selain itu, kemampuan *machine learning* dalam mengolah data dalam jumlah besar dan menemukan pola tersembunyi menjadikannya solusi yang relevan untuk menyelesaikan permasalahan prediksi keputusan nasabah terhadap produk deposito berjangka [3], [4].

Beberapa penelitian sebelumnya telah membuktikan bahwa pendekatan *machine learning* mampu menghasilkan performa yang baik dalam memprediksi minat nasabah terhadap deposito berjangka. Penelitian mengenai prediksi nasabah deposito tidak hanya berfokus pada pemilihan algoritma klasifikasi, tetapi juga pada proses seleksi fitur yang berpengaruh terhadap keberhasilan prediksi. Puteri dkk. (2021) menerapkan metode *Correlation-Based Feature Selection* yang dikombinasikan dengan *Multilayer Perceptron Neural Network* untuk mengidentifikasi atribut yang paling relevan dalam prediksi nasabah potensial deposito. Hasil penelitian menunjukkan bahwa atribut seperti *duration, previous, contact, month, age, job,* dan *housing* merupakan faktor dominan yang memengaruhi keputusan nasabah, dengan tingkat akurasi mencapai 80,5% [5]. Tentua dkk. (2022) melakukan komparasi algoritma *Naïve Bayes* dan *C4.5* dalam memprediksi pelanggan deposito berjangka menggunakan data kampanye pemasaran perbankan. Temuan penelitian mengindikasikan bahwa algoritma *Naïve Bayes* mampu menghasilkan performa klasifikasi yang lebih baik daripada algoritma C4.5 dengan tingkat akurasi sebesar 63,77%. Selain itu, Naïve Bayes dinilai memiliki mekanisme yang lebih sederhana dalam memodelkan hubungan antaratribut yang digunakan. [6].

Penelitian yang dilakukan oleh Mutiarachim dan Tyoso (2024) menerapkan algoritma *Logistic Regression* untuk memprediksi keberhasilan pemasaran deposito menggunakan dataset *Bank Marketing UCI*. Melalui proses seleksi atribut dan penyeimbangan data, model yang dihasilkan memperoleh akurasi sebesar 88,53% dan nilai *AUC* sebesar 93,4%, yang menunjukkan kemampuan model dalam mengidentifikasi nasabah potensial secara efektif [7]. Penelitian yang dilakukan oleh Riyyasy dkk. (2023) menerapkan *Random Forest, Logistic Regression, SVC,* dan *XGBoost* untuk memprediksi preferensi nasabah terhadap term deposit, dimana *Random Forest* dan *XGBoost* menghasilkan akurasi tertinggi sebesar 91,7% [2]. Selanjutnya, Syofyan dkk. (2023) membandingkan fungsi jarak *Manhattan* dan *Minkowski* pada algoritma *K-Nearest Neighbor* untuk memprediksi nasabah potensial deposito berjangka dan memperoleh akurasi terbaik sebesar 88,40% menggunakan fungsi jarak *Minkowski* dengan nilai K=7 [8]. Penelitian lain oleh Panggabean dan Aziz (2023) mengembangkan metode *Ensemble Least Square Support Vector Machine* dengan *AdaBoost* untuk klasifikasi nasabah potensial deposito dan berhasil mencapai akurasi sebesar 95,15% [9].

Pada tahun 2025, penelitian mengenai prediksi deposito berjangka semakin berkembang dengan penggunaan algoritma yang lebih kompleks. Aurelia dan Hiryanto (2025) mengembangkan sistem rekomendasi deposito berjangka menggunakan *Random Forest* dan memperoleh akurasi sebesar 91,56%, dimana variabel duration, age, dan balance merupakan faktor yang paling berpengaruh terhadap keputusan nasabah [4]. Prayoga dkk. (2025)

membandingkan beberapa metode *machine learning* dan *deep learning* seperti *SVM, Logistic Regression, Random Forest, BiGRU,* dan *BiLSTM* untuk memprediksi langganan produk keuangan, dengan hasil terbaik diperoleh oleh *BiGRU* yang mencapai akurasi 92,52% [3]. Saputra dkk. (2025) menggunakan *Regresi Logistik* dan *Support Vector Machine* untuk mengidentifikasi nasabah potensial deposito berjangka serta menekankan pentingnya penanganan ketidakseimbangan data melalui teknik *oversampling* dan *undersampling* [10]. Mangande dan Parmadi (2025) juga mengimplementasikan algoritma Naive Bayes dalam proses klasifikasi calon nasabah yang berpotensi membuka simpanan deposito. Model yang dikembangkan berhasil mencapai akurasi sebesar 86,64%, dengan variabel *duration, cons.conf.idx, nr.employed, emp.var.rate,* dan *euribor3m* sebagai faktor yang memberikan kontribusi terbesar terhadap hasil klasifikasi [11]. Gultom dkk. (2025) menerapkan kombinasi algoritma *Naïve Bayes* dan teknik *Synthetic Minority Over Sampling Technique (SMOTE)* untuk menangani ketidakseimbangan distribusi kelas pada dataset deposito bank. Pendekatan tersebut menghasilkan akurasi sebesar 78,10% dan meningkatkan kemampuan model dalam mengenali calon nasabah potensial melalui proses *data balancing* [12].

Penelitian terkini oleh Christian (2026) membandingkan algoritma *Decision Tree, Random Forest,* dan *XGBoost* dalam memprediksi penerimaan deposito berjangka pada kampanye telemarketing bank. Hasil penelitian menunjukkan bahwa *XGBoost* menghasilkan performa terbaik dengan akurasi di atas 90% setelah dilakukan optimasi menggunakan *Grid Search Cross Validation*. Penelitian ini juga menegaskan pentingnya pemanfaatan machine learning sebagai sistem pendukung keputusan dalam strategi pemasaran perbankan [13].

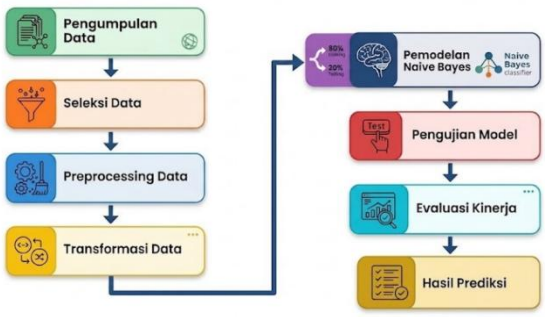
Berdasarkan penelitian-penelitian terdahulu, dapat diketahui bahwa sebagian besar penelitian berfokus pada peningkatan akurasi menggunakan algoritma *Random Forest, XGBoost, Logistic Regression, Support Vector Machine, Neural Network,* maupun *Deep Learning*. Penelitian yang menggunakan *Naïve Bayes* pada kasus deposito berjangka masih relatif terbatas dan umumnya hanya digunakan sebagai metode pembanding atau dikombinasikan dengan teknik lain seperti *SMOTE*. Selain itu, sebagian besar penelitian lebih menitikberatkan pada perbandingan performa berbagai algoritma, sedangkan penelitian yang secara khusus mengevaluasi kemampuan *Naïve Bayes Classifier* sebagai model utama untuk memprediksi keputusan nasabah berlangganan deposito berjangka masih belum banyak dilakukan. Oleh karena itu, penelitian ini berupaya mengisi celah tersebut dengan menerapkan algoritma *Naïve Bayes Classifier* sebagai model utama prediksi keputusan nasabah deposito berjangka serta menganalisis pengaruh atribut-atribut nasabah terhadap hasil prediksi. Selain itu, sebagian besar penelitian lebih menitikberatkan pada peningkatan akurasi model, sedangkan pembahasan mengenai penerapan model yang sederhana, mudah diinterpretasikan, dan efisien untuk membantu pengambilan keputusan pemasaran di lingkungan perbankan masih relatif sedikit. Oleh karena itu, terdapat celah penelitian *(research gap)* berupa kebutuhan untuk mengevaluasi kemampuan algoritma *Naive Bayes Classifier* dalam memprediksi keputusan nasabah berlangganan deposito berjangka menggunakan atribut yang telah disesuaikan dengan karakteristik nasabah perbankan di Indonesia, khususnya pada konteks Kota Padang.

Berdasarkan permasalahan dan *research gap* tersebut, penelitian ini bertujuan untuk membangun model prediksi keputusan nasabah dalam berlangganan deposito berjangka menggunakan algoritma *Naive Bayes Classifier*. Data yang digunakan terdiri dari atribut demografis, finansial, dan riwayat interaksi nasabah dengan bank, yaitu saldo tabungan, pekerjaan, usia, status perkawinan, pendidikan terakhir, kredit rumah, kredit pribadi, tunggakan, durasi panggilan, dan hasil kampanye sebelumnya. Penelitian ini diharapkan mampu menghasilkan model klasifikasi yang memiliki performa baik dalam memprediksi keputusan nasabah, sekaligus memberikan informasi mengenai faktor-faktor yang paling berpengaruh terhadap keputusan berlangganan deposito. Penelitian ini diharapkan mampu memberikan dukungan dalam proses pengambilan keputusan di sektor perbankan, khususnya dalam penyusunan strategi pemasaran yang lebih efektif, peningkatan kualitas promosi, dan optimalisasi peluang keberhasilan pemasaran produk deposito berjangka.

2. Metode Penelitian

2.1. Tahapan Penelitian

Penelitian ini bertujuan untuk membangun model prediksi keputusan nasabah dalam berlangganan deposito berjangka menggunakan algoritma *Naive Bayes Classifier*. Seluruh tahapan penelitian dirancang secara sistematis, mulai dari pengumpulan data, pengolahan data, hingga evaluasi model untuk menghasilkan prediksi yang optimal. Alur pelaksanaan penelitian ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Berdasarkan Gambar 1, alur penelitian terdiri atas beberapa tahapan, yaitu pengumpulan data, seleksi data, *preprocessing*, dan transformasi data. Setelah seluruh proses tersebut selesai dilakukan, dataset dibagi menjadi 80% data training dan 20% data testing sebagai dasar pembentukan model *Naive Bayes Classifier*. Selanjutnya, model diuji dan dievaluasi untuk mengukur kinerjanya, sehingga menghasilkan prediksi keputusan nasabah dalam berlangganan deposito berjangka.

2.2. Pengumpulan Data

Tahap pengumpulan data menjadi langkah awal yang krusial dalam penelitian, mengingat kualitas data yang digunakan sangat berpengaruh terhadap performa proses klasifikasi. Pada penelitian ini, data yang digunakan berupa dataset nasabah perbankan yang telah disesuaikan dengan karakteristik dan kondisi perbankan di Kota Padang. Dataset tersebut terdiri dari 515 data nasabah yang merepresentasikan berbagai karakteristik demografis, kondisi finansial, serta riwayat interaksi nasabah dengan pihak bank dalam kegiatan pemasaran produk deposito berjangka. Data yang digunakan merupakan hasil adaptasi dari dataset *Bank Marketing* yang kemudian disesuaikan ke dalam konteks perbankan Indonesia, khususnya Kota Padang, melalui penerjemahan atribut ke Bahasa Indonesia dan penyesuaian kategori variabel agar lebih relevan dengan kondisi nasabah lokal. Setiap data nasabah memiliki informasi yang dapat digunakan untuk menggambarkan profil dan perilaku nasabah dalam mengambil keputusan terhadap produk deposito berjangka.

Variabel yang digunakan dalam penelitian ini terdiri atas sepuluh variabel independen yang berperan sebagai prediktor serta satu variabel dependen yang menjadi target klasifikasi. Variabel prediktor mencakup faktor finansial, demografis, status kredit, serta riwayat komunikasi pemasaran yang memengaruhi keputusan nasabah untuk berlangganan deposito berjangka. Sementara itu, variabel target berupa keputusan_deposito, yang menunjukkan apakah nasabah memutuskan untuk berlangganan deposito berjangka atau tidak. Variabel dalam penelitian ini dapat dilihat pada Tabel 1.

Tabel 1. Variabel Penelitian

No	Variabel	Jenis Data
1	saldo_tabungan	Numerik
2	pekerjaan	Kategorikal
3	usia	Numerik
4	status_perkawinan	Kategorikal
5	pendidikan_terakhir	Kategorikal
6	kredit_rumah	Kategorikal
7	kredit_pribadi	Kategorikal
8	tunggakan	Kategorikal
9	durasi_panggilan	Numerik
10	hasil_kampanye_sebelumnya	Kategorikal
11	keputusan_deposito	Kategorikal (Target)

Berdasarkan variabel yang digunakan, penelitian ini memanfaatkan kombinasi atribut finansial, demografis, dan historis pemasaran untuk membangun model klasifikasi menggunakan algoritma *Naive Bayes Classifier*. Variabel-variabel tersebut dipilih berdasarkan hasil studi literatur dan penelitian terdahulu yang menunjukkan bahwa karakteristik nasabah, kondisi keuangan, serta efektivitas kampanye pemasaran memiliki pengaruh terhadap keputusan nasabah dalam memilih produk deposito berjangka. Dengan memanfaatkan variabel-variabel yang telah ditentukan, model yang dibangun diharapkan mampu memberikan hasil prediksi yang akurat serta mendukung pengambilan keputusan pihak perbankan dalam menentukan calon nasabah potensial untuk penawaran deposito berjangka.

2.3. Seleksi Data

Seleksi data dilakukan untuk menentukan atribut yang relevan dan memiliki pengaruh terhadap prediksi keputusan nasabah dalam berlangganan deposito berjangka. Proses ini bertujuan untuk mengurangi atribut yang tidak diperlukan sehingga model klasifikasi dapat bekerja lebih efektif dan menghasilkan performa yang optimal. Pemilihan atribut dilakukan berdasarkan studi literatur dan penelitian terdahulu yang menunjukkan bahwa faktor demografis, finansial, status kredit, serta riwayat kampanye pemasaran berpengaruh terhadap keputusan nasabah dalam memilih produk deposito. Berdasarkan proses seleksi data, diperoleh 10 atribut yang digunakan sebagai variabel input, yaitu saldo_tabungan, pekerjaan, usia, status_perkawinan, pendidikan_terakhir, kredit_rumah, kredit_pribadi, tunggakan, durasi_panggilan, dan hasil_kampanye_sebelumnya, sedangkan keputusan_deposito sebagai variabel target.

2.4. Preprocessing Data

Tahap *preprocessing* data dilakukan sebagai proses persiapan sebelum penerapan algoritma *Naive Bayes Classifier.* Proses ini bertujuan untuk memperbaiki kualitas data sehingga model klasifikasi yang dibangun mampu memberikan hasil yang lebih akurat dan optimal.Pada penelitian ini, *preprocessing* dilakukan melalui beberapa tahapan, yaitu *data cleaning*, *data encoding*, dan normalisasi data.

Data cleaning dilakukan untuk memeriksa konsistensi data serta memastikan tidak terdapat data duplikat maupun nilai yang tidak sesuai. Selanjutnya, *data encoding* dilakukan untuk mengubah atribut kategorikal seperti pekerjaan, status_perkawinan, pendidikan_terakhir, kredit_rumah, kredit_pribadi, tunggakan, hasil_kampanye_sebelumnya, dan keputusan_deposito ke dalam bentuk numerik agar dapat diproses oleh algoritma *Naive Bayes*. Setelah itu, dilakukan normalisasi data untuk atribut numerik seperti usia, saldo tabungan, dan durasi panggilan untuk menyamakan rentang nilai antar atribut. Hasil dari tahap *preprocessing* adalah dataset yang telah

bersih, terstruktur, dan siap digunakan pada proses pembentukan model *Naive Bayes Classifier*.

2.5. Transformasi Data

Transformasi data dilakukan untuk mengubah data hasil *preprocessing* ke dalam format yang sesuai untuk proses pembentukan model klasifikasi. Pada tahap ini, seluruh atribut yang telah melalui proses *encoding* disusun menjadi variabel input *(features)* dan variabel target *(label).* Variabel input terdiri dari saldo_tabungan, pekerjaan, usia, status_perkawinan, pendidikan_terakhir, kredit_rumah, kredit_pribadi, tunggakan, durasi_panggilan, dan hasil_kampanye_sebelumnya, sedangkan variabel target adalah keputusan_deposito. Selain itu, data numerik dan kategorikal yang telah ditransformasikan ke dalam bentuk numerik selanjutnya diperiksa kembali untuk memastikan tidak terdapat kesalahan format yang dapat memengaruhi proses klasifikasi. Hasil dari proses transformasi data berupa dataset yang telah tersusun secara sistematis sehingga siap digunakan untuk pembagian data training dan testing, serta sebagai masukan dalam pembentukan model *Naive Bayes Classifier.*

2.6. Pembagian Data

Setelah proses transformasi data selesai dilakukan, tahap selanjutnya adalah pembagian data *(data splitting)* menjadi data latih *(training data)* dan data uji *(testing data).* Pembagian data bertujuan untuk mengukur kemampuan model dalam mempelajari pola dari data latih dan mengevaluasi kinerjanya terhadap data yang belum pernah digunakan sebelumnya. Dengan demikian, hasil pengujian yang diperoleh dapat menggambarkan kemampuan model dalam melakukan prediksi pada data baru. Penelitian ini menerapkan metode *holdout validation* untuk membagi dataset yang berjumlah 515 data nasabah menjadi dua bagian, yaitu 80% *data training* dan 20% *data testing*. Sebanyak 412 data digunakan pada tahap pelatihan model *Naive Bayes Classifier*, sementara 103 data digunakan pada tahap pengujian guna mengukur performa model. Pembagian data dilakukan secara acak *(random splitting)* untuk mengurangi potensi bias pada proses pelatihan dan pengujian model.

2.7. Algoritma Naive Bayes Classifier

Algoritma *Naive Bayes Classifier* merupakan metode klasifikasi yang dibangun berdasarkan T*eorema Bayes* dan konsep probabilitas. Mekanisme kerjanya dilakukan dengan menghitung peluang suatu data berada pada kelas tertentu berdasarkan nilai atribut yang dimiliki. Keunggulan utama *Naive Bayes* adalah proses komputasinya yang sederhana, cepat, serta mampu menghasilkan performa yang baik pada dataset dengan jumlah data yang relatif besar dan atribut yang beragam. Dalam penelitian ini, algoritma *Naive Bayes* digunakan untuk memprediksi keputusan nasabah dalam berlangganan deposito berjangka berdasarkan atribut saldo_tabungan, pekerjaan, usia, status_perkawinan, pendidikan_terakhir, kredit_rumah, kredit_pribadi, tunggakan, durasi_panggilan, dan hasil_kampanye_sebelumnya. Prediksi dilakukan dengan menghitung probabilitas setiap kelas target, yaitu Berlangganan dan Tidak Berlangganan, kelas dengan probabilitas posterior terbesar kemudian ditentukan sebagai hasil prediksi model.

Dasar perhitungan *Naive Bayes* menggunakan *Teorema Bayes* dinyatakan pada Persamaan (1) [14].

$$P(Y|X) = \frac{P(X|Y) \cdot P(Y)}{P(X)} \qquad (1)$$

Dimana $P(Y|X)$ adalah probabilitas suatu kelas Y berdasarkan atribut X. $P(X|Y)$ adalah probabilitas kemunculan atribut X pada kelas Y. $P(Y)$ adalah probabilitas awal *(prior probability)* dari kelas Y. $P(X)$ adalah probabilitas kemunculan atribut X. X merupakan kumpulan atribut yang digunakan dalam proses klasifikasi. Y merupakan kelas target yang akan diprediksi.

Karena penelitian ini menggunakan lebih dari satu atribut prediktor, maka probabilitas setiap atribut diasumsikan saling independen *(conditional independence).* Oleh karena itu, probabilitas posterior dapat dihitung menggunakan persamaan (2).

$$P(Y|X) = P(Y) \prod_{i=1}^{n} P(x_i|Y) \qquad (2)$$

Dimana $P(Y|X)$ adalah probabilitas kelas berdasarkan seluruh atribut. $P(Y)$ adalah probabilitas awal kelas. $P(x_i|Y)$ adalah probabilitas atribut ke-i terhadap kelas Y. n adalah jumlah atribut yang digunakan dalam penelitian. $\prod$ menunjukkan operasi perkalian seluruh probabilitas atribut.

Berdasarkan persamaan tersebut, model akan menghitung probabilitas untuk masing-masing kelas target, yaitu Berlangganan dan Tidak Berlangganan. Kelas yang memiliki nilai probabilitas terbesar akan ditetapkan sebagai hasil prediksi keputusan deposito nasabah. Dengan pendekatan ini, *Naive Bayes Classifier* diharapkan mampu mengidentifikasi pola hubungan antara karakteristik nasabah dan keputusan berlangganan deposito secara efektif serta menghasilkan tingkat akurasi yang baik dalam proses klasifikasi.

2.8. Pengujian Model

Pengujian model dilakukan untuk mengetahui kemampuan algoritma *Naive Bayes Classifier* dalam memprediksi keputusan nasabah berlangganan deposito berjangka berdasarkan data yang telah diproses sebelumnya. Pada tahap ini, model yang telah dibangun menggunakan data training diuji

menggunakan data testing yang belum pernah digunakan pada proses pelatihan.

Proses pengujian dilakukan dengan membandingkan hasil prediksi yang dihasilkan model dengan data aktual pada dataset testing. Hasil pengujian kemudian disajikan dalam bentuk *Confusion Matrix* yang digunakan untuk mengetahui jumlah prediksi yang benar maupun prediksi yang salah pada masing-masing kelas. Confusion Matrix terdiri dari empat komponen utama, yaitu *True Positive (TP), True Negative (TN), False Positive (FP),* dan *False Negative (FN)* [15].

2.9. Evaluasi Kinerja

Berdasarkan nilai yang diperoleh dari *Confusion Matrix,* selanjutnya dilakukan pengukuran performa model menggunakan beberapa metrik evaluasi, yaitu *Accuracy, Precision, Recall,* dan *F1-Score* [16].
Nilai *Accuracy* digunakan untuk mengukur tingkat ketepatan prediksi model terhadap seluruh data pengujian dan dihitung menggunakan Persamaan (3).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

Nilai *Precision* digunakan untuk mengukur tingkat ketepatan prediksi pada kelas positif dan dihitung menggunakan Persamaan (4).

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

Nilai *Recall* digunakan untuk mengukur kemampuan model dalam menemukan seluruh data yang termasuk ke dalam kelas positif dan dihitung menggunakan Persamaan (5).

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

Sedangkan nilai *F1-Score* merupakan rata-rata harmonik antara *Precision* dan *Recall* yang dihitung menggunakan Persamaan (6).

$$F1 - Score = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall} \quad (6)$$

Hasil pengujian model yang diperoleh dari perhitungan *Accuracy, Precision, Recall,* dan *F1-Score* digunakan untuk mengevaluasi tingkat keberhasilan algoritma *Naive Bayes Classifier* dalam memprediksi keputusan nasabah berlangganan deposito berjangka. Semakin tinggi nilai yang dihasilkan pada setiap metrik evaluasi, maka semakin baik kemampuan model dalam melakukan klasifikasi terhadap data nasabah.

2.10. Hasil Prediksi

Tahap terakhir menghasilkan model klasifikasi yang mampu memprediksi keputusan nasabah dalam berlangganan deposito berjangka berdasarkan atribut yang dimiliki. Hasil prediksi diharapkan dapat membantu pihak perbankan dalam menentukan calon nasabah potensial sehingga strategi pemasaran deposito dapat dilakukan secara lebih efektif dan efisien.

3. Hasil dan Pembahasan

Bagian ini menyajikan hasil penelitian yang diperoleh dari penerapan algoritma *Naive Bayes Classifier* dalam memprediksi keputusan nasabah berlangganan deposito berjangka. Pembahasan diawali dengan penyajian karakteristik dataset yang digunakan, dilanjutkan dengan hasil preprocessing data, pembentukan model, pengujian model, serta evaluasi kinerja klasifikasi. Hasil yang diperoleh kemudian dianalisis untuk mengetahui kemampuan model dalam mengidentifikasi nasabah yang berpotensi berlangganan deposito berdasarkan atribut yang digunakan dalam penelitian.

3.1. Hasil Pengumpulan Data

Tahap pengumpulan data menghasilkan dataset yang terdiri dari 515 data nasabah perbankan yang telah disesuaikan dengan kondisi perbankan di Kota Padang. Dataset tersebut memuat informasi karakteristik nasabah yang meliputi aspek demografis, finansial, status kredit, serta riwayat interaksi pemasaran yang digunakan sebagai dasar dalam proses prediksi keputusan deposito. Variabel yang digunakan terdiri dari 10 variabel input dan 1 variabel target, yaitu keputusan_deposito. Sebagai gambaran terhadap data yang digunakan dalam penelitian, Tabel 2 menyajikan contoh dataset nasabah yang menjadi objek penelitian.

Tabel 2. Dataset Nasabah

Nasabah	Saldo Tabungan	Pekerjaan	Usia	Status Perkawinan	Pendidikan Terakhir	Kredit Rumah	Kredit Pribadi	Tunggakan	Durasi Panggilan	Hasil Kampanye Sebelumnya	Keputusan Deposito
1	13,282,500	pekerja lapangan	32	menikah	SMA/SMK	ada	tidak ada	tidak ada	317	belum dihubungi	berlangganan
2	26,792,500	pensiunan	70	cerai	SD/SMP	tidak ada	tidak ada	tidak ada	445	belum dihubungi	tidak berlangganan
3	32,200,000	pensiunan	65	cerai	SD/SMP	tidak ada	tidak ada	tidak ada	383	gagal	tidak berlangganan
4	6,037,500	pekerja lapangan	56	menikah	SMA/SMK	ada	tidak ada	tidak ada	605	belum dihubungi	tidak berlangganan

Submitted : 09-06-2026 | Reviewed : 17-06-2026 | Accepted : 05-08-2026

.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
515	60,655,000	manajer	31	menikah	SMA/SMK	ada	tidak ada	tidak ada	164	belum dihubungi	tidak berlangganan

Selanjutnya, ringkasan dataset disajikan pada Tabel 3 untuk menunjukkan jumlah data, jumlah atribut, dan kelas target yang digunakan dalam proses klasifikasi.

Tabel 3. Ringkasan Dataset

Keterangan	Nilai
Jumlah Data	515
Jumlah Variabel Input	10
Jumlah Variabel Target	1
Total Variabel	11
Kelas Target	Berlangganan, Tidak Berlangganan

Berdasarkan ringkasan dataset pada Tabel 3, data yang digunakan telah memenuhi kebutuhan penelitian karena memuat atribut yang relevan untuk menggambarkan karakteristik nasabah dan faktor-faktor yang memengaruhi keputusan berlangganan deposito berjangka. Dataset tersebut selanjutnya digunakan pada tahap *preprocessing* dan pembentukan model klasifikasi menggunakan algoritma *Naive Bayes Classifier*.

3.2. Hasil Seleksi Data

Berdasarkan hasil seleksi data, diperoleh 10 atribut yang digunakan sebagai variabel input, yaitu saldo_tabungan, pekerjaan, usia, status_perkawinan, pendidikan_terakhir, kredit_rumah, kredit_pribadi, tunggakan, durasi_panggilan, dan hasil_kampanye_sebelumnya. Sementara itu, atribut keputusan_deposito digunakan sebagai variabel target yang akan diprediksi oleh model. Pemilihan atribut tersebut didasarkan pada karakteristik nasabah, kondisi finansial, status kredit, serta riwayat kampanye pemasaran yang berpotensi memengaruhi keputusan nasabah dalam memilih produk deposito berjangka. Data yang telah melalui proses seleksi dalam penelitian ini ditampilkan pada Tabel 4.

Tabel 4. Hasil Seleksi Data

No	Variabel	Peran
1	saldo_tabungan	Input
2	pekerjaan	Input
3	usia	Input
4	status_perkawinan	Input
5	pendidikan_terakhir	Input
6	kredit_rumah	Input
7	kredit_pribadi	Input
8	tunggakan	Input
9	durasi_panggilan	Input
10	hasil_kampanye_sebelumnya	Input
11	keputusan_deposito	Target

Berdasarkan hasil seleksi data pada Tabel 4, seluruh atribut input yang dipilih selanjutnya digunakan dalam proses preprocessing dan pembentukan model *Naive Bayes Classifier*. Atribut-atribut tersebut diharapkan mampu merepresentasikan kondisi nasabah secara menyeluruh sehingga model dapat mengidentifikasi pola yang berhubungan dengan keputusan berlangganan deposito berjangka secara lebih efektif.

3.3. Hasil Preprocessing dan Transformasi Data

Tahap *preprocessing* bertujuan untuk meningkatkan kualitas data, menghilangkan inkonsistensi, serta mengubah atribut kategorikal menjadi format numerik yang dapat dikenali oleh sistem. Dataset yang telah melalui tahap seleksi data kemudian diperiksa untuk memastikan tidak terdapat data duplikat, data kosong *(missing value),* maupun nilai yang tidak sesuai dengan ketentuan penelitian. Selanjutnya, dilakukan proses transformasi data terhadap atribut yang masih berbentuk kategorikal, seperti pekerjaan, status_perkawinan, pendidikan_terakhir, kredit_rumah, kredit_pribadi, tunggakan, hasil_kampanye_sebelumnya, dan keputusan_deposito. Proses transformasi dilakukan menggunakan teknik label *encoding*, yaitu mengubah setiap kategori menjadi representasi numerik tanpa menghilangkan makna informasi yang terkandung di dalamnya. Sementara itu, atribut numerik seperti saldo_tabungan, usia, dan durasi_panggilan tetap digunakan sesuai nilai aslinya.

3.4. Hasil Pembagian Data

Pada penelitian ini, pembagian data dilakukan menggunakan metode holdout validation dengan perbandingan 80:20, di mana 80% data digunakan sebagai data training dan 20% sisanya digunakan sebagai data testing. Dataset yang terdiri atas 515 data nasabah dibagi menjadi 412 *data training* dan 103 *data testing*. Proses pembagian dilakukan menggunakan metode random sampling agar setiap data memiliki kesempatan yang sama untuk terpilih sebagai data pelatihan maupun data pengujian. Rincian hasil pembagian dataset dapat dilihat pada Tabel 5.

Tabel 5. Hasil Pembagian Dataset

Dataset	Jumlah Data	Persentase
Data Training	412	80%
Data Testing	103	20%
Total Data	515	100%

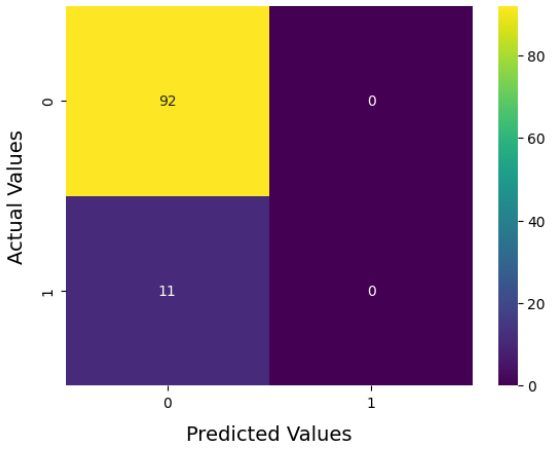
Berdasarkan hasil pembagian data pada Tabel 5, data training digunakan untuk membentuk model klasifikasi *Naive Bayes* dengan mempelajari pola hubungan antara atribut input dan variabel target. *Data testing* kemudian dimanfaatkan untuk menguji kemampuan model dalam memprediksi keputusan deposito nasabah berdasarkan data yang belum pernah dikenali sebelumnya. Hasil dari proses pengujian tersebut akan digunakan sebagai dasar

dalam mengevaluasi tingkat kinerja model pada tahap berikutnya.

3.5. Hasil Pengujian Model Naive Bayes Classifier

Model *Naive Bayes* dibangun menggunakan data training sebanyak 412 data. Algoritma menghitung probabilitas masing-masing kelas berdasarkan atribut yang dimiliki oleh nasabah dan menghasilkan model klasifikasi yang digunakan untuk memprediksi keputusan deposito. Pada tahap ini, algoritma *Naive Bayes* menghitung probabilitas masing-masing kelas target berdasarkan pola yang terdapat pada *data training*. Setiap atribut yang dimiliki nasabah akan digunakan untuk menentukan peluang suatu data termasuk ke dalam kelas Berlangganan atau Tidak Berlangganan. Algoritma kemudian membandingkan nilai probabilitas dari kedua kelas tersebut dan menetapkan kelas dengan probabilitas tertinggi sebagai hasil prediksi.

Untuk memperoleh gambaran yang lebih rinci mengenai kemampuan model dalam melakukan klasifikasi, hasil pengujian tidak hanya dievaluasi menggunakan nilai akurasi, tetapi juga dianalisis menggunakan *Confusion Matrix*. Evaluasi model dilakukan menggunakan *Confusion Matrix,* yang berfungsi untuk menunjukkan jumlah prediksi yang benar maupun prediksi yang salah. Matriks ini menyajikan distribusi nilai *True Positive (TP), True Negative (TN), False Positive (FP),* dan *False Negative (FN)* yang dihasilkan oleh model *Naive Bayes Classifier.* Hasilnya ditampilkan pada Gambar 2.



Gambar 2. Confusion Matrix Klasifikasi Naive Bayes Classifier

Berdasarkan *Confusion Matrix* pada Gambar 2, diperoleh bahwa model berhasil mengklasifikasikan 92 data pada kelas 0 dengan benar, sedangkan 11 data pada kelas 1 diklasifikasikan ke dalam kelas 0. Selain itu, tidak terdapat data yang diprediksi sebagai kelas 1, sehingga nilai pada kolom prediksi kelas 1 bernilai 0. Berdasarkan hasil yang diperoleh, model mampu mengenali data pada kelas mayoritas (kelas 0) dengan sangat baik. Namun, performa model dalam mengidentifikasi data pada kelas minoritas (kelas 1) masih belum optimal. Kondisi ini terlihat dari tidak adanya data kelas 1 yang berhasil diprediksi dengan benar *(True Positive)*. Fenomena ini umumnya dipengaruhi oleh distribusi data yang tidak seimbang *(imbalanced dataset),* sehingga jumlah data pada kelas 0 mendominasi dibandingkan data pada kelas 1.

Setelah proses pelatihan model selesai dilakukan, algoritma *Naive Bayes Classifier* diterapkan pada data testing untuk mengetahui kemampuannya dalam memprediksi keputusan nasabah berlangganan deposito berjangka. Hasil pengujian model menghasilkan nilai evaluasi berupa *accuracy, precision, recall,* dan *classification report* yang digunakan untuk mengukur tingkat kinerja model dalam melakukan klasifikasi. Hasil pengujian *Naive Bayes Classifier* dapat dilihat pada Gambar 3.

Classification Report:

	precision	recall	f1-score	support
1	0.89	1.00	0.94	92
2	0.00	0.00	0.00	11
accuracy			0.89	103
macro avg	0.45	0.50	0.47	103
weighted avg	0.80	0.89	0.84	103

Gambar 3. Hasil Pengujian Model Naive Bayes Classifier

3.6. Hasil Evaluasi

Hasil pengujian pada Gambar 2 menunjukkan bahwa model *Naive Bayes Classifier* mencapai nilai *accuracy* sebesar 89,32%. Nilai tersebut menunjukkan bahwa sebanyak 89,32% *data testing* berhasil diklasifikasikan dengan benar. Sementara itu, nilai *precision* sebesar 89,32% menunjukkan tingkat ketepatan prediksi model, sedangkan nilai *recall* sebesar 89,32% menggambarkan kemampuan model dalam mendeteksi data pada kelas target.

Hasil *classification report* menunjukkan bahwa pada kelas 1 model memperoleh nilai *precision* sebesar 0,89, *recall* sebesar 1,00, dan *f1-score* sebesar 0,94 dengan jumlah data *(support)* sebanyak 92 data. Nilai tersebut menunjukkan bahwa model sangat baik dalam mengenali data yang termasuk ke dalam kelas 1. Namun, pada kelas 2, model memperoleh nilai *precision, recall,* dan *f1-score* sebesar 0,00 dengan jumlah data sebanyak 11 data. Hal ini mengindikasikan bahwa model belum mampu mengidentifikasi data pada kelas 2 secara optimal.

Perbedaan nilai *macro average F1-score* (0,47) dan *weighted average F1-score* (0,84) menunjukkan bahwa performa model tidak merata pada masing-masing kelas yang diklasifikasikan. Kondisi ini mengindikasikan bahwa distribusi data pada variabel target cenderung tidak seimbang *(imbalanced dataset),* sehingga model lebih dominan memprediksi kelas yang memiliki jumlah data lebih banyak. Meskipun demikian, secara keseluruhan model *Naive Bayes Classifier* mampu menghasilkan tingkat akurasi yang cukup baik dalam memprediksi

keputusan nasabah berlangganan deposito berjangka dengan nilai akurasi mencapai 89,32%.

3.7. Pembahasan Hasil

Berdasarkan hasil pengujian yang telah dilakukan, algoritma *Naive Bayes Classifier* mampu menghasilkan nilai akurasi sebesar 89,32%, yang menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam memprediksi keputusan nasabah berlangganan deposito berjangka. Nilai akurasi tersebut mengindikasikan bahwa sebagian besar data testing berhasil diklasifikasikan dengan benar sesuai dengan kelas aktualnya. Hasil ini menunjukkan bahwa atribut yang digunakan dalam penelitian, yaitu saldo_tabungan, pekerjaan, usia, status_perkawinan, pendidikan_terakhir, kredit_rumah, kredit_pribadi, tunggakan, durasi_panggilan, dan hasil_kampanye_sebelumnya, memiliki hubungan terhadap keputusan nasabah dalam memilih produk deposito berjangka.

Hasil *classification report* menunjukkan bahwa model memiliki performa yang sangat baik pada kelas 1, yang ditunjukkan oleh nilai *precision* sebesar 0,89, *recall* sebesar 1,00, dan *F1-score* sebesar 0,94. Capaian *recall* sebesar 100% menunjukkan bahwa tidak terdapat data pada kelas tersebut yang gagal dikenali oleh model. Hasil ini mengindikasikan bahwa algoritma *Naive Bayes* mampu mempelajari pola karakteristik nasabah yang termasuk dalam kelas mayoritas dengan baik sehingga proses klasifikasi dapat dilakukan secara optimal.

Meskipun demikian, performa model pada kelas 2 masih tergolong rendah karena nilai *precision, recall,* dan *F1-score* yang diperoleh seluruhnya sebesar 0,00. Hal ini menunjukkan bahwa model belum mampu mengklasifikasikan data pada kelas tersebut dengan benar. Salah satu penyebab utama kondisi ini adalah adanya ketidakseimbangan jumlah data *(class imbalance)* pada variabel target. Berdasarkan hasil pengujian, jumlah data pada kelas 1 mencapai 92 data, sedangkan kelas 2 hanya terdiri dari 11 data. Perbedaan distribusi yang cukup besar menyebabkan model lebih cenderung mempelajari pola pada kelas mayoritas dibandingkan kelas minoritas.

Walaupun performa model pada salah satu kelas masih belum optimal, nilai *weighted average F1-score* sebesar 0,84 mengindikasikan bahwa model secara keseluruhan masih mampu melakukan klasifikasi dengan baik terhadap dataset penelitian. Hasil ini membuktikan bahwa algoritma *Naive Bayes Classifier* mampu digunakan sebagai metode prediksi keputusan nasabah berlangganan deposito berjangka karena memiliki proses perhitungan yang sederhana, cepat, dan menghasilkan tingkat akurasi yang cukup tinggi.

Variabel yang digunakan dalam penelitian terdiri dari faktor finansial, demografis, dan riwayat pemasaran. Di antara variabel tersebut, saldo_tabungan, durasi_panggilan, dan hasil_kampanye_sebelumnya memiliki pengaruh yang lebih besar terhadap keputusan nasabah dalam berlangganan deposito berjangka. Selain itu, variabel usia, pekerjaan, dan pendidikan_terakhir turut memberikan kontribusi dalam proses klasifikasi karena berkaitan dengan kondisi ekonomi dan kemampuan finansial nasabah. Kombinasi atribut tersebut memungkinkan model mengidentifikasi pola yang berhubungan dengan keputusan deposito secara lebih efektif.

3.8. Perbandingan dengan Penelitian Sebelumnya

Sebagai bentuk evaluasi terhadap kinerja model, hasil penelitian ini dibandingkan dengan sejumlah penelitian terdahulu yang membahas prediksi nasabah deposito berjangka menggunakan berbagai metode klasifikasi. Perbandingan dilakukan berdasarkan nilai akurasi yang diperoleh dari masing-masing penelitian. Hasil perbandingan dapat dilihat pada Tabel 6.

Tabel 6. Perbandingan Hasil Penelitian

Peneliti	Metode	Akurasi
Tentua dkk. (2022)	Naive Bayes	63,77%
Syofyan dkk. (2023)	K-Nearest Neighbor	88,40%
Mutiarachim dan Tyoso (2024)	Logistic Regression	88,53%
Gultom dkk. (2025)	Naive Bayes + SMOTE	78,10%
Penelitian Ini	**Naive Bayes Classifier**	**89,32%**

Berdasarkan Tabel 6, algoritma *Naive Bayes Classifier* yang digunakan dalam penelitian ini menghasilkan akurasi sebesar 89,32%. Nilai tersebut lebih tinggi dibandingkan penelitian Tentua dkk. (2022) yang memperoleh akurasi 63,77% menggunakan metode *Naive Bayes* serta penelitian Gultom dkk. (2025) yang memperoleh akurasi 78,10% menggunakan *Naive Bayes* yang dikombinasikan dengan *SMOTE*. Hasil penelitian ini juga sedikit lebih tinggi dibandingkan *Logistic Regression* yang memperoleh akurasi 88,53% pada penelitian Mutiarachim dan Tyoso (2024).

Berdasarkan hasil perbandingan dengan penelitian terdahulu, *Naive Bayes Classifier* masih menunjukkan kemampuan yang kompetitif dalam melakukan prediksi terhadap keputusan nasabah untuk berlangganan deposito berjangka. Dengan akurasi sebesar 89,32%, metode ini dapat menjadi alternatif yang efektif untuk diterapkan pada sistem pendukung keputusan di lingkungan perbankan, terutama ketika dibutuhkan model yang sederhana, mudah diimplementasikan, dan memiliki performa klasifikasi yang baik.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma *Naive Bayes Classifier* mampu diterapkan untuk memprediksi keputusan nasabah dalam berlangganan deposito berjangka berdasarkan karakteristik demografis,

finansial, dan riwayat pemasaran nasabah. Penelitian menggunakan dataset sebanyak 515 data nasabah perbankan yang telah disesuaikan dengan kondisi perbankan di Kota Padang, dengan variabel saldo_tabungan, pekerjaan, usia, status_perkawinan, pendidikan_terakhir, kredit_rumah, kredit_pribadi, tunggakan, durasi_panggilan, dan hasil_kampanye_sebelumnya sebagai atribut prediktor.

Berdasarkan hasil pengujian, algoritma *Naive Bayes Classifier* berhasil mencapai tingkat accuracy sebesar 89,32%. Nilai tersebut menunjukkan bahwa model mampu melakukan klasifikasi keputusan deposito nasabah dengan tingkat ketepatan yang baik. Selain itu, variabel yang berkaitan dengan kondisi finansial, aktivitas pemasaran, dan karakteristik nasabah terbukti berkontribusi dalam proses prediksi keputusan berlangganan deposito berjangka.

Walaupun model menunjukkan kinerja yang baik secara keseluruhan, hasil *Confusion Matrix* dan *Classification Report* mengindikasikan bahwa model masih mengalami penurunan performa pada kelas minoritas karena distribusi data yang tidak seimbang. Oleh sebab itu, penelitian berikutnya disarankan menerapkan teknik data balancing, seperti *SMOTE*, atau mengeksplorasi algoritma klasifikasi lain untuk meningkatkan kemampuan model dalam mengenali seluruh kelas.

ecara umum, hasil penelitian membuktikan bahwa algoritma *Naive Bayes Classifier* mampu memberikan kinerja yang baik sebagai metode klasifikasi dalam mengidentifikasi calon nasabah yang berpotensi berlangganan deposito berjangka. Dengan karakteristiknya yang sederhana dan efisien, algoritma ini dapat dimanfaatkan sebagai pendukung penyusunan strategi pemasaran yang lebih efektif.

Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada Universitas Putra Indonesia YPTK Padang, khususnya Fakultas Ilmu Komputer, atas dukungan dan fasilitas yang diberikan selama pelaksanaan penelitian ini. Penulis juga menyampaikan terima kasih kepada semua pihak yang telah memberikan bantuan, masukan, dan motivasi sehingga penelitian ini dapat diselesaikan dengan baik. Semoga hasil penelitian ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya di bidang *machine learning* dan penerapannya dalam sektor perbankan.

Daftar Rujukan

[1] I. Indrastata, "Strategi Promosi Perbankan dengan Analisis dan Implementasi Algoritma Random Forest," *AnoaTIK J. Teknol. Inf. dan Komput.*, vol. 1, no. 2, pp. 54–58, 2023, doi: 10.33772/anoatik.v1i2.8.

[2] M. R. A. Riyyasy, W. N. Aghniya, dan H. Tantyoko, "Penerapan Algoritma Machine Learning Untuk Memprediksi Term Deposit Nasabah Perbankan," *LEDGER J. Inform. Inf. Technol.*, vol. 2, no. 3, pp. 145–156, 2023, doi: 10.20895/ledger.v2i3.1416.

[3] H. E. Prayoga, D. R. I. M. Setiadi, dan H. Kurniawan, "Penerapan Machine Learning dan Deep Learning untuk Memprediksi Langganan Produk Keuangan Berdasarkan Fitur Pelanggan," *J. Inovtek Polbeng - Seri Inform.*, vol. 10, no. 2, pp. 670–681, 2025, doi: 10.35314/cyvzwk14.

[4] Aurelia dan L. Hiryanto, "Sistem Rekomendasi Deposito Berjangka Menggunakan Metode Random Forest: Studi Kasus pada Bank Marketing UCI," *Comput. J. Comput. Sci. Inf. Syst.*, vol. 9, no. 1, pp. 1–8, 2025, doi: 10.24912/computatio.v9i1.32667.

[5] A. N. Puteri, Arizal, dan A. D. Achmad, "Feature Selection Correlation-Based pada Prediksi Nasabah Bank Telemarketing untuk Deposito," *Matrik J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 20, no. 2, pp. 335–342, 2021, doi: 10.30812/matrik.v20i2.1183.

[6] M. N. Tentua, Turiyan, dan K. Nurbagja, "Komparasi Metode Naïve Bayes dan C45 pada Prediksi Pelanggan Deposito Berjangka," *J. Din. Inform.*, vol. 11, no. 1, pp. 92–99, 2022, [Daring]. Tersedia pada: https://jdi.upy.ac.id/index.php/jdi/article/view/136.

[7] A. Mutiarachim dan J. S. P. Tyoso, "Optimasi Prediksi Pemasaran Nasabah Deposito Bank dengan Metode Klasifikasi Logistic Regression," *J. Cakrawala Inf. J.*, vol. 4, no. 1, pp. 20–28, 2024, doi: 10.54066/jci.v4i1.390.

[8] M. T. Syofyan, N. Amalita, D. Vionanda, dan D. Fitria, "Comparison of Distance Function in K-Nearest Neighbor Algorithm to Predict Prospective Customers in Term Deposit Subscriptions," *UNP J. Stat. Data Sci.*, vol. 1, no. 3, pp. 105–111, 2023, doi: 10.24036/ujsds/vol1-iss3/47.

[9] B. L. E. Panggabean dan F. Aziz, "Klasifikasi Nasabah Potensial menggunakan Algoritma Ensemble Least Square Support Vector Machine dengan AdaBoost," *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, pp. 269–274, 2023, doi: 10.30591/jpit.v8i3.5675.

[10] M. H. Saputra, H. Kurniawan, dan R. A. Andini, "Prediksi Nasabah Potensial Deposito Berjangka Menggunakan Regresi Logistik dan Support Vector Machine," *JOISTICS J. Inf. Syst. Data Anal.*, vol. 1, no. 1, pp. 1–12, 2025, [Daring]. Tersedia pada: https://journal.uyr.ac.id/index.php/JOISTICS/article/view/821/0.

[11] P. Mangande dan E. H. Parmadi, "Klasifikasi Data Nasabah yang Berpotensi Membuka Simpanan Deposito Menggunakan Algoritme Naive Bayes," in *Prosiding Seminar Nasional Teknologi dan Sains*, 2025, vol. 4, pp. 108–117, doi: 10.29407/8xyzw712.

[12] M. D. Gultom, R. D. Destianti, dan G. Noviamanti, "Penerapan Naïve Bayes dengan Balancing Data Smote untuk Peningkatan Prediksi Nasabah Deposito Bank," *J. Comput. Sci. Artif. Intell.*, vol. 2, no. 1, pp. 1–8, 2025, [Daring]. Tersedia pada: https://ruangjurnal.or.id/index.php/jcsai/article/view/121.

[13] M. A. Christian, "Komparasi Model Machine Learning untuk Prediksi Penerimaan Deposito Berjangka pada Kampanye Telemarketing Bank," in *Prosiding Seminar Nasional Indonesia*, 2026, vol. 4, no. 1, pp. 28–37, doi: 10.5281/zenodo.20309740.

[14] F. Izma dan H. Firmansyah, "Penerapan Algoritma Naive Bayes untuk Memprediksi Keputusan Berlangganan Deposito Berjangka pada Kampanye Pemasaran Langsung," *J. Din. Inform.*, vol. 15, no. 1, pp. 140–148, 2026, doi: 10.31316/jdi.v15i1.427.

[15] A. R. N. P. Wahyudi, "Penerapan Algoritma Decision Tree untuk Mengidentifikasi Karakteristik Nasabah yang Cenderung Berlangganan Deposito Berjangka," *EJECTS J. Comput. Technol. Informations Syst.*, vol. 5, no. 1, pp. 7–12, 2025, [Daring]. Tersedia pada: https://jurnal.unda.ac.id/index.php/ejects/article/view/556.

[16] R. Wahyudi, K. I. Manik, M. Alfin, J. B. Henrydunan, Arnita, dan F. Ramadhani, "Analisis Faktor Keputusan Nasabah Berlangganan Term Deposit: Perbandingan Random Forest dan XGBoost," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 4, pp. 5600–5607, 2025, doi: 10.36040/jati.v9i4.13881.

Submitted : 09-06-2026 | Reviewed : 17-06-2026 | Accepted : 05-08-2026