

**CREDIT SCORING MENGGUNAKAN ALGORITMA
CLASSIFICATION AND REGRESSION TREE (CART)
DAN *ARTIFICIAL BEE COLONY***



OLEH:
INDRA IRAWAN
09042681721009

**PROGRAM STUDI MAGISTER ILMU KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2021**

LEMBAR PENGESAHAN

*CREDIT SCORING MENGGUNAKAN ALGORITMA CLASSIFICATION
AND REGRESSION TREE (CART) DAN ARTIFICIAL BEE COLONY*

TESIS


Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Magister

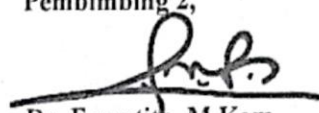
OLEH :

INDRA IRAWAN, S.Kom.
09042681721009

Palembang, Juli 2021
Pembimbing 2,

Pembimbing 1,


Dian Palupi Rini, M.Kom., Ph.D.
NIP. 197802232006042002


Dr. Ermatira, M.Kom.
NIP. 19670913200602000

Mengetahui,
Kaprodi Magister Ilmu Komputer


Dian Palupi Rini, M.Kom., Ph.D.
NIP. 197802232006042002



HALAMAN PERSETUJUAN

Pada hari Sabtu tanggal 9 Januari 2021 telah dilaksanakan ujian tesis II secara daring oleh Magister Ilmu Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Indra Irawan
NIM : 09042681721009
Judul : Credit Scoring Menggunakan Algoritma Clasification And Regression Tree (CART) Dan Artificial Bee Colony

1. Pembimbing I

Dian Palupi Rini, M.kom., Ph.D
NIP 197802232006042002



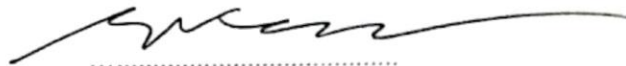
2. Pembimbing II

Dr. Ermatita, M.Kom.
NIP 196709132006042001



3. Penguji I

Dr.Ir.Sukemi, M.T
NIP 196612032006041001



4. Penguji II

Dr. Yusuf Hartono, M.Sc.
NIP 196411161990031002



Palembang, Juli 2021

Mengetahui,

Koordinator Program Studi Magister Ilmu Komputer



Dian Palupi Rini, M. kom., Ph.D
NIP 197802232006042002

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Indra Irawan
NIM : 09042681721009
Program Studi : Magister Ilmu Komputer
Judul Tesis : Credit Scoring Menggunakan Algoritma Classification
And Regression Tree (Cart) Dan Artificial Bee Colony

Hasil Pengecekan Software iThenticate/Turnitin : 19 %

Menyatakan bahwa laporan tesis saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan tesis ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, Juli 2021



Indra Irawan

NIM. 09042681721009

KATA PENGANTAR

Puji dan syukur penulis haturkan kehadiran Allah SWT yang telah melimpahkan rahmat, hidayah, dan kesehatan kepada penulis sehingga dapat menyelesaikan Tesis yang berjudul **“Credit Scoring Menggunakan Algoritma Classification And Regression Tree (Cart) Dan Artificial Bee Colony”**.

Pada kesempatan ini, penulis juga hendak mengucapkan banyak terimakasih kepada pihak-pihak yang telah banyak membantu penyelesaian Tesis ini, diantaranya :

1. Orang tua, Bapak Suroso dan Ibu Sri Lestari
2. Istriku Fitri Indahsari,S.Pd
3. Dosen Pembimbing 1, Ibu Dian Palupi Rini, M.Kom., Ph.D.
4. Dosen Pembimbing 2, Ibu Dr. Ermatita, M.Kom.
5. Pejuang Tesis Pandito Dewa Putra,Andre Hardoni dan Lia Andiani

Penulis menyadari bahwa masih banyak kekurangan dalam penulisan Tesis ini. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi kemajuan karya tulis khususnya yang berkenaan dengan Tesis ini. Penulis berharap semoga Tesis ini dapat bermanfaat bagi semua pihak khususnya Magister Ilmu Komputer Universitas Sriwijaya.

Palembang, 7 Juli 2021

Penulis
Indra Irawan

CREDIT SCORING USING ALGORITHM CLASSIFICATION AND REGRESSION TREE (CART) AND ARTIFICIAL BEE COLONY

Indra Irawan

Abstract

Various statistical techniques and machine learning have been used to develop financial prediction models. In this case, credit rating is closely related in terms of prediction of creditworthiness. Because there is no general agreement on financial ratios as an input feature for model development, many studies consider feature selection as a pre-consideration step in data mining before creating a model. This study examines the effect of feature selection using Artificial Bee Colony on the performance improvement of the CART algorithm. The experimental results show that ABC is the best combination of feature selection in improving CART algorithm performance. Compared with some PSO and Ant Colony optimization algorithms, the proposed method shows outstanding performance. The test results with 2 types of datasets show the level of accuracy of CART + ABC with an accuracy of 80% and 82%. With this research it is expected to be a reference in terms of assessment. credit, allowing banks to reject prospective borrowers with poor creditworthiness.

Keywords: Credit Scoring, CART, ABC, Feature Selection, Data Mining

ABSTRAK

Berbagai teknik statistik dan *machine learning* telah digunakan untuk mengembangkan model prediksi keuangan. Dalam hal ini Penilaian kredit sangat berkaitan erat dalam hal prediksi kelayakan kredit. Karena tidak ada kesepakatan umum atas rasio keuangan sebagai fitur masukan untuk pengembangan model, banyak penelitian menganggap pemilihan fitur sebagai langkah pra-pemrosesan dalam data mining sebelum membuat model. Penelitian ini menguji pengaruh seleksi fitur menggunakan *Artificial Bee Colony* terhadap peningkatan performa algoritma CART. Hasil eksperimental menunjukkan bahwa ABC merupakan kombinasi terbaik dari pemilihan fitur dalam meningkatkan performa algoritma CART. Dibandingkan dengan beberapa algoritma optimasi PSO dan *Ant Colony* metode yang diusulkan menunjukkan kinerja yang luar biasa. Hasil pengujian dengan 2 jenis dataset menunjukkan tingkat akurasi CART + ABC dengan akurasi 80% dan 82%. Dengan adanya penelitian ini diharapkan dapat menjadi sebuah rujukan dalam hal penilaian kredit, memungkinkan bank untuk menolak calon peminjam dengan kelayakan kredit yang buruk.

Kata kunci: *Credit Scoring, CART, ABC, Seleksi Fitur, Data Mining*

DAFTAR ISI

	Halaman
HALAMAN JUDUL	
LEMBAR PENGESAHAN	ii
LEMBAR PERSETUJUAN	iii
LEMBAR PERNYATAAN	iv
KATA PENGANTAR	v
ABSTRACT	vi
ABSTRAK	vii
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA	6
2.1 Landasan Teori	6
2.1.1 <i>Data mining</i>	6
2.1.2 Klasifikasi	6
2.1.3 Seleksi Fitur (<i>Feature Selection</i>)	7
2.1.4 <i>Credit Scoring</i>	7
2.1.5 Algoritma <i>Classification and Regression Tree (CART)</i>	7
2.1.6 Ukuran Pemilihan Atribut	8

2.1.7 <i>Artificial Bee Colony</i>	11
2.1.7.1 <i>Fase Initial</i>	12
2.1.7.2 <i>Improvement Solution</i>	13
2.1.7.3 <i>Swap Operator</i>	13
2.1.7.5 <i>Insert Operator</i>	14
2.1.7.6 <i>Insert Secuence</i>	15
2.1.8 Evaluasi Model	15
2.2 Tinjauan Studi	17
BAB III METODOLOGI PENELITIAN	23
3.1 Pengumpulan Data	23
3.2 Modifikasi <i>Dataset</i>	30
3.3 Instrumen Penelitian	30
3.4 Metode Penelitian	31
3.5 Proses Klasifikasi <i>Algoritma Classification And Regression Tree (CART</i>	33
3.6 Pemilihan Fitur menggunakan <i>Artificial Bee Colony (ABC)</i>	35
3.7 Modifikasi <i>Binary ABC</i>	37
BAB IV HASIL DAN ANALISIS	39
4.1 Hasil Eksperimen Awal Algoritma CART	39
4.1.1 <i>Confution Matrix</i> Algoritma CART	39
4.1.1.1 <i>Confution Matrix Dataset german.data-numeric</i>	39
4.1.1.2 <i>Confution Matrix Dataset germancredit.xls</i>	40
4.1.2 Performa Algoritma CART.	41
4.1.2.1 Performa <i>Dataset German.credit-numeric.</i>	41
4.1.2.2 Kurva ROC <i>Dataset German.credit-numeric</i>	43
4.1.2.3 Performa <i>Dataset Germancredit.xls</i>	43
4.1.2.4 Kurva ROC <i>Dataset Germancredit.xls</i>	44
4.2 Hasil Penelitian algoritma CART + ABC	44
4.2.1. Penelitian <i>dataset german.data-numeric</i>	44
4.2.2. Penelitian <i>dataset germancredit.xls</i>	46

4.2.3	<i>Confusion Matrix</i> algoritma CART + ABC	48
4.2.3.1	<i>Confusion Matrix CART + ABC german.data-numeric</i>	48
4.2.3.2	<i>Confusion Matrix CART + ABC Dataset germandata.xls</i>	49
4.2.4	Performa algoritma CART + ABC	51
4.2.4.1	Performa <i>Dataset german-data-numeric</i>	51
4.2.4.2	Kurva ROC CART+ABC <i>Dataset german-data-numeric</i>	52
4.2.4.3	Performa <i>Dataset Germancredit.xls</i>	52
4. 3	Pembahasan Hasil Penelitian	53
BAB V KESIMPULAN DAN SARAN		57
5.1	Kesimpulan	57
5.2	Saran	57
DAFTAR PUSTAKA		58

DAFTAR TABEL

	Halaman
Tabel 2.1 Kriteria Pengukuran Evaluasi	16
Tabel 2.2 <i>Confusion Matrix</i>	17
Tabel 3.1 Atribut Dataset <i>German .data</i>	24
Tabel 3.2 Atribut Dataset <i>Germancredit.xls</i>	24
Tabel 3.3 Ilustrasi <i>Employed Bee ABC</i>	38
Tabel 4.1 <i>Confusion Matrix Dataset “German.data-numeric”</i>	39
Tabel 4.2 <i>Confusion Matrix Dataset “Germancredit.xls”</i>	40
Tabel 4.3 Performa Algoritma CART (<i>German.data-numeric</i>)	42
Tabel 4.5 Performa Algoritma CART (<i>Germancredit.xls</i>)	43
Tabel 4.6 Hasil Penelitian CART + ABC	44
Tabel 4.7 Hasil Penelitian CART + ABC	46
Tabel 4.8 <i>Confusion Matrix Dataset “German.credit-numeric”</i>	49
Tabel 4.9 <i>Confusion Matrix Dataset “Germandata.xls”</i>	50
Tabel 4.10 Performa Algoritma CART + ABC	51
Tabel 4.11 Performa Algoritma CART + ABC (<i>Germancredit.xls</i>)	52
Tabel 4.12 Perbandingan Hasil Penelitian	53
Tabel 4.13 Perbandingan Akurasi Metode	55

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Diagram Penelitian	31
Gambar 3.2 Diagram Alir Proses Klasifikasi Algoritma <i>Classification And Regression Tree</i>	33
Gambar 3.3 <i>Swap Operator</i>	36
Gambar 3.4 <i>Insert Operator</i>	36
Gambar 3.5 Diagram Alir Seleksi <i>Fitur Artificial Bee Colony</i>	37
Gambar 4.1 Kurva ROC CART <i>Dataset German.data-numeric</i>	43
Gambar 4.2 Kurva ROC CART <i>Dataset Germancredit.xls</i>	44
Gambar 4.3 Kurva ROC CART <i>Dataset German.data-numeric</i>	52
Gambar 4.4 Kurva ROC CART <i>Dataset Germancredit.xls</i>	53
Gambar 4.5 Grafik Perbandingan Tingkat Akurasi	56

DAFTAR LAMPIRAN

LAMPIRAN 1. Source Code Python

LAMPIRAN 2. Publikasi Ilmiah

LAMPIRAN 3. Hasil Pengecekan Software Inthenticate/Turniting

LAMPIRAN 4. Form Konsultasi

LAMPIRAN 5. Form Revisi Ujian Tesis II

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Credit scoring adalah proses penilaian kredit yang sering dilakukan oleh pihak lembaga keuangan. Melalui proses ini, ditentukan apakah calon debitur yang mengajukan kredit diklasifikasikan sebagai calon debitur yang layak untuk diberikan pinjaman atau sebaliknya. Kesalahan dalam proses *credit scoring*, pada akhirnya akan mengakibatkan kerugian dari lembaga keuangan tersebut. Kesalahan proses yang umum terjadi adalah kesalahan hasil dari prosedur *credit scoring* tersebut. Kesalahan yang dinamakan *error* tipe I dan tipe II (Li & Zhong, 2012). Kesalahan tipe pertama yang umum terjadi adalah calon debitur yang layak, dikategorikan sebagai calon debitur yang tidak layak. Kesalahan ini akan merugikan lembaga keuangan tersebut, dimana akan mengurangi keuntungan lembaga keuangan tersebut. Kesalahan tipe kedua, yang mana calon debitur bermasalah dan tidak layak, setelah melalui prosedur *credit scoring*, menjadi calon debitur yang layak dan diberikan pinjaman kredit. Ini berdampak besar bagi *cash flow* perusahaan, dimana apabila dibiarkan maka jumlah kredit macet akan meningkat. Hasilnya adalah lembaga keuangan akan menjadi lembaga keuangan dengan kategori tidak sehat.

Pada akhirnya *history* dari debitur, baik dalam bentuk data pembayaran tagihan dan data lainnya bisa membuktikan apakah prosedur *credit scoring* menghasilkan hasil yang benar atau tidak. Sumber pengetahuan tersebut, apabila bisa dimanfaatkan secara benar, akan sangat membantu di dalam prosedur *credit scoring* calon-calon debitur lainnya di kemudian hari. Ini tentunya akan sangat membantu manajemen lembaga keuangan sebagai dasar pertimbangan untuk membuat keputusan.

Bidang ilmu dari teknologi informasi, yang bisa membantu *credit scoring* adalah data *mining*. Salah satu algoritma yang bisa digunakan di dalam data

mining adalah *Classification And Regresion Tree* (CART). Penggunaan algoritma ini untuk *credit scoring* akan menghemat waktu, usaha dan biaya serta dengan cepat, tepat dan efektif menganalisis kelayakan calon debitur. Algoritma CART sudah digunakan dalam berbagai bidang data mining, seperti web mining, educational mining, medical data mining, sampai dengan credit scoring. Hasil penelitian sebelumnya, menggunakan private dataset dari *Export Development Bank of Iran* dalam jangka waktu dua tahun. Implementasinya pada algoritma klasifikasi menunjukkan bahwa algoritma CART (*Classification and Regression Tree*) sebagai single classification algorithm dengan nilai akurasi 87,6 % (Koutanaei et al., 2015). Tentunya semakin baik tingkat akurasi, maka akan semakin baik juga output diagnosa yang dihasilkan. Hasil penelitian sebelumnya algoritma CART dan Particle Swarm Optimization diterapkan pada credit scoring hasilnya adalah mencapai tingkat akurasi 78 % (Malik & Hermawan, 2018).

Algoritma CART adalah salah satu dari sepuluh algoritma terbaik untuk digunakan di dalam *data mining* (Wu et al., s2008). Ini dikarenakan kelebihan dari algoritma ini yang tergolong metode nonparametrik, sehingga tidak memerlukan asumsi awal, atau spesifikasi. Kelebihan lainnya adalah algoritma ini yang bisa secara otomatis mengidentifikasi variable yang signifikan dan membuang variable yang tidak signifikan. Yang terpenting adalah, kelebihan algoritma ini yang bisa mengatasi data noise. Data noise ini biasanya akan sangat sering terjadi pada data financial (Roman Timofeev, 2004).

Metode CART juga memiliki kekurangan, salah satu kekurangannya adalah, apabila struktur data yang digunakan kompleks atau memiliki banyak atribut/feature, maka ada kemungkinan CART tidak bisa mengidentifikasi secara baik struktur data tersebut. Semua atribut data tidak selalu memberikan kontribusi yang sama pada hasil akhir. Atribut-atribut utama harus diidentifikasi agar dapat berkontribusi dan digunakan sebagai perwakilan data (Pratiwi et al., 2014). Atribut yang tidak relevan dapat mengurangi tingkat akurasi pada proses pengklasifikasian. Untuk mengatasi hal ini, diperlukan pemilihan fitur (*Feature Selection*) supaya menghilangkan atribut yang tidak relevan untuk meningkatkan kinerja metode yang digunakan pada proses klasifikasi (Shen & Yu, 2013). Pemilihan fitur membantu

mempercepat proses klasifikasi dengan mengekstraksi informasi yang relevan dan berguna dari dataset.

Seleksi atribut sendiri, adalah suatu proses yang membutuhkan *effort* yang besar. Akan tetapi dengan perkembangan teknologi informasi, ada banyak sekali algoritma yang bisa digunakan sebagai dasar untuk seleksi atribut. Salah satu algoritma optimasi yaitu Algoritma ABC merupakan algoritma yang handal dalam memberikan solusi optimal, dalam ruang solusi yang besar. Sebagai algoritma optimasi, algoritma ini sudah pernah digunakan untuk meningkatkan akurasi dari algoritma klasifikasi di berbagai bidang data mining, Seperti di dalam medical mining, pada proses *feature selection* algoritma ABC dikomparasikan dengan algoritma Ant Bee Colony untuk melihat tingkat akurasi dengan 10 jenis data set jenis penyakit. Hasilnya adalah peningkatan akurasi ABC lebih baik dibandingkan dengan Ant Colony Optimization (Schiezaro & Pedrini, 2013). Peningkatan hasil akurasi juga ditunjukkan pada penelitian dengan algoritma ABC yang di hybrid dengan algoritma DE, Metode hybrid yang diterapkan untuk masalah pencarian dan optimisasi lainnya kinerjanya untuk pemilihan fitur lebih baik daripada optimasi *Artificial Bee Colony* murni. Penelitian menggunakan 15 dataset dari repositori UCI (Zorarpaci & Özel, 2016).

Oleh karena itu, berdasarkan penelitian sebelumnya, peneliti akan mencoba untuk menerapkan algoritma CART pada credit scoring. Kemudian, peneliti akan mencoba meningkatkan tingkat akurasi tersebut, dengan proses seleksi atribut/feature dengan menggunakan algoritma *Artificial Bee Colony* dengan menggunakan public dataset. Perbandingan akan dibuat, untuk mengetahui berapa besar kenaikan persentase akurasi menggunakan algoritma CART, sebelum dan sesudah menggunakan seleksi atribut oleh ABC.

1.2 Perumusan Masalah

Berlandaskan latar belakang masalah yang telah dijelaskan, permasalahan dapat diuraikan sebagai berikut :

1. Bagaimana cara menentukan credit scoring dengan menggunakan metode CART dan ABC?
2. Bagaimana seleksi fitur dapat meningkatkan akurasi dari CART setelah menggunakan algoritma *Artificial Bee Colony* ?

3. Bagaimana hasil akurasi dari CART dan CART yang dioptimisasi dengan *Artificial Bee Colony* pada data *credit scoring*.

1.3 Batasan Masalah

Penelitian dibatasi pada klasifikasi *credit scoring*, dengan seleksi atribut menggunakan algoritma *Artificial Bee Colony*. Dataset yang digunakan adalah *dataset public*. Sehingga nantinya hasil penelitian bisa digunakan oleh peneliti lain. Dataset yang digunakan adalah dataset standar dari UCI (*University of California*), yaitu *German.data* dan *Germancredit.xls* dari Massachusetts Institute of Technology.

1.4 Tujuan Penelitian

Adapun tujuan yang ingin dicapai pada penelitian ini adalah :

1. Menerapkan algoritma CART dan ABC untuk prosedur *credit scoring*.
2. Menggunakan algoritma optimasi *Artificial Bee Colony* (ABC), untuk seleksi atribut / *feature* dengan tujuan meningkatkan akurasi algoritma CART.
3. Membandingkan akurasi dari *credit scoring* menggunakan CART, dan *credit scoring* menggunakan CART dengan seleksi atribut berbasis *Artificial Bee Colony*(ABC).

1.5 Manfaat Penelitian

Sedangkan, manfaat yang dapat diambil dari penelitian ini adalah :

1. Memberikan kontribusi penelitian di dalam bidang *credit scoring* menggunakan *data mining*.
2. Memberikan alternatif bagi lembaga keuangan untuk melakukan prosedur *credit scoring* dengan menggunakan teknik data mining.

1.6 Sistematika Penulisan

Tahapan sistematika penulisan tesis ini adalah sebagai berikut :

- 1. BAB I Pendahuluan**
Bagian ini menjelaskan tentang mengenai latar belakang masalah, perumusan dan batasan masalah, tujuan dan manfaat penelitian, dan sistematika penulisan tesis.
- 2. BAB II Tinjauan Pustaka**
Bagian ini menjelaskan tentang landasan teori, tinjauan studi yang terkait dengan penelitian yang akan dibahas dalam tesis.
- 3. BAB III Metodologi Penelitian**
Bagian ini menjelaskan mengenai tahap-tahap serta kerangka kerja dalam menyelesaikan tesis.
- 4. BAB IV Hasil dan Analisis**
Bagian ini menampilkan hasil dan analisa penelitian, dalam bentuk tabel, grafik dan pengukuran performa metode yang diusulkan.
- 5. BAB V Kesimpulan dan Saran**
Bagian ini berisikan kesimpulan dari hasil penelitian serta saran untuk penelitian selanjutnya.

DAFTAR PUSTAKA

- Agarwal, S. (2014). Data mining: Data mining concepts and techniques. In *Proceedings - 2013 International Conference on Machine Intelligence Research and Advancement, ICMIRA 2013*.
<https://doi.org/10.1109/ICMIRA.2013.45>
- Chen, J., & Xu, L. (2016). A method of improving credit evaluation with support vector machines. *Proceedings - International Conference on Natural Computation, 2016-Janua*, 615–619.
<https://doi.org/10.1109/ICNC.2015.7378060>
- Chen, M. Y., Chen, C. C., & Liu, J. Y. (2013). Credit rating analysis with support vector machines and artificial bee colony algorithm. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 7906 LNAI*, 528–534.
https://doi.org/10.1007/978-3-642-38577-3_54
- Chen, W., Ma, C., & Ma, L. (2009). Mining the customer credit using hybrid support vector machine technique. *Expert Systems with Applications, 36*(4), 7611–7616. <https://doi.org/10.1016/j.eswa.2008.09.054>
- Gordon S. Linoff, & Michael J. A. Berry. (2013). Data Mining Techniques. Third Edition For Marketing, Sales, and Customer Relationship Management. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9).
- Han, J., Kamber, M., & Pei, J. (2012). Introduction. In *Data Mining*.
<https://doi.org/10.1016/b978-0-12-381479-1.00001-0>

- Han, T., Xie, W., Zisserman, A., Jakab, T., Gupta, A. A., Bilen, H., Vedaldi, A., Graves, A., Wayne, G., Danihelka, I., Collier, M., Beel, J., Maninis, K. K., Caelles, S., Chen, Y. Y., Pont-Tuset, J., Leal-Taixe, L., Cremers, D., Van Gool, L., ... Hwang, S. J. (2019). A systematic approach to variable selection methods used in datamining: A practice in epidemiology. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017-Janua(6), 1–16. <http://arxiv.org/abs/1708.01547><http://arxiv.org/abs/1503.02531><http://dx.doi.org/10.1016/j.procs.2015.02.020><http://dx.doi.org/10.1007/s00500-014-1334-5><http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed17&NEWS=N&AN=72274363><https://>
- Hens, A. B., & Tiwari, M. K. (2012). Computational time reduction for credit scoring: An integrated approach based on support vector machine and stratified sampling method. *Expert Systems with Applications*, 39(8), 6774–6781. <https://doi.org/10.1016/j.eswa.2011.12.057>
- Koutanaei, F. N., Sajedi, H., & Khanbabaei, M. (2015). A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for credit scoring. *Journal of Retailing and Consumer Services*, 27, 11–23. <https://doi.org/10.1016/j.jretconser.2015.07.003>
- Lee, T. S., Chiu, C. C., Chou, Y. C., & Lu, C. J. (2006). Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics and Data Analysis*, 50(4), 1113–1130. <https://doi.org/10.1016/j.csda.2004.11.006>
- Li, X.-L., & Zhong, Y. (2012). An Overview of Personal Credit Scoring: Techniques and Future Work. *International Journal of Intelligence Science*, 02(04), 181–189. <https://doi.org/10.4236/ijis.2012.224024>
- Malik, R. F., & Hermawan, H. (2018). Credit Scoring Using Classification and

Regression Tree (CART) Algorithm and Binary Particle Swarm Optimization. *International Journal of Electrical and Computer Engineering (IJECE)*, 8(6), 5425. <https://doi.org/10.11591/ijece.v8i6.pp5425-5431>

Ping, Y., & Yongheng, L. (2011). Neighborhood rough set and SVM based hybrid credit scoring classifier. *Expert Systems with Applications*, 38(9), 11300–11304. <https://doi.org/10.1016/j.eswa.2011.02.179>

Pratiwi, F. E., Pratiwi, F. E., & Zain, I. (2014). Klasifikasi Pengangguran Terbuka Menggunakan CART (Classification and Regression Tree) di Provinsi Sulawesi Utara. *Jurnal Sains Dan Seni ITS*, 3(1), D54–D59. http://www.ejurnal.its.ac.id/index.php/sains_seni/article/view/6129

Schiezaro, M., & Pedrini, H. (2013). Data feature selection based on Artificial Bee Colony algorithm. *EURASIP Journal on Image and Video Processing*, 2013(1), 1–8. <https://doi.org/10.1186/1687-5281-2013-47>

Shen, X., & Yu, L. (2013). CU splitting early termination based on weighted SVM. *Eurasip Journal on Image and Video Processing*, 2013, 1–11. <https://doi.org/10.1186/1687-5281-2013-4>

Steinberg, D. (2015). *CART: Classification and Regression Trees*. January 2009.

Wang, G., Ma, J., Huang, L., & Xu, K. (2012). Two credit scoring models based on dual strategy ensemble trees. *Knowledge-Based Systems*, 26, 61–68. <https://doi.org/10.1016/j.knosys.2011.06.020>

Wu, X., Kumar, V., Ross, Q. J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z. H., Steinbach, M., Hand, D. J., & Steinberg, D. (2008). Top 10 algorithms in data mining. In *Knowledge and Information Systems* (Vol. 14, Issue 1). <https://doi.org/10.1007/s10115-007-0114-2>

- Xia, Y., Liu, C., Li, Y. Y., & Liu, N. (2017). A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. *Expert Systems with Applications*, 78, 225–241. <https://doi.org/10.1016/j.eswa.2017.02.017>
- Zhang, Z., Gao, G., & Shi, Y. (2014). Credit risk evaluation using multi-criteria optimization classifier with kernel, fuzzification and penalty factors. *European Journal of Operational Research*, 237(1), 335–348. <https://doi.org/10.1016/j.ejor.2014.01.044>
- Zorapacı, E., & Özel, S. A. (2016). A hybrid approach of differential evolution and artificial bee colony for feature selection. *Expert Systems with Applications*, 62, 91–103. <https://doi.org/10.1016/j.eswa.2016.06.004>