

Prediksi Keberhasilan Pemasaran Layanan Jasa Perbankan Menggunakan Algoritma Logistic Regreesion

Sari Dewi^{1*}, Hanggaro Aji Al Kautsar², Dwi Yuni Utami³

¹Program Studi Sistem Informasi PSDKU Kota Pontianak, Fakultas Teknik dan Informatika,
Universitas Bina Sarana Informatika,
Jl. Abdul Rahman Saleh No.18, Kalimantan Barat 78124, Indonesia

^{2,3}Program Studi Teknologi Komputer, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika
Jl. Kramat Raya No 98 Jakarta Pusat 10450, Indonesia

e-mail: 1sari.sre@bsi.ac.id, 2Hanggoro.hgr@bsi.ac.id, 3Dwi.dyu@bsi.ac.id

(*) Corresponding Author

Artikel Info : Diterima : 02-05-2023 | Direvisi : 28-06-2023 | Disetujui : 27-07-2023

Abstrak - Menentukan minat masyarakat dalam memasarkan jasa perbankan dengan menggunakan teknik data mining. Segmentasi prospek merupakan salah satu proses yang digunakan dalam strategi pemasaran industri perbankan. Dukungan data mining memainkan peran penting dalam mengklasifikasikan nasabah bank potensial dan mengevaluasi keberhasilan pemasaran layanan mereka. Hal ini penting untuk mendukung kesimpulan tentang tingkat keberhasilan telemarketer dalam menjalankan tugas pemasaran bank. produk yang cara kerjanya membutuhkan informasi tentang pelanggan potensial. Ini adalah teknik klasifikasi yang sering digunakan untuk mengklasifikasikan prospek menggunakan regresi logistik menurut peta penelitian dukungan data mining prospek. Mendefinisikan algoritma klasifikasi penambangan data yang akurat untuk memprediksi keberhasilan telemarketing berdasarkan eksperimen 2010. dalam pemasaran produk jasa perbankan, hasil proses evaluasi algoritma ini ditentukan dengan cross-validation, Confusion Matrix, ROC curve dan T-test. Algoritma regresi logistik lebih akurat dengan akurasi 92,32 % dan nilai AUC 0,962, sehingga algoritma yang digunakan termasuk dalam kelompok klasifikasi yang baik.

Kata Kunci : Algoritma, data mining, *regresi logistik*.

Abstracts - Determining public interest in marketing banking services using data mining techniques. Prospect segmentation is one of the processes used in the marketing strategy of the banking industry. Data mining support plays an important role in classifying potential bank customers and evaluating the success of marketing their services. This is important to support the conclusion about the success rate of telemarketers in carrying out bank marketing tasks. a product whose way of working requires information about potential customers. This is a classification technique that is often used to classify prospects using logistic regression according to research maps supporting prospect data mining. Defining an accurate data mining classification algorithm to predict telemarketing success based on a 2010 experiment. In marketing banking service products, the results of the evaluation process of this algorithm are determined by cross-validation, Confusion Matrix, ROC curve and T-test. The logistic regression algorithm is more accurate with an accuracy of 92.32% and an AUC value of 0.962, so the algorithm used is included in the good classification group.

Keywords: Algoritma, Data Mining, *Logistic regression*.

PENDAHULUAN

Jasa perbankan adalah jasa yang ditawarkan kepada nasabah yang menarik perhatian dan memuaskan keinginan dan kebutuhannya. Dipercayai bahwa manusia kurang memiliki kemampuan untuk memahami kebutuhan nasabah bank (Sigit dan Soliha, 2017). Pemasaran juga memenuhi permintaan dan keinginan konsumen. Fungsi telemarketing dalam proses pemasaran sangat erat kaitannya. Penggunaan teknologi telekomunikasi sebagai komponen pemasaran yang rutin dan terstruktur disebut telemarketing. Prospecting, menjual ke pelanggan saat ini dan menerima pesanan dan menjawab pertanyaan melalui telepon, telemarketing melalui telepon dan pusat panggilan. Dengan bantuan telemarketing, perusahaan dapat meningkatkan penjualan, mengurangi biaya kegiatan



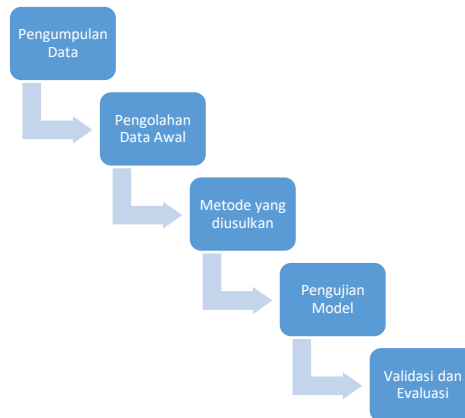
penjualan dan meningkatkan kepuasan pelanggan. Dengan kata lain, telemarketing menggunakan telepon sebagai alat pemasaran langsung untuk berbagai kegiatan riset penjualan dan pasar. Pelanggan penyandang disabilitas yang dibatasi oleh waktu atau jarak masih dapat membeli program perlindungan yang diperlukan untuk melindungi pelanggan dan keluarganya berkat penawaran yang dilakukan melalui saluran telemarketing. Layanan ini diberikan untuk mendukung kelancaran penggalangan dana dan penyaluran dana untuk pembiayaan langsung dan tidak langsung (Yoga Religia et al., 2021).

Perusahaan menggunakan penambangan data untuk mengubah data yang tidak terstruktur menjadi data yang informatif. Dalam dunia bisnis, pembelajaran yang lebih dalam dapat digunakan untuk membantu calon pelanggan mengembangkan strategi pemasaran yang lebih sukses, meningkatkan penjualan, dan memangkas biaya dengan menggunakan perangkat lunak untuk menemukan data bersama. Pengumpulan, penyimpanan, dan pemrosesan komputer yang efektif merupakan prasyarat untuk penambangan data. Penambangan data sering disebut sebagai penemuan data dan penemuan data. Business Intelligence dan riset pemasaran memungkinkan penambangan data dalam pemasaran (Oktanisa dan Supianto, 2018). Mencapai keunggulan kompetitif memerlukan rencana untuk merampingkan proses pemasaran. Untuk data mining strategi pemasaran, salah satunya adalah dengan menggunakan database pemasaran untuk menemukan informasi baru untuk pengambilan keputusan (Wardhani et al., 2019). Untuk meningkatkan loyalitas nasabah bank, metode database market mining digunakan untuk menganalisis prospek nasabah potensial, mengoptimalkan saluran penjualan, produk konkuren, karakteristik konsumen dan lainnya. Oleh karena itu, bagian pemasaran harus dapat menentukan strategi pemasaran yang terbaik untuk produk dan layanan perbankan yang ditawarkan kepada nasabah, khususnya yang berkaitan dengan kredit. Ini dimungkinkan jika informasi yang dikumpulkan cukup untuk membuat keputusan pemasaran. Salah satunya meramalkan kriteria pelanggan untuk mencapai tujuan pemasaran. Mengoptimalkan tujuan telemarketing merupakan masalah besar di sektor perbankan di bawah tekanan pertumbuhan omset dan pengurangan biaya..

Pemasaran dapat menggunakan penambangan data untuk menilai pasar agar lebih memahami pasar saat ini atau menemukan peluang bisnis yang belum dimanfaatkan, seperti mengantisipasi kebutuhan pasar sasaran. Kurva ROC (Receiver Operating Characteristic) secara visual membandingkan klasifikasi dan menunjukkan akurasi. Di ROC, matriks kebingungan diekspresikan (Dewi et al., 2022). Positif palsu ditampilkan sebagai garis horizontal di plot ROC, dan positif sejati digunakan untuk membandingkan kinerja berbagai pendekatan. Metode klasifikasi adalah salah satu teknik penambangan data yang digunakan di sebagian besar aplikasi, dan kurva ROC adalah alat lain untuk mengevaluasi kinerja klasifikasi (Pangaribuan et al., 2017). Dua jenis data diperlukan untuk proses klasifikasi: data pelatihan dan data uji. Penambangan data menggunakan data pelatihan. Teknisi mengklasifikasikan data uji dan data lainnya untuk mempelajari klasifikasi menggunakan metrik (Ardi Ramdani et al., 2022). Kami membagi dua kumpulan data artikel teks menjadi kumpulan pelatihan dan pengujian. Untuk mengetahui apakah prospek tersebut berpotensi menjadi nasabah kredit produktif atau tidak, penelitian ini berfokus pada penggunaan data mining untuk memprediksi keberhasilan telemarketing bank untuk menarik calon nasabah bank ke berbagai produk layanan bank. Hal ini dilakukan dengan membandingkan algoritma klasifikasi yang digunakan pada penelitian sebelumnya yang mengukur keakuratan dataset pemasaran bank. Dengan mengklasifikasikan preferensi nasabah melalui data mining, penelitian ini mengkaji efektivitas telemarketing yang dilakukan bank (Subarkah et al., 2020). Pohon klasifikasi dan regresi (CART) dan algoritma Naive Baye digunakan sebagai metode. Hasil ini menunjukkan bahwa algoritma CART lebih baik dalam memprediksi keputusan prabayar pelanggan telemarketing. Algoritma yang digunakan adalah CART dan Naive Bayes digunakan untuk menyelesaikan masalah tersebut. Oleh karena itu, penulis bermaksud menguji keefektifan pemodelan regresi logistik.

METODE PENELITIAN

Proses pemecahan masalah penelitian meliputi pengumpulan data, perumusan hipotesis atau proposisi, pengujian hipotesis, pembacaan hasil, dan penarikan kesimpulan lain yang dievaluasi secara mandiri. Meskipun empat metode penelitian banyak digunakan yaitu penelitian tindakan, eksperimen, studi kasus dan survei (Putro et al., 2021). Karya ini menggunakan penelitian eksperimental. Penelitian eksperimen meneliti pengolahan parameter atau variabel dari sudut pandang peneliti, dengan menggunakan metodologi penelitian berikut dan menggunakan tes yang dilakukan oleh peneliti sendiri: (Ardi Ramdani et al., 2022)



Sumber : (Widaningsih, 2019)

Gambar 1. Alur metode Penelitian.

Pengumpulan data (data yang diperoleh dari dataset Uci) dan pengolahan data di Excel merupakan langkah awal dalam metodologi penelitian. Data atau atribut yang digunakan kemudian digunakan untuk menghitung nilai bersih yang digunakan. Identifikasi dan pemilihan atribut, penanganan nilai atribut yang tidak lengkap atau hilang (menangani nilai yang hilang), dan prosedur diskritisasi nilai merupakan bagian dari tahap pra-pemrosesan data (Yolanda & Fahmi, 2021). Metode yang digunakan yaitu Regresi Logistik termasuk dalam kategori pengelolaan data. Selanjutnya dilakukan perancangan pengujian menggunakan Metode Rapid, dan didapatkan nilai AUC dan nilai akurasi. Setelah diperoleh, kita bandingkan dengan menggunakan metode T-test. (Hendrian, 2018)

3.1. Pengumpulan Data

Pengumpulan data primer dan pengumpulan data sekunder adalah dua kategori metode pengumpulan data yang berbeda. Data yang pertama kali dikumpulkan untuk mengetahui apa yang sebenarnya terjadi disebut data primer. Data sekunder adalah data yang dibuat dan dipublikasikan atau tidak. Variabel yang digunakan dalam contoh ini adalah informasi yang diperoleh dari dataset UCI.

Tabel 1 Atribut Dataset yang di pakai

No	Atribut	Nilai
1.	Age	Numeric
2.	Job	Categorical
3.	Education	Categorical
4.	Default Credit	Categorical
5.	Housing Loan	Categorical
6.	Personal Loan	Categorical
7.	Contact Communication	Categorical
8.	Contact mount in years	Categorical
9.	Day of week work	Categorical
10.	Duration	Numeric
11.	Campaign	Numeric
12.	Pday	Numeric
13.	Previous	Numeric
14.	Poutcome	Categorical
15.	Employment variation rate	Numeric
16.	Customers Price Index	Numeric
17.	Customers Confidence Index	Numeric
18.	Euribor 3 Mount rate	Numeric
19.	Number of employes	Numeric
20.	Has the client subcrable a term deposit	binary

Sumber: dataset UCI

3.2. Pengolahan Awal Data

Beberapa metode digunakan untuk mendapatkan data berkualitas tinggi (Yoga Religia et al., 2021) sebagai berikut:

1. Validasi data untuk menemukan dan menghilangkan data outlier/berisik, data yang tidak konsisten dan hilang (missing value). Catatan pasien yang tersedia ditinjau untuk menentukan apakah berisi data yang tidak biasa (outlier/noise), data yang tidak konsisten, atau nilai yang hilang.
2. Integrasi dan transformasi data untuk meningkatkan akurasi dan efisiensi algoritma. Penelitian ini menggunakan data kategorikal. Perangkat lunak yang disebut Rapidminer mengubah data (Hozairi et al., 2021).
3. Reduksi dan analisis data untuk membuat kumpulan data dengan atribut yang lebih sedikit dan tetap bermanfaat

Pada penelitian sebelumnya, penelitian ini mengklasifikasikan keputusan klien menggunakan data mining untuk menilai efektivitas telemarketing bank (Subarkah et al., 2020). Classification and Regression Trees (CART) dan algoritma Naive Baye adalah proses yang digunakan. Algoritma CART lebih unggul dalam memperkirakan keputusan klien telemarketing ketika penawaran deposit digunakan untuk mengatasi masalah ini, menurut studi ini (General et al., 2020). Teknik yang akan diterapkan saat menguji data ditentukan oleh penulis. Pendekatan yang dipilih, sebagaimana ditentukan oleh studi sebelumnya. Metode Algoritma Regresi Logistik digunakan oleh penulis.

1.3 Metode yang Diusulkan

Metode yang digunakan penulis adalah Logistic Regression yang tujuannya untuk meningkatkan akurasi saat pengujian dengan metode T-test (Abdillah et al., 2020).

3.4 Eksperimen dan Pengujian Metode

Dalam percobaan yang dilakukan oleh para peneliti, framework RapidMiner 6.4 digunakan untuk mengolah data agar mendapatkan nilai akurasi yang akurat. Anda dapat melihat nilai akurasi dan plot AUC di browser cepat ini dan dilengkapi dengan uji-T (Farid et al., 2022).

3.5 Evaluasi dan Validasi Hasil

1. Cross Validation

Metode mengevaluasi dan menguji keakuratan model yang dibuat dengan kumpulan data tertentu disebut validasi silang. Menurut (Angraini et al., 2019), pemodelan biasanya mencoba memprediksi atau mengklasifikasikan data baru yang mungkin tidak pernah muncul dalam dataset. Data yang digunakan untuk memvalidasi model disebut data uji atau dataset, sedangkan data yang digunakan dalam proses pembuatan model disebut data pelatihan. Tes tipikal lainnya untuk memperkirakan tingkat kesalahan adalah validasi silang. Data dalam setiap kategori dibagi secara acak menjadi beberapa kelompok dengan ukuran yang sama. Semua prosedur pelatihan dan pengujian dilakukan beberapa kali dengan banyak sampel untuk mengurangi bias spesifik sampel. Tingkat kesalahan semua pengulangan dijumlahkan untuk menentukan tingkat kesalahan total (Butsianto dan Mayangwulan, 2020).

2. Confusion Matrix

Teknik yang biasa digunakan untuk menghitung keakuratan ide penambahan data adalah matriks kebingungan. Ingat, presisi, akurasi dan tingkat kesalahan adalah empat hasil perhitungan menggunakan rumus ini. Tes untuk menentukan benar dan salah item menjadi dasar evaluasi model klasifikasi (Hendrian, 2018). Urutan pengujian diformat sebagai matriks kebingungan, dengan kelas yang diprediksi ditampilkan di bagian atas matriks dan di sebelah kelas yang diamati di sebelah kiri. Angka pada setiap sel menunjukkan seberapa banyak kelas yang diamati diprediksi oleh nilai sebenarnya (Nugroho et al., 2022).

3. ROC Curve

Salah satu metode pengujian model klasifikasi yang dibuat menggunakan kurva ROC. Saat menganalisis properti yang diinginkan dari model pengklasifikasi, kurva ROC digunakan untuk menunjukkan parameter model yang sesuai. Tingkat TP dan tingkat FP diwakili oleh sumbu Y dan sumbu X dari kurva ROC, yang dibagi menjadi dua dimensi (Anisya, 2020). AUC atau Area Under the ROC Curve, didefinisikan sebagai probabilitas, adalah metode yang digunakan untuk menghitung area di bawah kurva ROC. AUC mengukur kinerja diskriminasi dengan memperkirakan probabilitas hasil untuk sampel yang dipilih secara berbeda dari populasi, semakin tinggi AUC, semakin baik klasifikasinya, semakin kuat klasifikasi yang digunakan. Nilai AUC selalu antara 0,0 dan 1,0. (Dewi et al., 2022.)

Algoritma klasifikasi dapat dievaluasi berdasarkan kriteria seperti presisi, kecepatan, ketergantungan, skalabilitas, dan interpretabilitas (Hendrian, 2018). Tingkat akurasi dapat dinilai setelah pemrosesan data untuk mengukur

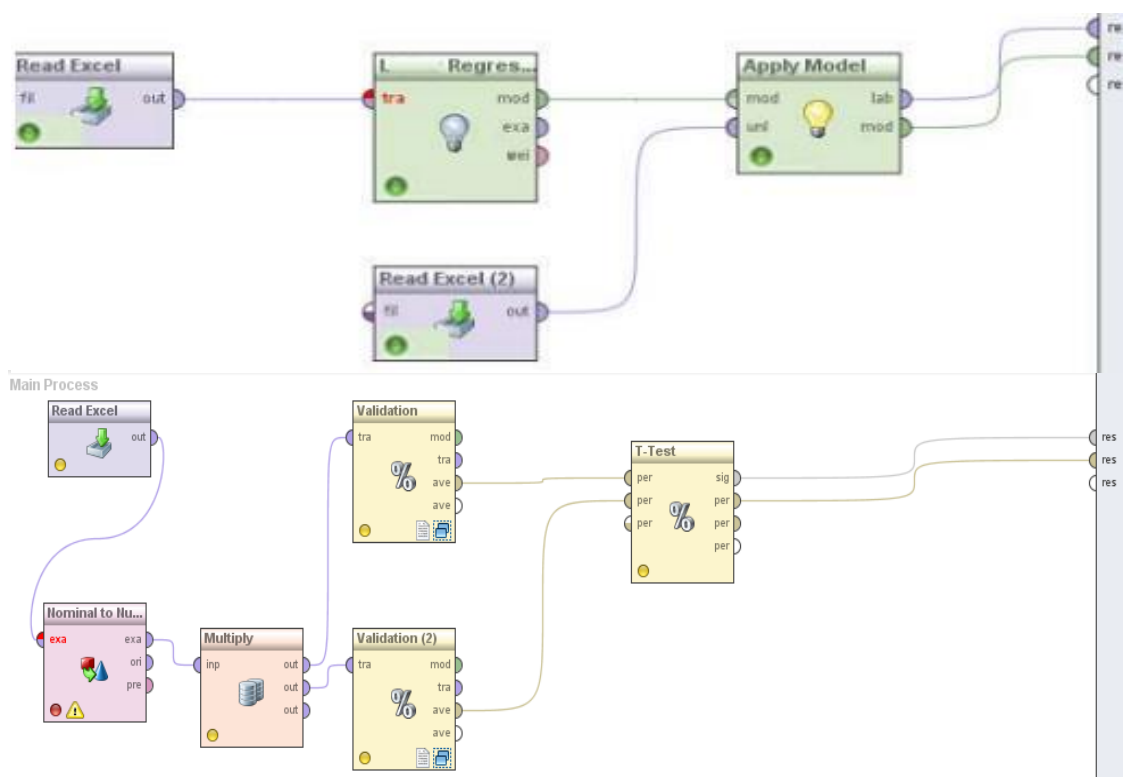
seberapa baik kinerja setiap strategi. Penelitian ini berusaha untuk memastikan apakah kriteria yang dimiliki oleh calon nasabah menguntungkan atau tidak menguntungkan dalam hal prediktabilitas tingkat keberhasilan penjualan produk dan layanan perbankan melalui telemarketing. (Ariadi, 2020).

4. Validasi

Pendekatan sistematis diperlukan untuk mengevaluasi efektivitas strategi. Penilaian kelas didasarkan pada pengujian benar dan salah item (Riany dan Testiana, 2023). Menggunakan data pelatihan untuk melatih kurikulum untuk memaksimalkan pemanfaatan pengetahuan, validasi data digunakan untuk menentukan jenis model pembelajaran yang tepat untuk digunakan (Liliana et al., 2021).

HASIL DAN PEMBAHASAN

Hasil pengujian model yang diuji mencari nilai algoritma regresi logistik untuk mendapatkan akurasi nilai yang diperoleh dan meningkatkan akurasi dengan algoritma T-test software RapidMiner dengan desain model sebagai berikut.:



Sumber : Penelitian (2023)

Gambar 2. Diagram pada Pengujian Validation dan T-test pada Rapidminer

Data pelatihan dimasukkan ke dalam modul Read Excel pada metode pengujian tersebut di atas, yang terkait dengan modul validasi yang berisi metode yang akan digunakan. Metode divalidasi menggunakan k-folds menggunakan alat RapidMiner setelah menghitung data training dan mendapatkan hasilnya. Kami dapat mengamati nilai AUC dan hasil validasi selama proses berlangsung (Rizaldi & Mustakim, 2020). Oleh karena itu, jika kita menghitung nilai AUC dan akurasi pendekatan, Regresi Logistik adalah pilihan terbaik karena akurasi yang tinggi dibandingkan dengan K-NN dan nilai AUC, tetapi jika kita melihat kinerja akurasi yang dihasilkan perlu diperhatikan. diperiksa lagi, khusus menggunakan t-test. (Annisa, 2019).

.Berikut ini adalah panduan keakuratan klasifikasi dengan menggunakan AUC

a. Hasil AUC Logistic Resresion



Sumber : Penelitian (2023)

Gambar 3 Grafik AUC Logistic Regression

Kurva ROC yang diperoleh dengan menguji data pada gambar di atas menunjukkan bahwa akurasi ditingkatkan dengan regresi logistik dengan 92,32 % serta AUC 0,962.

A. Pengujian dengan T-Test

T-test ini mengevaluasi metode klasifikasi ini untuk mencapai nilai terbaik, dan nilai 0,05 dianggap sebagai hasil tes terbaik. (Ratnawati dan Sunendiari, 2021)

Hasil uji-t algoritma pohon keputusan dengan regresi logistik

Tabel 2. Grafik T-Test Logistic Regression

T-Test Significance	0,893 +- 0,008	0,893 +- 0,004
0,893 +- 0,008		0,962
0,893 +- 0,004		

Sumber : Penelitian (2023)

B. Analisis Hasil Komparasi

Hasil yang diperoleh berdasarkan dari analisi pengujian metode LR maka dapat diperoleh hasilnya seperti Tabel berikut

Tabel 3. Performace Algoritma LR

Performace	LR
Accuracy	92,32%
AUC	0.962
T-Test	0.893

Sumber : Penelitian (2023)

Tabel di atas menunjukkan hasil perhitungan yang diperoleh dengan menerapkan klasifikasi kinerja akurasi AUC yaitu terdapat dua metode yang termasuk dalam kategori klasifikasi Baik (Tampill et al., 2017) yaitu untuk LR. metode dengan nilai AUC sebesar 0,962, juga dapat dilihat dari hasil pada tabel diatas bahwa nilai akurasi metode Logistic Regression Classification Algorithm dengan nilai akurasi sebesar 92,32%.

KESIMPULAN

Penelitian ini menggunakan data pasar perbankan untuk membangun model menggunakan algoritma klasifikasi regresi logistik. Model yang dihasilkan dibandingkan untuk melihat algoritma mana yang berkinerja terbaik dalam mengidentifikasi tanggapan positif dari konsumen potensial. Validasi silang dan metode uji ROC digunakan untuk mengevaluasi kinerja algoritma klasifikasi, dan akurasi pengukuran matriks konfusi Algoritma Regresi Logistik adalah 92,32%. Alhasil, masalah identifikasi tingkat keberhasilan telemarketing pemasaran jasa perbankan dapat diselesaikan dengan menggunakan metode regresi logistik. Memprediksi keberhasilan telemarketing bank Bank menggunakan algoritma data mining untuk memprediksi efektivitas pemasaran bank, yang dapat digunakan untuk mengumpulkan data dari calon nasabah. Keakuratan teknik LR ditunjukkan dengan dataset yang digunakan dalam penelitian ini. Hal ini terlihat dari hasil evaluasi. Berdasarkan pengamatan tersebut, dapat disimpulkan bahwa regresi logistik merupakan metode yang dapat diterima untuk peramalan data dan dapat memberikan solusi untuk masalah identifikasi pelanggan potensial. Disarankan untuk penelitian selanjutnya untuk memilih atribut dengan membandingkan metode klasifikasi data mining dengan bank data tambahan, karena atribut material ada pada metode algoritma klasifikasi lainnya. Untuk menentukan apakah suatu fungsi diperlukan atau tidak, metode algoritma tidak berpengaruh (karena nilainya sama). Investigasi tersebut dapat dilakukan di unit bisnis lain atau unit bisnis sejenis. Algoritma lain dapat digunakan untuk mengembangkan penelitian ini, seperti teknik statistik lainnya seperti support vector machine atau neural network.

REFERENSI

- Abdillah, N., Defit, S., & Sumijan. (2020). Penggunaan metode klasifikasi Data Metode Naïve Bayes merupakan metode klasifikasi Data Mining dengan membandingkan algoritma Decision Tree dan Naive Bayes . Penelitian ini bertujuan. *Seminar Nasional Syedza Saintika*, 281–287.
- Anggraini, R. A., Widagdo, G., Budi, A. S., & Qomaruddin, M. (2019). Penerapan Data Mining Classification untuk Data Blogger Menggunakan Metode Naïve Bayes. *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)*, 7(1), 47. <https://doi.org/10.26418/justin.v7i1.30211>
- Anisya. (2020). DATA MINING DALAM PREDIKSI PASOKAN KELAPA SAWIT. *Rang Teknik Journal*, 3(1), 7–13. http://repository.radenintan.ac.id/11375/1/PERPUS_PUSAT.pdf%0Ahttp://business-law.binus.ac.id/2015/10/08/pariwisata-syariah/%0Ahttps://www.ptonline.com/articles/how-to-get-better-mfi-results%0Ahttps://journal.uir.ac.id/index.php/kiat/article/view/8839
- Annisa, R. (2019). Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Penderita Penyakit Jantung. *Jurnal Teknik Informatika Kaputama (JTJK)*, 3(1), 22–28. <https://jurnal.kaputama.ac.id/index.php/JTJK/article/view/141/156>
- Ardi Ramdani, Christian Dwi Sofyan, Fauzi Ramdani, Muhamad Fauzi Arya Tama, & Muhammad Angga Rachmatsyah. (2022). Algoritma Klasifikasi Data Mining Untuk Memprediksi Masyarakat Dalam Menerima Bantuan Sosial. *Jurnal Ilmiah Sistem Informasi*, 1(2), 39–47. <https://doi.org/10.51903/juisi.v1i2.363>
- Ariadi, F. (2020). Analisa Perbandingan Algoritma DT C.45 dan Naïve Bayes Dalam Prediksi Penerimaan Kredit Motor. *KERNEL: Jurnal Riset Inovasi Bidang Informatika Dan Pendidikan Informatika*, 1(1), 1–8. <https://doi.org/10.31284/j.kernel.2020.v1i1.1183>
- Butsianto, S., & Mayangwulan, N. T. (2020). Penerapan Data Mining Untuk Prediksi Penjualan Mobil Menggunakan Metode K-Means Clustering. *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, 3(3), 187–201. <https://doi.org/10.32672/jnkti.v3i3.2428>
- Dewi, S. P., Nurwati, N., & Rahayu, E. (2022). Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor. *Building of Informatics, Technology and Science (BITS)*, 3(4), 639–648. <https://doi.org/10.47065/bits.v3i4.1408>
- Farid, M., Wibowo, S., Puspitasari, N. F., & Satya, B. (2022). PENERAPAN DATA MINING DAN ALGORITMA NAÏVE BAYES UNTUK METODE KLASIFIKASI Abstraksi Pendahuluan Tinjauan Pustaka Metode Penelitian. *Journal of Information System Management (JOISM)*, 3(2).
- Hendrian, S. (2018). Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan. *Faktor Exacta*, 11(3), 266–274. <https://doi.org/10.30998/faktorexacta.v11i3.277>
- Hozairi, H., Anwari, A., & Alim, S. (2021). Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes. *Network Engineering Research Operation*, 6(2), 133. <https://doi.org/10.21107/nero.v6i2.237>

- Liliana, D. Y., Maulana, H., & Setiawan, A. (2021). Data Mining untuk Prediksi Status Pasien Covid-19 dengan Pengklasifikasi Naïve Bayes. *Multinetics*, 7(1), 48–53. <https://doi.org/10.32722/multinetics.v7i1.3786>
- Nugroho, B. I., Ma'arif, Z., & Arif, Z. (2022). Tinjauan Pustaka Sistematis: Penerapan Data Mining Metode Klasifikasi Untuk Menganalisa Penyalahgunaan Sosial Media. *Jurnal Sistem Informasi Dan ...*, 3(2), 46–51. <http://journal.peradaban.ac.id/index.php/jsitp/article/download/1265/860>
- Putro, M. F., Prayitno, E., Siregar, J., & Muharrom, M. (2021). Penerapan Data Mining dengan Naive Bayes untuk Klasifikasi Siswa Sekolah Menengah Dalam Penentuan Perguruan Tinggi. *Jurnal Akrab Juara*, 6, 307–309.
- Ratnawati, S., & Sunendiari, S. (2021). Penggunaan Metode Logistic Regression Ensemble (LORENS) pada Klasifikasi Leukemia Akut. *Prosiding Statika*, 7, 56–63. <http://dx.doi.org/10.29313/v7i1.25555>
- Riany, A. F., & Testiana, G. (2023). Penerapan Data Mining untuk Klasifikasi Penyakit Stroke Menggunakan Algoritma Naïve Bayes. *Jurnal SAINTEKOM*, 13(1), 42–54. <https://doi.org/10.33020/saintekom.v13i1.352>
- Tampil, Y., Komaliq, H., & Langi, Y. (2017). Analisis Regresi Logistik Untuk Menentukan Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa FMIPA Universitas Sam Ratulangi Manado. *D'CARTESIAN*, 6(2), 56. <https://doi.org/10.35799/dc.6.2.2017.17023>
- Wardhani, R. P. S., Sudarno, S., & Maruddani, D. A. I. (2019). OPTIMASI REGRESI LOGISTIK MENGGUNAKAN ALGORITMA GENETIKA UNTUK PEMODELAN FAKTOR-FAKTOR YANG MEMPENGARUHI PENGGOLONGAN KREDIT BANK (Studi Kasus: Debitur di PT BPR Gunung Lawu Klaten Periode Tahun 2017). *Jurnal Gaussian*, 8(4), 506–517. <https://doi.org/10.14710/j.gauss.v8i4.26751>
- Widaningsih, S. (2019). Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm. *Jurnal Tekno Insentif*, 13(1), 16–25. <https://doi.org/10.36787/jti.v13i1.78>
- Yoga Religia, Agung Nugroho, & Wahyu Hadikristanto. (2021). Klasifikasi Analisis Perbandingan Algoritma Optimasi pada Random Forest untuk Klasifikasi Data Bank Marketing. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 187–192. <https://doi.org/10.29207/resti.v5i1.2813>
- Yolanda, I., & Fahmi, H. (2021). Penerapan Data Mining Untuk Prediksi Penjualan Produk Roti Terlaris Pada PT . Nippon Indosari Corpindo Tbk Menggunakan Metode K-Nearest Neighbor. *JIKOMSI*, 3(3), 9–15.