JITE (Journal of Informatics and Telecommunication Engineering)

Available online <http://ojs.uma.ac.id/index.php/jite> DOI : 10.31289/jite.v5i2.6220

Received: 30 November 2021

Accepted: 05 January 2022

Published: 26 January 2022

Implementation of Logistic Regression Classification Algorithm and Support Vector Machine for Credit Eligibility Prediction

Amrin¹⁾, Omar Pahlevi²⁾ *

1) Program Studi Teknologi Komputer, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Indonesia

2) Program Studi Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Indonesia

*Corresponding Email: omar.opi@bsi.ac.id

Abstrak

Kredit merupakan penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, pemberian pinjaman atau kredit. Analisa kredit yang baik sangat diperlukan dikarenakan merupakan salah satu proses yang sangat penting, berupa penyelidikan mengenai lancar atau kurang lancarnya dalam pengembalian kredit. Tahapan mengidentifikasi dan memprediksi nasabah secara baik dan benar dapat dilakukan pada saat sebelum proses pemberian pinjaman. Hal ini dilakukan dengan cara memeriksa data historis pinjaman nasabah tersebut. Kegiatan ini merupakan upaya yang dilakukan oleh industri perbankan pada masa ini dalam menghadapi masalah resiko kredit. Pada penelitian ini, peneliti akan mengaplikasikan dan menerapkan beberapa metode klasifikasi data mining, diantaranya algoritma Logistic Regression dan Support Vector Machine untuk prediksi kelayakan kredit. Dataset yang digunakan sebanyak 481 record data kredit kendaraan bermotor baik yang bermasalah maupun yang tidak bermasalah. Variabel input pada penelitian ini terdiri atas tiga belas variabel, diantaranya Status Perkawinan, Jumlah Tanggungan, Umur, Status Tempat Tinggal, Kepemilikan Rumah, Pekerjaan, Status Pekerjaan, Status Perusahaan, Penghasilan, Uang Muka, Pendidikan, Lama Tinggal, dan Kondisi Rumah. Hasil penelitian dan pengujian yang diperoleh menunjukkan bahwa performa model Logistic Regression untuk prediksi kelayakan kredit memberikan tingkat akurasi kebenaran sebesar 94,81% dengan nilai area under the curva (AUC) sebesar 0,987. Kemudian performa model Support Vector Machine memberikan tingkat akurasi kebenaran sebesar 94,19% dengan nilai area under the curve (AUC) sebesar 0,978. Berdasarkan hasil T-Test maka metode Logistic Regression memiliki performance atau kinerja yang sama dibandingkan dengan Support Vector Machine.

Kata Kunci: Logistic Regression, Support Vector Machine, Confusion Matrix, ROC Curva.

Abstract

Credit is a provision of money or bills that can be equated with it, the provision of loans or credit. A good credit analysis is very necessary, because it is one of the most important processes in the form of an investigation regarding the smooth or substandard credit repayments. The stages of identifying and predicting customers properly and correctly can be done before the loan process. This is done by examining the historical data of the customer's loan. At this time this activity is an effort made by the banking industry in dealing with credit risk problems. In this research, researchers will apply several data mining classification methods, including Logistic Regression algorithms and Support Vector Machines to predict creditworthiness. The dataset used 481 record motorized vehicle loan data, both problematic and non-problematic. The input variables in this study consisted of thirteen variables, including marital status, number of dependents, age, residence status, home ownership, occupation, employment status, company status, income, down payment, education, length of stay, and housing conditions. From the results of research and testing, the performance of the Logistic Regression model for predicting creditworthiness provided an accuracy rate of 94.81% with an area under the curve (AUC) value of 0.987. While the performance of the Support Vector Machine model provides an accuracy of 94.19% with an area under the curve (AUC) value of 0.978. Based on the T-Test test, the Logistic Regression method has the same performance compared to the Support Vector Machine.

Keywords: Logistic Regression, Support Vector Machine, Confusion Matrix, ROC Curva.

How to Cite : Amrin, & Pahlevi, O. (2022). Implementation of Logistic Regression Classification Algorithm and Support Vector Machine for Credit Eligibility Prediction. *JITE (Journal of Informatics and Telecommunication Engineering)*, 5(2), 433-441

I. PENDAHULUAN

Status perkembangan ekonomi dan tren perkembangan perekonomian nasional merupakan dua aspek yang diutamakan pada pasar keuangan, dikarenakan dapat menjadi faktor perubahan tren pasar keuangan. Kemudian kondisi keadaan operasional yang tidak stabil, memiliki tingkat periodisitas yang jelas, dan mengalami kondisi pada tahapan sirkulasi tingkat tinggi dan rendah memiliki peran pada status perkembangan ekonomi dan tren perkembangan perekonomian nasional. Berdasarkan hal ini, analisa teoritis yang sederhana tidak cukup untuk menyimpulkan resiko sistemik pasar keuangan (Ma & Lv, 2019).

Moda transportasi kendaraan pribadi roda empat pada saat ini sangatlah dibutuhkan dalam melakukan kegiatan khususnya untuk mobilitas, akan tetapi terdapat beberapa faktor yang menyebabkan seseorang terkendala dalam memilikinya antara lain faktor pendapatan seseorang dan harga yang mahal. Berdasarkan faktor-faktor tersebut, pembelian secara kredit menjadi solusi dalam membeli kendaraan roda empat (Amrin, 2017).

Pada tahapan pemberian pinjaman atau kredit memerlukan analisa kredit, dikarenakan merupakan salah satu proses yang sangat penting berupa penyelidikan mengenai lancar atau kurang lancarnya dalam pengembalian kredit (Setiadi, 2017). Pada tahapan identifikasi kadang mengalami kegagalan, sehingga mengakibatkan masalah antara lain hilangnya pendapatan dan resiko kredit macet yang dapat mengancam profitabilitas. Timbulnya permasalahan seperti kehilangan nasabah, ketidakpastian pengembalian pinjaman, bahkan ketidakmampuan nasabah untuk mengembalikan pinjaman merupakan contoh dari kesalahan pada analisa kredit (Sugiyarto et al., 2019).

Tahapan mengidentifikasi dan memprediksi nasabah secara baik dan benar dapat dilakukan pada saat sebelum proses pemberian pinjaman dengan cara memeriksa data historis pinjaman nasabah tersebut. Kegiatan ini merupakan upaya yang dilakukan oleh industri perbankan pada masa ini dalam menghadapi masalah resiko kredit. Pada kegiatan perbankan metode klasifikasi resiko kredit memegang peranan penting, apabila terdapat kesalahan yang terjadi pada proses klasifikasi nasabah yang berhutang dapat menimbulkan adanya kredit yang bermasalah (Bawono, 2019).

Suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam basis data dikenal dengan *data mining*. *Data mining* menghasilkan pengetahuan (*knowledge*) yang tidak diketahui sebelumnya dari suatu sekumpulan data. Kegiatan pengolahan data tersebut di dalam basis data, *data warehouse*, atau media penyimpanan informasi yang lain melalui proses pencarian pola-pola yang tersembunyi (*hidden pattern*) (Parapat & Sinaga, 2018). Dalam beberapa tahun terakhir, ada peningkatan minat penelitian dalam menerapkan *Logistic Regression* yang diharapkan dapat menghasilkan model resiko kredit yang lebih akurat (Fitzpatrick & Mues, 2016). Sedangkan *Support Vector Machines* (SVM) biasanya digunakan karena memiliki hasil yang lebih baik daripada pengklasifikasi lainnya (Teles et al., 2021).

Pada penelitian ini, peneliti akan menerapkan dua buah metode klasifikasi *data mining*, diantaranya yaitu algoritma klasifikasi *Logistic Regression* dan *Support Vector Machine* untuk prediksi kelayakan kredit. Adapun alasan peneliti menerapkan metode klasifikasi *Logistic Regression* dikarenakan merupakan salah satu algoritma yang populer dan sangat berguna dalam pembelajaran mesin untuk masalah klasifikasi. Algoritma ini dapat digunakan untuk mendeskripsikan data dan digunakan untuk menjelaskan hubungan antara variabel biner tunggal dan variabel tingkat nominal tunggal atau ganda, ordinal yang bersifat independen (Sheikh et al., 2020). Kemudian alasan peneliti menerapkan metode *Support Vector Machine* dalam penelitian ini, dikarenakan dalam masalah penilaian kredit (*credit scoring*) sering menggunakan model *Support Vector Machine*. Model algoritma ini didukung dengan adanya kemampuan penyamarataan data yang baik serta mampu memecahkan masalah pada jumlah data tertentu (Hasrul, 2020).

Beberapa penelitian terdahulu yang terkait dengan tema metode-metode yang pernah digunakan untuk menyelesaikan prediksi kelayakan kredit, diantaranya penelitian yang dilakukan oleh (Doswalwar et al., 2021), penelitian ini menggunakan algoritma klasifikasi *Logistic Regression*, *Decision Tree Classifier*, *K Neighbors Classifier*, *Naive Bayes*, *Random Forest Classifier*, *Support Vector Machine*, dan *XGBoost Classifier*. Pada penelitian ini terdapat tiga belas karakteristik atribut antara lain Jenis Kelamin, Pendidikan, Jumlah Tanggungan, Status Perkawinan, Pekerjaan, Skor Kredit, Jumlah Pinjaman, dan lain-lain dengan target apakah mereka memenuhi syarat untuk mendapatkan pinjaman atau tidak. Dari hasil eksperimen yang dilakukan, disimpulkan bahwa *Logistic Regression* memberikan akurasi yang lebih baik untuk prediksi ketersediaan pinjaman dengan nilai akurasi sebesar 0.785. Selanjutnya penelitian yang dilakukan oleh (Alankar & Alam, 2021), penelitian ini menggunakan *Logistic Regression* dan *k-Nearest Neighbor*. Evaluasi pada penelitian ini dilakukan berdasarkan *confusion matrix*, akurasi, dan jumlah nilai *misclassified*. *Confusion matrix* mengenai sampel yang diprediksi dengan benar dengan algoritma *Logistic Regression* adalah 1379

(1329+50) dan sampel yang diprediksi dengan benar oleh algoritma *k-Nearest Neighbor* 1368 (1331+37). Hasil pada penelitian ini algoritma *Logistic Regression* menghasilkan hasil bahwa lebih tepat dibandingkan dengan *k-Nearest Neighbor*. Pada bagian akurasi, dapat ditarik kesimpulan bahwa akurasi *Logistic Regression* sama dengan 0,919 dan *k-nearest neighbor* sama dengan 0,912, yang berarti *Logistic Regression* memberikan model yang lebih akurat. Nilai ini berdasarkan *confusion matrix*, *accuracy score* dan nilai *misclassified* prediksi yang diperoleh dari evaluasi model menggunakan bahasa pemrograman python.

Kemudian penelitian yang dilakukan oleh (Horak et al., 2020), pada penelitian ini menggunakan metode algoritma *Support Vector Machine* dan *Artificial Neural Networks* untuk memprediksi kepailitan pada perusahaan industri yang beroperasi di Republik Ceko selama 5 tahun terakhir. Tujuan dilakukannya penelitian ini untuk mengidentifikasi perusahaan yang diperkirakan akan menghadapi kesulitan keuangan di masa depan. Model SVM menunjukkan kemampuan yang hebat untuk memprediksi situasi kedua yang berlawanan pada pandangan pertama, yaitu kemampuan perusahaan untuk bertahan dari kemungkinan kesulitan keuangan. Dalam hal ini, prediksi model sebesar 99,39% serta kemampuan memprediksi kebangkrutan berada pada level 8,22%. Penelitian terkait berikutnya merupakan penelitian yang dilakukan oleh (Moula et al., 2017) yang menggunakan dataset *benchmarking credit* negara Australia, Jerman, dan Jepang dari repositori UCI *machine learning database*, serta dataset bank komersil China. Tujuan dari penelitian ini adalah untuk mengatasi masalah prediksi *default* kredit dengan *Support Vector Machine* (SVM). Pada tahapan penelitian ini membandingkan antara model SVM dengan model CART dengan *Discriminant Analysis*. Hasil eksperimen pada penelitian ini menunjukkan bahwa model SVM sedikit lebih unggul dari CART dengan *Discriminant Analysis* (DA). Model SVM memiliki nilai akurasi 0.7700 nilai presisi 89.44%, dan nilai AUC 72.99%.

II. STUDI PUSTAKA

A. Kredit

Pengertian kredit menurut UU Perbankan No.10 Tahun 1998 dalam (Nurelasari, 2016) merupakan penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, dimana pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam melunasi hutangnya setelah jangka waktu tertentu dengan pemberian bunga menjadi dasar persetujuannya. Menurut Kasmir (2004) dalam (Wahyono & Cahyono, 2015), kredit berarti memperoleh barang dengan membayar cicilan atau angsuran di kemudian hari atau memperoleh pinjaman uang yang pembayarannya dilakukan di kemudian hari dengan cicilan atau angsuran sesuai dengan perjanjian.

B. Data Mining

Sebuah perangkat lunak yang memiliki fungsi antara lain menemukan pola tersembunyi, tren, maupun aturan-aturan yang terdapat dalam basis berukuran besar dimana dapat menghasilkan aturan-aturan yang digunakan untuk memperkirakan perilaku di masa mendatang (Wahyuningsih & Utari, 2018). Menurut (Han & Kamber, 2006) dalam (Pahlevi, 2018), *data mining* adalah rangkaian proses untuk menggali nilai tambah berupa informasi yang belum terekplorasi dari sebuah basis data, melakukan eksplorasi dengan cara-cara tertentu untuk memanipulasi data menjadi informasi yang lebih berharga dengan cara mengekstraksi dan mengenali pola penting dari basis data.

C. Klasifikasi

Klasifikasi merupakan proses menempatkan obyek atau konsep tertentu, kedalam satu set kategori berdasarkan sifat obyek atau konsep yang bersangkutan. Pengertian klasifikasi *data mining* adalah salah satu metode pembelajaran yang dapat digunakan memprediksi nilai dari sekelompok atribut dalam menggambarkan dan membedakan kelas data atau konsep yang bertujuan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui (Rifai, 2016). Menurut (Gorunescu, 2011) untuk tingkat akurasi klasifikasi data mining dapat dibagi menjadi beberapa kelompok:

1. 0.90-1.00 = klasifikasi sangat baik
2. 0.80-0.90 = klasifikasi baik
3. 0.70-0.80 = klasifikasi cukup
4. 0.60-0.70 = klasifikasi buruk

5. 0.50-0.60 = klasifikasi salah

D. Confusion Matrix

Confusion matrix memberikan keputusan yang diperoleh dalam *training* dan *testing*, *confusion matrix* memberikan penilaian *performance* klasifikasi berdasarkan objek dengan benar atau salah (Gorunescu, 2011). *Confusion matrix* menghasilkan nilai akurasi, presisi, *recall*. Akurasi dalam klasifikasi merupakan presentase ketepatan *record* data yang diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi. *Precision* atau *confidence* merupakan proporsi kasus yang diprediksi positif yang juga positif benar pada data yang sebenarnya. *Recall* atau *sensitivity* merupakan proporsi kasus positif yang sebenarnya yang diprediksi positif secara benar.

E. Logistic Regression

Merupakan salah satu algoritma klasifikasi yang digunakan untuk memberikan penilaian secara diskrit. *Logistic Regression* pada dasarnya adalah algoritma klasifikasi *supervised* (Srinivas et al., 2020). Tujuan utama dari penerapan *logistic regression* adalah klasifikasi secara individual dalam kelompok yang berbeda (Alheety et al., 2021).

F. Support Vector Machine

Merupakan teknik yang digunakan dalam klasifikasi data dan analisis regresi, dimana tujuan utamanya adalah untuk memisahkan beberapa kelas dalam set pelatihan (Cervantes et al., 2020). Menurut (Bellotti & Crook, 2007) dalam (Rahmatullah, 2016), *Support Vector Machine* (SVM) adalah metode *learning machine* yang bekerja atas prinsip *Structural Risk Minimization* (SRM) dengan tujuan menemukan *hyperplane* terbaik yang memisahkan dua buah *class* pada *input space*.

G. Rapid Miner

Menurut Lee dan Juan (2010:101) dalam (Indriani & Tanjung, 2017) merupakan perangkat lunak *open source* untuk *knowledge discovery* dan *data mining* dimana memiliki kurang lebih 400 prosedur (operator) *data mining*, termasuk operator untuk masukan, *output*, *data preprocessing* dan visualisasi. *Rapid Miner* pertama kali dinamai *Yet Another Learning Environment* atau singkat YALE, kemudian pada tahun 2007 akhirnya diganti namanya menjadi *Rapid Miner* (Nofitri & Irawati, 2019).

III. METODE PENELITIAN

Penelitian ini terdiri atas beberapa tahap seperti terlihat pada gambar 1 kerangka pemikiran. Permasalahan pada penelitian ini adalah belum diketahui algoritma yang akurat untuk prediksi kelayakan kredit. Untuk itu dibuat *approach* (model) yaitu algoritma klasifikasi *Logistic Regression* dan *Support Vector Machine* untuk memecahkan permasalahan kemudian dilakukan pengujian terhadap kinerja dari metode tersebut. Pengujian menggunakan metode *Cross Validation*, *Confusion Matrix* dan kurva ROC. Pada penelitian ini menggunakan aplikasi *Rapid Miner*.

Gambar 1. Kerangka Pemikiran Penelitian

IV. HASIL DAN PEMBAHASAN

A. Analisa Data

Dataset yang digunakan sebanyak 481 record data kredit kendaraan bermotor baik yang bermasalah maupun yang tidak bermasalah. Variabel input pada penelitian ini terdiri dari 13 variabel, yaitu: 1) Status Perkawinan, 2) Jumlah Tanggungan, 3) Umur 4) Status Tempat Tinggal, 5) Kepemilikan Rumah, 6) Pekerjaan, 7) Status Pekerjaan, 8) Status Perusahaan, 9) Penghasilan, 10) Uang Muka, 11) Pendidikan, 12) Lama Tinggal, dan 13) Kondisi Rumah. Gambar 2 menunjukkan sampel dataset yang digunakan untuk menguji algoritma yang akan diuji.

status perkawinan	jumlah tanggungan	pendidikan terakhir	usia	kepemilikan rumah	lama tinggal	kondisi rumah
menikah	2-3	SLTP	21-55	ortu	>5	permanen
belum menikah	tdk ada	SLTA	21-55	ortu	1-3	permanen
menikah	2-3	SLTA	21-55	KPR	3-5	permanen
belum menikah	tdk ada	SLTA	<21/>60	ortu	1-3	permanen
menikah	>3	SD	21-55	atas nama	>5	non permanen
menikah	2-3	SLTP	21-55	milik sendiri	>5	permanen
menikah	2-3	SLTA	21-55	milik sendiri	>5	permanen

Gambar 2. Sampel Dataset

B. Pengujian Model

Penelitian ini dilakukan dengan eksperimen pengujian pada model yang diusulkan. Kemudian dilakukan evaluasi dan validasi model untuk menghasilkan nilai *accuracy* dan AUC. Pengujian menggunakan *Rapid Miner* dengan operator *10-fold cross-validation* untuk mendapatkan hasil *accuracy* dan AUC pada setiap algoritma yang diuji. Langkah-langkah pengujian model ditunjukkan oleh gambar 3 berikut ini:

Gambar 3. Pengujian Model yang Dibangun

Evaluasi yang dilakukan adalah dengan *Confusion Matrix* dan *ROC Curve* atau *Area Under Curve* (AUC).

1. *Confusion Matrix*

a. Algoritma *Logistic Regression*

Tabel 1 adalah *confusion matrix* untuk algoritma *Logistic Regression*. Diketahui 175 data diklasifikasi “bad” diprediksi sesuai dengan data sebenarnya, lalu 4 data diprediksi “good” tetapi ternyata “bad”. Kemudian 281 data diklasifikasi “good” diprediksi sesuai, dan 21 data diprediksi “bad” ternyata “good”.

Tabel 1. Model *Confusion Matrix* untuk *Logistic Regression*

Accuracy: 94.81% +/- 2.46% (mikro: 94.80%)			
	true bad	true good	class precision
pred. bad	175	21	89.29%
pred. good	4	281	98.60%
class recall	97.77%	93.05%	

b. Algoritma *Support Vector Machine*

Tabel 2 adalah *confusion matrix* untuk algoritma *support vector machine*. Diketahui 158 data diklasifikasi “bad” tepat sesuai dengan data sebenarnya, lalu 21 data diprediksi “good” tetapi ternyata “bad”. Kemudian 295 data diklasifikasi “good” diprediksi tepat sesuai dengan kenyataan dan 7 data diprediksi “bad” ternyata “good”.

Tabel 2. Model *Confusion Matrix* untuk Algoritma *Support Vector Machine*

Accuracy: 94.19% +/- 3.43% (mikro: 94.18%)			
	true bad	true good	class precision
pred. bad	158	7	95.76%
pred. good	21	295	93.35%
class recall	88.27%	97.68%	

2. Kurva *Receiver Operating Characteristic* (ROC)

a. Algoritma *Logistic Regression*

Kurva ROC untuk algoritma *Logistic Regression* seperti ditunjukkan oleh gambar 4 di bawah ini.

Gambar 4. Kurva ROC Algoritma Logistic Regression

Kurva ROC pada gambar 4 mengekspresikan *confusion matrix*. Garis horizontal adalah *false positives* dan garis vertikal *true positives*.

b. Algoritma Support Vector Machine

Kurva ROC untuk algoritma *Support Vector Machine* seperti ditunjukkan oleh gambar 5 di bawah ini.

Gambar 5. Kurva ROC Algoritma Support Vector Machine

3. Analisis Hasil Komparasi

Perbandingan nilai *accuracy*, *precision*, *recall* dan AUC untuk algoritma *Logistic Regression* dan *Support Vector Machine* ditunjukkan oleh tabel 3 di bawah ini.

Tabel 3. Komparasi Nilai Accuracy dan AUC

Algoritma	Accuracy	Precision	Recall	AUC
<i>Logistic Regression</i>	94.81%	98.65%	93.06%	0.987
<i>Support Vector Machine</i>	94.19%	93.51%	97.69%	0.978

Tabel 3 membandingkan *accuracy* dan AUC dari tiap algoritma. Secara sekilas terlihat bahwa nilai *accuracy* algoritma *Logistic Regression* sedikit lebih tinggi, begitu pula dengan nilai AUC-nya memiliki akurasi yang sedikit lebih tinggi dibandingkan algoritma *Support Vector Machine*. Untuk menjamin klaim atau dugaan tersebut, maka dilakukan evaluasi menggunakan metode yang paling umum dalam metode statistik tradisional, yaitu dengan uji *t-Test*. Dapat dilihat pada Gambar 6 hasil *performance t-test*. Dari hasil *t-test* dapat dilihat bahwa tidak ada perbedaan signifikan antara *Logistic Regression* dengan *Support vector Machine*, karena nilai $\alpha = 0,650 > 0,05$. Hal tersebut menunjukkan bahwa Algoritma *Logistic Regression* memiliki *performance* atau kinerja yang sama dibandingkan dengan Algoritma *Support Vector Machine*.

Gambar 6. Performance t-test

Berdasarkan pengelompokkan klasifikasi maka dapat disimpulkan bahwa model algoritma *Logistic Regression* dan *Support Vector Machine* termasuk katagori klasifikasi sangat baik.

V. SIMPULAN

Dari hasil penelitian dan pengujian bahwa performa model *Logistic Regression* untuk prediksi kelayakan kredit memberikan tingkat akurasi kebenaran sebesar 94,81% dengan nilai *area under the curva* (AUC) sebesar 0,987. Sedangkan Performa model *Support Vector Machine* memberikan tingkat akurasi kebenaran sebesar 94,19% dengan nilai *area under the curve* (AUC) sebesar 0,978. Berdasarkan *t-test* maka metode *Logistic Regression* memiliki *performance* atau kinerja yang sama dibandingkan dengan *Support Vector Machine*.

DAFTAR PUSTAKA

- Alankar, B., & Alam, I. (2021). Predictive Modeling and Analysis of Logistic Regression and k-Nearest Neighbor for Personal Loan Campaign. *ICIDSSD 2020*, 1(1), 1–12. <https://doi.org/10.4108/eai.27-2-2020.2303232>
- Alheety, M. I., Månsson, K., & Golam Kibria, B. M. (2021). A new kind of stochastic restricted biased estimator for logistic regression model. *Journal of Applied Statistics*, 48(9), 1559–1578. <https://doi.org/10.1080/02664763.2020.1769576>
- Amrin, A. (2017). Analisa Kelayakan Pemberian Kredit Mobil Dengan Menggunakan Metode Neural Network Model Radial Basis Function. *Paradigma*, 19(102), 1410–5063. <http://ejournal.bsi.ac.id/ejurnal/index.php/paradigma/article/view/2283>
- Bawono, B. (2019). Perbandingan Metode Regresi Logistik Biner dan Naive Bayes Dalam Klasifikasi Debitur Berdasarkan Kualitas Kredit Nasabah. Universitas Muhammadiyah Semarang.
- Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, xxxx. <https://doi.org/10.1016/j.neucom.2019.10.118>
- Dosalwar, S., Kinkar, K., Sannat, R., & Pise, D. N. (2021). Analysis of Loan Availability using Machine Learning Techniques. *International Journal of Advanced Research in Science, Communication and Technology*, September, 15–20. <https://doi.org/10.48175/ijarsct-1895>
- Fitzpatrick, T., & Mues, C. (2016). An empirical comparison of classification algorithms for mortgage default prediction: Evidence from a distressed mortgage market. *European Journal of Operational Research*, 249(2), 427–439. <https://doi.org/10.1016/j.ejor.2015.09.014>
- Gorunescu, F. (2011). *Data Mining: Concepts, Models, and Techniques* (Verlag Berlin Heidelberg (ed.)). Springer.
- H, M. H. (2020). Credit Scoring Menggunakan Support Vector Machine Berbasis Evolution Strategies. *JURNAL IT Media Informasi IT STMIK Handayani*, 11(2), 72–77.

- Horak, J., Vrbka, J., & Suler, P. (2020). Support Vector Machine Methods and Artificial Neural Networks Used for the Development of Bankruptcy Prediction Models and their Comparison. *Journal of Risk and Financial Management*, 13(3), 60. <https://doi.org/10.3390/jrfm13030060>
- Indriani, K., & Tanjung, Q. (2017). Sistem Pendukung Keputusan Kelayakan Kredit Motor Menggunakan Metode NAÏVE BAYES Pada NSC FINANCE Cikampek. *Publikasi Jurnal Penelitian Teknik Informatika Universitas Prima Indonesia*, 1((UNPRI) Medan), 6–11.
- Ma, X., & Lv, S. (2019). Financial credit risk prediction in internet finance driven by machine learning. *Neural Computing and Applications*, 31(12), 8359–8367. <https://doi.org/10.1007/s00521-018-3963-6>
- Moula, F. E., Guotai, C., & Abedin, M. Z. (2017). Credit default prediction modeling: An application of support vector machine. *Risk Management*, 19(2), 158–187. <https://doi.org/10.1057/s41283-017-0016-x>
- Nofitri, R., & Irawati, N. (2019). Integrasi Metode Neive Bayes Dan Software Rapidminer Dalam Analisis Hasil Usaha Perusahaan Dagang. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 6(1), 35–42. <https://doi.org/10.33330/jurteksiv6i1.393>
- Nurelasari, E. (2016). Penentuan Nilai Kredit Dengan Algoritma Klasifikasi Support Vector Machine Berbasis Particle Swarm Optimization. *Paradigma*, XVIII(1), 13–20.
- Pahlevi, O. (2018). Data Mining Penentuan Aturan Asosiasi Penjualan Makanan di Amaria Hotel Jakarta Menggunakan Algoritma Apriori. *Jurnal Sistem Informasi STMIK Antar Bangsa*, 2, 137–142.
- Parapat, J. S., & Sinaga, A. S. (2018). Data Mining Algoritma C4.5 Pada Klasifikasi Kredit Koperasi Simpan Pinjam. *Jurnal Ilmu Teknik Elektro Komputer Dan Informatika (JITEKI)*, 4(2), 144–154.
- Rahmatullah, S. (2016). Komparasi Algoritma C4.5 dan SVM Berbasis Particle Swarm Optimazation dalam Penentuan Kredit. *Paradigma*, XVIII(1), 79–87.
- Rifai, A. (2016). Kajian Algoritma C4.5, Naive Bayes, Neural Network dan SVM dalam Penentuan Kelayakan Pemberian Kredit. *Sistem Informasi STMIK Antar Bangsa*, 5(2), 176–182.
- Setiadi, A. (2017). Penerapan Algoritma Radial Basis Functions untuk Prediksi Kelayakan Pemberian Kredit. *Konferensi Nasional Ilmu Sosial & Teknologi (KNiST)*, 607–612.
- Sheikh, M. A., Goel, A. K., & Kumar, T. (2020). An Approach for Prediction of Loan Approval using Machine Learning Algorithm. *Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020, Icesc*, 490–494. <https://doi.org/10.1109/ICESC48915.2020.9155614>
- Srinivas, K., Madhukar Rao, G., Vengatesan, K., Shivkumar Tanesh, P., Kumar, A., & Yuvaraj, S. (2020). An implementation of subsidy prediction system using machine learning logistical regression algorithm. *Advances in Mathematics: Scientific Journal*, 9(6), 3407–3415. <https://doi.org/10.37418/amsj.9.6.21>
- Sugiyarto, I., Sudarsono, B., & Faddillah, U. (2019). Performance Comparison of Data Mining Algorithm to Predict Approval of Credit Card. *Sinkron*, 4(1), 149. <https://doi.org/10.33395/sinkron.v4i1.10181>
- Teles, G., Rodrigues, J. J. P. C., Rabêlo, R. A. L., & Kozlov, S. A. (2021). Comparative study of support vector machines and random forests machine learning algorithms on credit operation. *Software - Practice and Experience*, 51(12), 2492–2500. <https://doi.org/10.1002/spe.2842>
- Wahyono, T., & Cahyono, A. D. (2015). Pengembangan Model Mitigasi Resiko Kredit Berbasis Komputasional Untuk Meningkatkan Kemampuan Manajemen Resiko Bagi Koperasi. *Jurnal Sistem Komputer*, 5(1), 2087–4685.
- Wahyuningsih, S., & Utari, D. R. (2018). Perbandingan Metode K-Nearest Neighbor, Naive Bayes dan Decision Tree untuk Prediksi Kelayakan Pemberian Kredit. *Konferensi Nasional Sistem Informasi 2018 STMIK Atma Luhur Pangkalpinang*, 1(1), 619–623.