



Prediksi Nasabah Potensial Deposito Berjangka Menggunakan Regresi Logistik dan Support Vector Machine

M. Hardi Saputra^{*1}, Hendra Kurniawan², dan Rekha Aprilia Andini³

¹⁾ Program Studi Sains Data, Fakultas Ilmu Komputer, Institut Informatika dan Bisnis Darmajaya, kota Bandar Lampung, Indonesia

muhammadhardisaputra@gmail.com^{1*}; Rekhaaprilia06@gmail.com²

* Corresponding author: hendra.kurniawan@darmajaya.ac.id

Abstract

This research aims to identify potential time deposit customers using machine learning techniques based on logistic regression. Customer data is analyzed to determine the factors that influence the decision to choose time deposits. The target label distribution shows significant data imbalance, where the majority of customers do not choose time deposits. Therefore, data balancing techniques such as oversampling and undersampling were applied to improve the accuracy of the model. Model evaluation was performed using ROC curves to evaluate the predictive performance on the minority class. The results show that logistic regression combined with cross-validation techniques are able to provide accurate results and good generalizability. This research makes an important contribution in helping banks effectively identify potential customers so that they can optimize their time deposit marketing strategies.

Keywords : Logistic Regression, Term Deposit, Machine Learning

Abstrak

Penelitian ini bertujuan untuk mengidentifikasi calon nasabah deposito berjangka menggunakan teknik pembelajaran mesin berdasarkan regresi logistik. Data nasabah dianalisis untuk mengetahui faktor-faktor yang mempengaruhi keputusan memilih deposito berjangka. Distribusi label sasaran menunjukkan ketidakseimbangan data yang signifikan, dimana mayoritas nasabah tidak memilih deposito berjangka. Oleh karena itu, teknik penyeimbangan data, seperti oversampling dan undersampling diterapkan untuk meningkatkan akurasi model. Evaluasi model dilakukan dengan menggunakan kurva ROC untuk mengevaluasi kinerja prediktif pada kelas minoritas. Hasil penelitian menunjukkan bahwa regresi logistik yang dikombinasikan dengan teknik validasi silang mampu memberikan hasil yang akurat dan generalisasi yang baik. Penelitian ini memberikan kontribusi penting dalam membantu bank mengidentifikasi secara efektif pelanggan potensial sehingga mereka dapat mengoptimalkan strategi pemasaran deposito berjangka mereka.

Kata Kunci : Regresi Logistik, Deposito Berjangka, Pembelajaran Mesin

1. PENDAHULUAN



Industri perbankan Indonesia mengalami perkembangan yang signifikan, seiring dengan meningkatnya kebutuhan masyarakat akan layanan keuangan yang lebih inovatif. Salah satu produk yang banyak diminati adalah deposito, yang dipandang sebagai pilihan investasi yang aman dengan imbal hasil yang menarik[1]. Tantangan utama bagi bank adalah mengidentifikasi nasabah yang berpotensi untuk memilih produk ini, sehingga dapat meningkatkan efektivitas pemasaran dan kepuasan nasabah [2]

Metode analitik seperti logistic regression telah terbukti efektif dalam memprediksi keputusan biner, seperti pilihan nasabah terhadap produk deposito. Dengan memanfaatkan data demografis, pendapatan, riwayat transaksi, dan preferensi investasi, model ini dapat memberikan prediksi yang akurat mengenai perilaku nasabah [3]. Penelitian sebelumnya menunjukkan bahwa faktor-faktor seperti usia, pendidikan, dan jenis pekerjaan berpengaruh signifikan terhadap keputusan investasi nasabah [4]

Untuk memastikan keandalan model prediksi, teknik cross-validation diterapkan. Teknik ini membagi dataset menjadi beberapa subset untuk melatih dan menguji model secara berulang, sehingga meminimalkan risiko overfitting dan meningkatkan kemampuan generalisasi model pada data baru [5]. Dengan adanya pendekatan ini, bank dapat lebih mengenal preferensi nasabah dan merancang strategi pemasaran yang lebih tepat sasaran[6].

Penelitian ini bertujuan untuk menerapkan kombinasi metode logistic regression dan cross-validation dalam mengidentifikasi nasabah potensial untuk produk tabungan deposito. Hasil dari penelitian ini diharapkan dapat memberikan wawasan baru bagi bank di Indonesia dalam pengambilan keputusan strategis yang lebih berbasis data dan sesuai dengan kebutuhan pasar[7].

2. TINJAUAN PUSTAKA

Data yang digunakan dalam penelitian ini berasal dari kampanye pemasaran langsung yang dilakukan oleh sebuah lembaga perbankan di Portugal. Kampanye pemasaran tersebut menggunakan metode panggilan telepon untuk menjangkau klien. Dalam banyak kasus, diperlukan lebih dari satu kali kontak dengan klien yang sama untuk menentukan apakah mereka akan menyetujui berlangganan produk yang ditawarkan, yaitu deposito berjangka bank (dengan hasil 'Ya') atau tidak menyetujui ('Tidak').

Tabel 2.1 Informasi Variabel

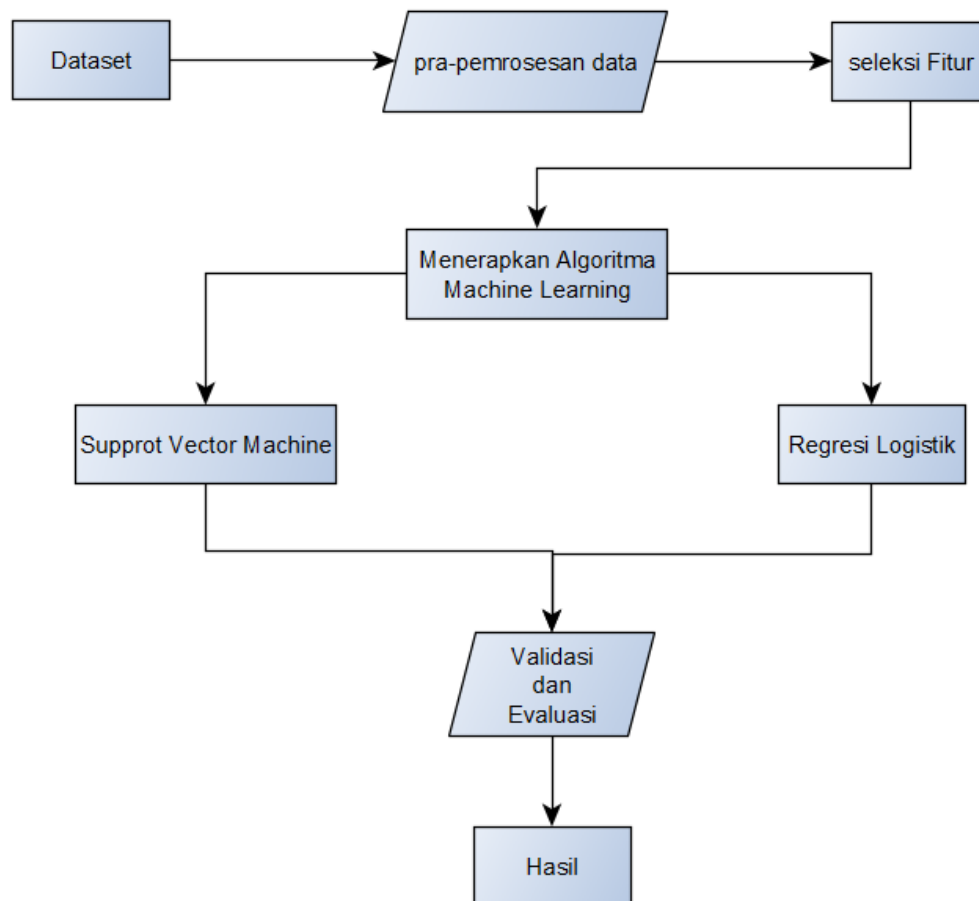
Nama Variabel	Keterangan
Umur	jenis pekerjaan (kategori: 'admin', 'pekerja kasar', 'wirausahawan', 'pembantu rumah tangga', 'manajemen', 'pensiunan', 'wiraswasta', 'jasa', 'mahasiswa', 'teknisi', 'pengangguran', 'tidak diketahui')



Pekerjaan	jenis pekerjaan (kategori: 'admin', 'pekerja kasar', 'wirausahawan', 'pembantu rumah tangga', 'manajemen', 'pensiunan', 'wiraswasta', 'jasa', 'mahasiswa', 'teknisi', 'pengangguran', 'tidak diketahui')
Pendidikan	dasar.4y', 'dasar.6y', 'dasar.9y', 'sekolah menengah', 'buta.huruf', 'kursus profesional', 'gelar universitas', 'tidak diketahui'
Status	status perkawinan (kategori: 'bercerai', 'menikah', 'lajang', 'tidak diketahui'; catatan: 'bercerai' berarti bercerai atau janda/duda)
Default	Memiliki kredit macet?

Sumber: <https://archive.ics.uci.edu/dataset/222/bank+marketing>

Langkah-langkah penelitian untuk mengidentifikasi nasabah potensial deposito berjangka dilakukan secara terstruktur. Penelitian ini menggunakan metode regresi logistik dan Support Vector Machine (SVM), dapat dilihat pada Gambar 1 berikut.



Gambar 1. Kerangka Penelitian



Gambar 1 menggambarkan langkah-langkah analisis data untuk memprediksi apakah klien akan setuju atau tidak berlangganan deposito berjangka bank dengan hasil('Ya' atau 'Tidak'). Proses dimulai dengan pengolahan data awal dan preprocessing untuk mempersiapkan data yang bersih dan siap dianalisis. Selanjutnya, dilakukan seleksi variabel untuk memilih faktor-faktor yang paling relevan dalam keputusan klien. Setelah itu, algoritma prediksi seperti regresi logistik dan Support Vector Machine (SVM) dipilih untuk membangun model prediksi. Validasi model dilakukan untuk mengukur akurasi dan performa menggunakan data uji, dan akhirnya, evaluasi hasil dilakukan untuk menarik kesimpulan serta memberikan rekomendasi berbasis analisis.

3. METODE

Pendekatan ini menggunakan pembelajaran mesin, yang memanfaatkan metode statistik untuk mengenali pola data, dengan regresi logistik dan SVM sebagai model yang sering diterapkan dalam klasifikasi calon pelanggan.[8], [9].

3.1. *Algoritma Machine Learning*

Pada bagian ini menjelaskan secara mendalam dua pendekatan yang digunakan untuk mengidentifikasi nasabah potensial deposito berjangka, yaitu metode Regresi Logistik dan Support Vector Machine (SVM). Penjelasan berikut akan menguraikan prinsip dasar, penerapan, dan kelebihan masing-masing algoritma dalam konteks analisis data [9], [10], [11], [12], [13], [14].

1. *Support Vector Machine (SVM)*

Support Vector Machine (SVM) merupakan teknik pembelajaran mesin yang berlandaskan pada teori statistika dan digunakan untuk tujuan klasifikasi serta regresi. Metode ini berfokus pada pemisahan dua kelas data dengan cara memaksimalkan jarak antara keduanya menggunakan hiperbidang yang optimal[15]. SVM dapat menangani data yang tidak dapat dipisahkan secara linier dengan menggunakan fungsi kernel, seperti kernel linear, polinomial, dan radial basis function (RBF), yang memungkinkan pemetaan data ke dimensi yang lebih tinggi[16]. Algoritma ini sangat efektif untuk bekerja dengan data berdimensi tinggi, tetapi memerlukan penyesuaian parameter dan pengolahan data untuk menghindari masalah overfitting. SVM diakui karena kemampuannya dalam membangun model yang kuat dan efisien, meskipun penyetelan parameter yang tepat sangat penting untuk memperoleh hasil yang optimal[17].

2. *Logistic Regression*

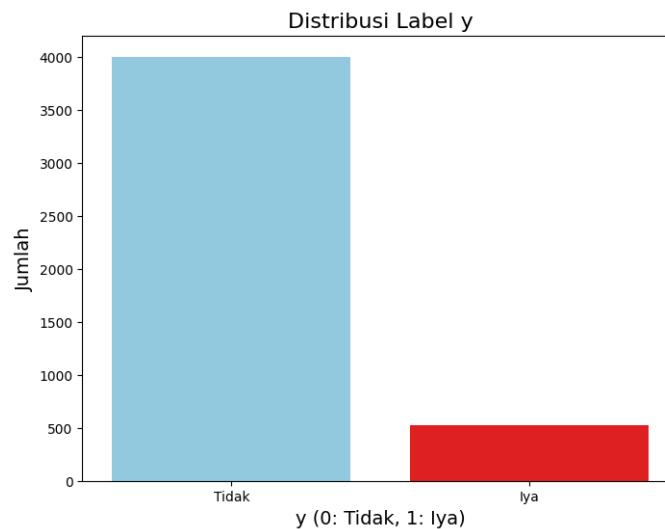
Regresi logistik adalah metode statistika yang digunakan untuk memperkirakan probabilitas terjadinya suatu kejadian yang memiliki dua kemungkinan hasil,yaitu berdasarkan satu atau lebih variabel independen[18]. Metode ini dipilih karena kemampuannya untuk memberikan interpretasi dalam bentuk probabilitas serta koefisien yang menunjukkan seberapa kuat hubungan antara variabel independen dan variabel yang diprediksi[19]. Regresi logistik biasanya digunakan ketika



hasil prediksi berupa nilai biner, dengan kemungkinan hasil yang dapat diukur dalam rentang probabilitas, seperti (0) untuk tidak terjadi dan (1) untuk terjadi[20].

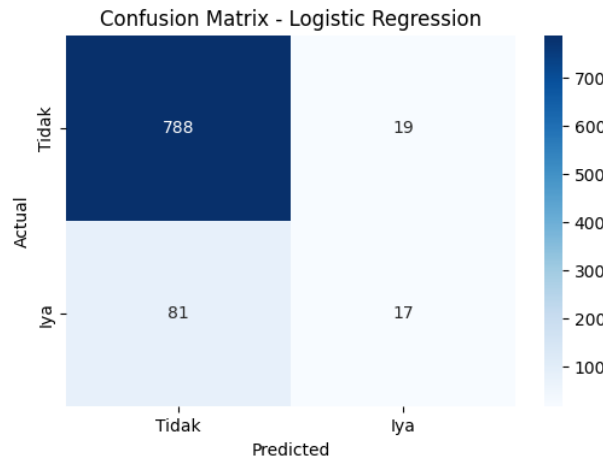
4. HASIL DAN PEMBAHASAN

Dalam analisis machine learning ini, grafik digunakan untuk menunjukkan distribusi label Y, yang menggambarkan jumlah deposito yang disetujui dan tidak disetujui. Dengan menggunakan grafik batang, perbandingan antara kedua kategori dapat dilihat secara jelas, serta mengidentifikasi ketidakseimbangan kelas dalam data, seperti yang ditampilkan pada Gambar 1 berikut:

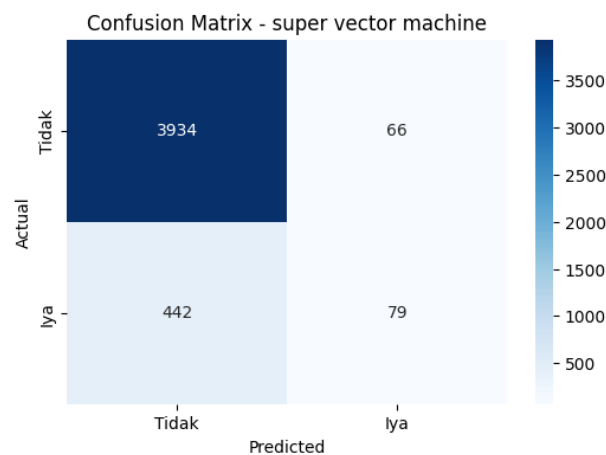


Gambar 1. Distribusi tabel Y

Gambar 1 menunjukkan sebaran label target pada data nasabah bank. Sebagian besar nasabah, sekitar 4000 orang, termasuk dalam kategori “Tidak” (tidak mempunyai deposito). Sebaliknya, kurang dari 500 orang termasuk dalam kategori “Ya” (memiliki deposito). Ketidakseimbangan ini membuat model prediksi menjadi sulit karena cenderung lebih akurat pada kelas mayoritas dan kurang akurat pada kelas minoritas. Untuk mengatasi masalah ini, teknik seperti oversampling pada kelas minoritas, sub-sampling pada kelas mayoritas, atau SMOTE dapat diterapkan. Algoritma seperti Random Forest dan XGBoost juga efektif dengan penyetelan parameter yang tepat. Sejalan dengan hal tersebut, metode yang diterapkan dalam analisis ini adalah Regresi Logistik dan SVM, yang dilatih untuk mengidentifikasi nasabah potensial deposito. Evaluasi model dilakukan dengan data uji menggunakan metrik. Oleh karena itu, Hasil evaluasi disajikan dalam *confusion matrix*, yang memberikan gambaran kinerja model dalam mengklasifikasikan nasabah yang berpotensi atau tidak. Perhatikan hasil *confusion matrix* pada Gambar 2 dan Gambar 3.



Gambar 2. Confusion Matrix Regresi Logistik



Gambar 3. Confusion Matrix SVM

Berdasarkan hasil yang ditunjukkan pada confusion matrix yaitu pada Gambar 2 (Regresi Logistik) dan Gambar 3 (SVM), dapat dianalisis bahwa kedua model memperlihatkan kinerja yang berbeda dalam mengolah data yang sangat tidak seimbang, berkaitan dengan sebaran nasabah bank pada kategori "Ya" (memiliki deposito) dan "Tidak" (tidak memiliki deposito). Pada Gambar 2, regresi logistik menunjukkan kinerja yang lebih baik dalam meminimalkan kesalahan prediksi pada kategori "Ya", menghasilkan 17 true positif (TP) yang artinya, 17 nasabah yang melakukan penyetoran telah diprediksi dengan benar. Selain itu, model ini mampu mengidentifikasi 788 true negative (TN), yaitu nasabah yang sebenarnya tidak memiliki simpanan dan prediksinya tepat. Namun model ini masih menghadapi tantangan yang cukup besar, terbukti dengan masih tingginya angka false negative (FN) yaitu 81 nasabah yang memiliki simpanan



salah memprediksi tidak ada simpanan, serta 19 false positif (FP) yang menunjukkan kepada nasabah bahwa simpanan mereka diperkirakan salah karena tidak ada simpanan.

Sedangkan pada Gambar 3, model SVM menunjukkan kemampuan yang sangat baik dalam mengenali kategori mayoritas, sehingga menghasilkan jumlah true negative (TN) yang jauh lebih besar yaitu 3.934, menunjukkan bahwa model ini masih mampu memprediksi nasabah tanpa simpanan. Di sisi lain, model ini juga menghasilkan true positif (TP) sebesar 79 yang berarti model berhasil. Kenali pelanggan dengan simpanan lebih baik daripada regresi logistik. Namun tantangan utama SVM adalah banyaknya false negative (FN) yang mencapai 442 yang menandakan banyak nasabah simpanan yang salah karena tidak memiliki simpanan, dan juga 66 positif (FP) yaitu nasabah tanpa simpanan. salah berencana untuk tidak memiliki deposito. meskipun regresi logistik membuat kesalahan perkiraan lebih sedikit dalam kategori "Ya", masalah dengan banyaknya negatif palsu tetap penting. Sebaliknya, SVM unggul dalam mengenali kategori mayoritas ("tidak"), namun terdapat kesulitan besar dalam kategori minoritas ("Ya"), dengan FN yang jauh lebih tinggi. Pengoperasian kedua model ini menyoroti tantangan utama dalam menangani ketimpangan data, terutama dalam mengenali kategori minoritas yang menjadi perhatian utama analisis ini. Oleh karena itu, diperlukan upaya tambahan seperti penyeimbangan data melalui teknik Oversampling, undersampling atau penerapan algoritma khusus seperti SMOTE, serta penyesuaian ambang batas untuk meningkatkan sensitivitas terhadap kategori minoritas dan memastikan hasil prediksi yang lebih akurat di kedua kelas.

Analisis data ini menggunakan algoritma regresi logistik dan SVM. Kedua model dilatih dengan data pelatihan dan dievaluasi dengan data uji untuk mengukur akurasi serta kinerjanya menggunakan metrik seperti akurasi, presisi, recall, dan AUC-ROC. Perhatikan Tabel 1 berikut.

Tabel 1. Data Peforma Machine Learning

Evaluasi/Model	Model	
	LR	SVM
Akurasi	0.83	0.83
Presisi	0.82	0.81
Recall	0.63	0.61
F1-scores	0.71	0.69

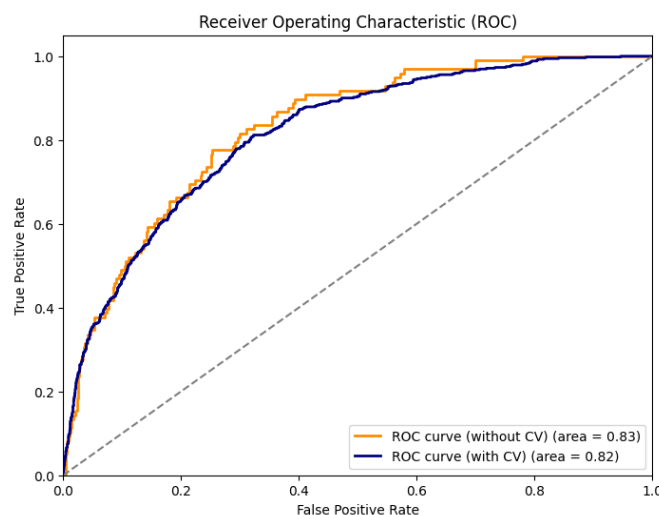
Tabel 1 menunjukkan perbandingan kinerja antara dua metode pembelajaran mesin, yaitu Regresi Logistik dan Support Vector Machine (SVM), untuk memprediksi properti simpanan nasabah bank, yang penting untuk mengidentifikasi calon nasabah dengan lebih tepat. Kedua



model tersebut menunjukkan tingkat akurasi yang sama yaitu 0,83 yang berarti mampu memprediksi dengan tepat pada 83% kasus apakah nasabah memiliki deposito berjangka (kategori “Ya”) atau (kategori “Tidak”). Metrik Akurasi yang hampir serupa menunjukkan kemampuan kedua model dalam menghasilkan prediksi positif yang akurat, meskipun terdapat perbedaan yang signifikan dalam ukuran recall dan F1-scores, mencerminkan bahwa beberapa nasabah yang saat ini memiliki simpanan (“Ya”) masih hilang dari proses identifikasi.

Akurasi yang tinggi dari kedua model mencerminkan potensinya dalam memberikan informasi berharga untuk membantu bank mengidentifikasi nasabah potensial adalah baik untuk memiliki deposito yang tetap. Namun, ketidakmampuan model untuk secara konsisten mengenali pelanggan dalam kategori “Ya”, yang dibuktikan dengan rendahnya nilai recall pada kategori ini, dapat mengurangi efektivitas perkiraan dalam menentukan target yang relevan untuk promosi atau penawaran. Oleh karena itu, diperlukan upaya optimasi model, seperti penyeimbangan data melalui teknik oversampling, penyesuaian ambang batas, atau penggunaan algoritma yang lebih spesifik pada data. tidak seimbang, sehingga kedua model tersebut tidak hanya mampu mempertahankan akurasi yang tinggi, namun juga meningkatkan kinerja dalam pengenalan kategori minoritas (nasabah dengan simpanan tetap). Optimalisasi ini penting untuk memastikan pengambilan keputusan berbasis data yang lebih tepat dan mendukung strategi bisnis bank secara keseluruhan.

Selanjutnya, pada *Receiver Operating Characteristic* (ROC) dan *Area Under the Curve* (AUC) merupakan metode penting untuk menilai kinerja model. Kurva ROC memberikan gambaran mengenai kemampuan model dalam membedakan antara kelas positif dan negatif. Perhatikan Gambar 4 berikut:



Gambar 4. Kurva ROC untuk Dua Algoritma Machine Learning



Berdasarkan grafik ROC (Receiver Operating Characteristic) yang diberikan, berikut adalah tabel perbandingan untuk kurva ROC dengan dan tanpa Cross-Validation (CV):

Tabel 2. Perbandingan Dengan dan Tanpa Cross-Validation

Metode	AUC (Area Under Curve)	Keunggulan
ROC Curve (tanpa CV)	0.83	Performa model lebih tinggi tanpa CV.
ROC Curve (dengan CV)	0.82	Model lebih stabil dan generalisasi baik.

Gambar 4 menunjukkan kurva ROC untuk dua model prediksi nasabah yang melakukan penarikan simpanan, dengan model Random Forest berwarna biru dan model regresi logistik berwarna oranye. Model Random Forest dengan nilai AUC sebesar 0,87 menunjukkan kinerja yang sedikit lebih baik dibandingkan model regresi logistik yang memiliki nilai AUC sebesar 0,85. Meskipun perbedaan AUC kedua model tidak terlalu besar, namun model Random Forest lebih unggul dari segi hasil. mengidentifikasi nasabah yang akan menarik simpanan, dengan true positive rate (TPR) yang lebih tinggi dan false positive rate (FPR) yang lebih rendah. Artinya model Random Forest lebih efektif dalam mengidentifikasi nasabah yang sebaiknya memanfaatkan penawaran khusus untuk menghemat dananya, sekaligus meminimalkan kesalahan saat menawarkan penawaran kepada nasabah yang tidak menarik simpanan. Di sisi lain, meskipun model regresi logistik juga memberikan hasil yang baik dengan AUC mendekati 0,85, model ini cenderung menghasilkan lebih banyak false positif, yaitu mengklasifikasikan nasabah yang tidak menarik simpanannya sebagai nasabah yang menginginkannya. Hal ini dapat menyebabkan bank menawarkan penawaran yang tidak perlu kepada nasabah sehingga menjadi kurang efisien. Namun model regresi logistik masih memberikan kinerja yang memadai dalam memprediksi nasabah akan menarik simpanan, dan kinerjanya juga dapat ditingkatkan dengan penyesuaian hyperparameter atau mengeksplorasi model lain yang lebih sesuai dengan karakteristik data.

5. KESIMPULAN

Penelitian ini menunjukkan bahwa perbandingan kinerja antara model regresi logistik dan SVM memiliki potensi besar dalam memprediksi deposito, di mana kedua model menunjukkan akurasi yang tinggi, yang berarti keduanya dapat memprediksi nasabah dengan atau tanpa deposito dengan baik. Namun, meskipun keduanya mencapai akurasi yang tinggi, terdapat perbedaan dalam kemampuannya untuk mengidentifikasi nasabah yang memiliki deposito tepat waktu (kategori "Ya"). Pengukuran recall yang lebih rendah dan F1-scores yang lebih kecil menunjukkan bahwa meskipun nasabah tanpa deposito lebih sering diprediksi dengan benar, masih banyak nasabah yang memiliki deposito yang terlewat.



Kondisi ini menunjukkan adanya tantangan dalam pengelolaan ketimpangan data, dimana kategori minoritas (nasabah yang mempunyai simpanan) lebih sulit diidentifikasi. Oleh karena itu, meskipun kedua model menunjukkan potensi yang baik untuk mengidentifikasi nasabah yang tidak memiliki simpanan, namun perlu dilakukan langkah perbaikan seperti penyeimbangan data dengan teknik oversampling atau undersampling, serta penyesuaian ambang batas untuk meningkatkan sensitivitas model terhadap minoritas kategori. Langkah ini penting untuk memastikan bahwa model tersebut tidak hanya memiliki presisi tinggi, tetapi juga mampu mengenali pelanggan. Terdapat potensi untuk memiliki simpanan berjangka yang lebih akurat. Dengan demikian, bank dapat memanfaatkan prediksi tersebut untuk merancang strategi yang lebih tepat dalam menawarkan produk deposito berjangka kepada nasabah yang sesuai, yang pada gilirannya akan meningkatkan efektivitas dan efisiensi dalam pengelolaan produk keuangan.

DAFTAR PUSTAKA

- [1] N. Wahyu and F. & Segaf, “SEBERAPA PENGARUH PENETAPAN NISBAH BAGI HASIL, INFLASI, DAN JUMLAH UANG BEREDAR TERHADAP DEPOSITO MUDHARABAH DI INDONESIA”.
- [2] A. Imran, ; Bakkareng, ; Yulistia, J. Manajemen, and F. Ekonomi, “ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI PERILAKU TERHADAP KEPUTUSAN NASABAH DALAM MEMILIH PRODUK PINJAMAN DI PT. BRI UNIT TIKU CABANG BUKITTINGGI,” *JM*, vol. 3, no. 4, pp. 662–677, 2021.
- [3] “DOC-20241209-WA0032.”.
- [4] M. Rasikh, A. Riyyasy, W. Nouval Aghniya, and H. Tantyoko, “LEDGER: Journal Informatic and Information Technology Penerapan Algoritma Machine Learning Untuk Memprediksi Term Deposit Nasabah Perbankan.”
- [5] H. Muhammad Nur, V. Maarif, F. Fandi Dwi Imaniawan, and E. Saputro, “Prediksi Keputusan Berdonasi Pada Website Charity Box dengan Toko Waralaba Menggunakan Operator Cross Validation dan Algoritma Decision Tree,” *Jurnal Sains dan Manajemen*, vol. 12, no. 1, 2024.
- [6] L. A. Afifah, A. M. Kosim, H. Hakiem, U. Ibn, K. Bogor ', and L. Com, “037 | V o l u m e 4 N o m o r 4 2 0 2 3 Strategi Pemasaran Produk Pembiayaan Cicil Emas di Bank Syari'ah Indonesia: Studi Kasus Bank Syari'ah Indonesia KCP Sudirman,” vol. 4, no. 4, 2023.
- [7] R. N. Irawan, K. M. Hindrayani, and M. Idhom, “Penerapan Cross Validation sebagai Analisis Sentimen Pelayanan Publik Kereta Api Lokal Daop 8 Menggunakan Metode Multinomial Naïve Bayes,” *G-Tech: Jurnal Teknologi Terapan*, vol. 8, no. 2, pp. 954–963, Apr. 2024, doi: 10.33379/gtech.v8i2.4117.



- [8] M. E. Purbaya, A. F. Nugraha, S. Gustina, and M. K. Azis, “Meta-Algorithms untuk Meningkatkan Kinerja Klasifikasi dalam Keberhasilan Telemarketing Perbankan,” *Techno.Com*, vol. 19, no. 4, pp. 385–396, 2020, doi: 10.33633/tc.v19i4.3725.
- [9] N. Charles, A. Bernard Wijaya, and D. Jollyta, “ANALISIS DATA SEWA SEPEDA DI SEOUL DENGAN VARIATIF METODE KLASIFIKASI,” *Jurnal Algoritme*, vol. 4, no. 2, pp. 53–62, 2024, doi: 10.35957/algoritme.xxxx.
- [10] C. Agustina, “Analisa Nasabah Potensial Tabungan Deposito Berjangka Menggunakan Teknik Klasifikasi Data Mining,” *Jurnal Teknologi Informasi dan Terapan*, vol. 5, no. 2, pp. 105–112, 2019, doi: 10.25047/jtit.v5i2.88.
- [11] N. Purwati, W. H. Pramujati, S. Syakur, and E. Safitri, “Prediksi Pasien Pusat Kesehatan Masyarakat Menggunakan Machine Learning,” *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 12, no. 3, p. 578, Jul. 2024, doi: 10.26418/justin.v12i3.80135.
- [12] N. Purwati, W. H. Pramujati, S. Syakur, and E. Safitri, “Prediksi Pasien Pusat Kesehatan Masyarakat Menggunakan Machine Learning,” *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 12, no. 3, p. 578, Jul. 2024, doi: 10.26418/justin.v12i3.80135.
- [13] D. Rofianto, E. Safitri, K. Amaliah, J. Fitra, and A. Hijriani, “Cyber Threat Detection Using an Ensemble Model Approach for Phishing Website Identification ARTICLE INFORMATION ABSTRACT,” 2024. [Online]. Available: <http://innovatics.unsil.ac.id>
- [14] E. Safitri, R. Rizalnul Fikri, and R. Nurlistiani, “Application of Ensemble Machine Learning for Infectious Diseases with Vaccine Intervention: A Global COVID-19 Case Study,” vol. 16, no. 4, pp. 819–836, 2024, doi: 10.20895/INFOTEL.V16I3.1263.
- [15] A. Setiawan and S. Mulyati, “ANALISIS SENTIMEN PENGGUNA SHOPEEPAYLATER PADA TWITTER MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM),” 2024AD, [Online]. Available: <https://journal.mediapublikasi.id/index.php/logic>
- [16] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, “Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 153–160, Oct. 2023, doi: 10.57152/malcom.v3i2.897.
- [17] T. Tan, H. Sama, G. Wijaya, and O. E. Aboagye, “Studi Perbandingan Deteksi Intrusi Jaringan Menggunakan Machine Learning: (Metode SVM dan ANN) Comparative Study of Network Intrusion Detection Using Machine Learning: (SVM and ANN Method),” vol. 13, 2023, doi: 10.34010/jati.v13i2.
- [18] A. T. R. DANI, V. RATNASARI, L. NI’MATUZZAHROH, I. C. AVIANTHOLIB, R. NOVIDIANTO, and N. Y. ADRIANINGSIH, “ANALISIS KLASIFIKASI ARTIST MUSIC MENGGUNAKAN MODEL REGRESI LOGISTIK BINER DAN ANALISIS DISKRIMINAN,”



JOISTICS: Journal of Information Systems and Data Analytics

Copyright ©2025, Vol. 1, No. 1, Mei 2025, pp. 001-012

p-ISSN: xxxx-xxxx, e-ISSN: xxxx-xxxx, DOI: xx.xxxxx/xxxx.xxxx.x

On-line: <https://journal.uyr.ac.id/index.php/JOISTICS>

Jambura Journal of Probability and Statistics, vol. 3, no. 1, pp. 1–10, May 2022, doi: 10.34312/jjps.v3i1.13708.

- [19] M. Valentina, R. Hadi, Y. Rosaripatria, S. I. Oktora, and P. S. Stis, “DETERMINAN PENGANGGURAN TERDIDIK DI PROVINSI NUSA TENGGARA TIMUR (NTT) TAHUN 2018 MENGGUNAKAN REGRESI LOGISTIK BINER,” 2021.
- [20] A. Ilham Sofiyat and A. Tjalla, “PEMODELAN REGRESI LOGISTIK BINER TERHADAP PENERIMAAN PEGAWAI DI PT XYZ JAKARTA,” 2023.