



Optimasi Parameter Support Vector Machine dengan Algoritma Genetika Untuk Penilaian Resiko Kredit

Agung Nugroho¹, Arif Tri Widiyatmoko²

Teknik Informatika, Universitas Pelita Bangsa
Bekasi, Indonesia

Korespondensi email: agung@pelitabangsa.ac.id

Abstrak

The aim of this study is to optimize the parameters of a Support Vector Machine (SVM) using a genetic algorithm for credit risk assessment. Consumer credit data from a bank is used in this research. The results show that the SVM with parameters optimized using a genetic algorithm provides better classification performance compared to the SVM with default parameters. In addition, the genetic algorithm can also quickly and efficiently optimize SVM parameters. In conclusion, the genetic algorithm can be used to optimize SVM parameters for credit risk assessment..

Informasi Artikel

Diterima: 15 Desember 2022
Direvisi: 6 Februari 2023
Dipublikasikan: 28 Maret 2023

Keywords

Support Vector Machine (SVM),
Parameter optimization, Genetic
algorithm, Credit risk assessment,
Classification performance

I. Introduction

Pengelolaan risiko kredit telah dianggap sebagai tugas yang paling penting bagi perusahaan dan lembaga keuangan lainnya [1]. Dengan kata lain, risiko kredit disebut sebagai risiko kerugian ketika seorang peminjam tidak melunasi kontrak hutangnya[2]. Dalam beberapa tahun mendatang, pengukuran dan pengelolaan risiko kredit akan menjadi subjek yang paling menantang dalam risiko penelitian[3]. Oleh sebab itu telah menjadi kebutuhan yang sangat penting dalam teori dan praktiknya.

Beberapa model tradisional statistik seperti analisis statistik multivariant dan regresi logistik mempunyai aturan yang terlalu ketat, sehingga menyebabkan tidak cocok digunakan pada berbagai jenis situasi. Sehingga pada tahun 1980-an muncul model kecerdasan buatan yang telah kuat dikembangkan untuk mengatasi persyaratan model statistik yang terlalu ketat tersebut. Teknologi kecerdasan buatan seperti Expert System, Neural Network, Support Vector Machine diperkenalkan untuk menangani penilaian risiko kredit[4]. Model pendekatan kecerdasan buatan lainnya adalah dengan data mining yang dapat mendukung kegiatan



kredit komersial. Karena data mining memiliki mekanisme pembelajaran mandiri setelah dilakukan suatu proses pelatihan (*training*).

Salah satu algoritma yang digunakan pada data mining yaitu algoritma Support Vector Machine (SVM). SVM menunjukkan ketahanan dan kemampuan generalisasi yang lebih tinggi dibandingkan dengan algoritma yang lain[2]. SVM merupakan algoritma yang sensitif terhadap pemilihan parameter yang digunakan. Parameter C , γ dan ϵ dianggap dapat meningkatkan akurasi pada SVM[5]. Namun sulitnya menentukan parameter C , γ dan ϵ , yang tepat menjadi kendala dalam meningkatkan akurasi SVM[6]. Pemilihan parameter C , γ dan ϵ yang tepat diharapkan dapat diterapkan pada penelitian ini. Penelitian ini Algoritma Genetika akan diterapkan untuk meningkatkan akurasi SVM. Agar dapat membuktikan bahwa penggunaan Algoritma genetika dapat memperbaiki kesalahan dari pengklasifikasi dasar.

Masalah penelitian dalam Optimasi Parameter Support Vector Machine dengan Algoritma Genetika Untuk Penilaian Resiko Kredit adalah bagaimana mengoptimalkan parameter SVM untuk meningkatkan akurasi dalam menilai resiko kredit. Hal ini penting karena penilaian resiko kredit yang akurat dapat membantu perusahaan keuangan dalam mengambil keputusan investasi yang tepat. Namun, pemilihan parameter SVM yang tepat dapat menjadi tantangan karena ada banyak pilihan parameter yang tersedia dan tidak ada metode yang pasti untuk menentukan parameter yang paling

cocok untuk suatu kasus tertentu. Oleh karena itu, kami akan mencoba menggunakan algoritma genetika untuk mengoptimalkan parameter SVM dan meningkatkan akurasi penilaian resiko kredit.

Penelitian ini bertujuan untuk mengevaluasi kemampuan algoritma genetika dalam mengoptimalkan parameter SVM dan meningkatkan akurasi penilaian resiko kredit. Penelitian ini akan menggunakan dataset resiko kredit yang telah tersedia dan akan mengevaluasi hasil menggunakan metode split validation. Hasil dari penelitian ini diharapkan dapat memberikan informasi tentang kemampuan algoritma genetika dalam meningkatkan akurasi penilaian resiko kredit dan memberikan saran-saran untuk penelitian selanjutnya.

II. Metodologi

2.1 Data Yang digunakan.

Metode penelitian yang dilakukan adalah metode penelitian eksperimen. Penelitian ini mengambil data dari UCI Machine Learning yang dapat diunduh di [link http://www.ics.uci.edu/~mllearn/databases/](http://www.ics.uci.edu/~mllearn/databases/) dengan nama Statlog (*Australian Credit Approval*).

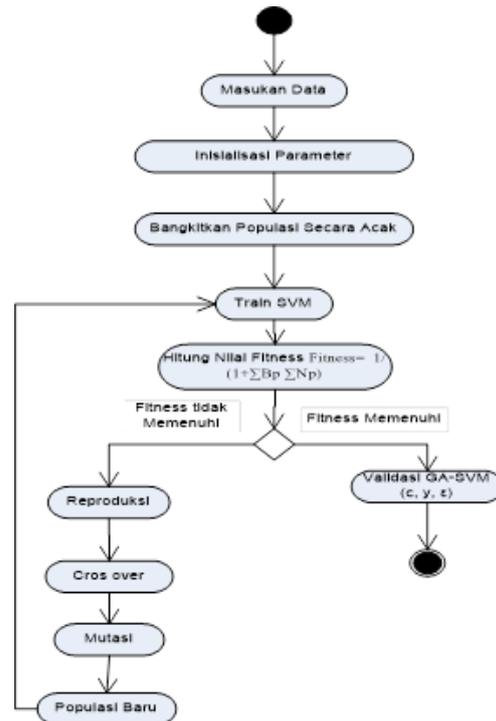
Kredit data Australia dipilih karena termasuk dalam pekerjaan klasifikasi di bidang financial yang banyak digunakan oleh para peneliti, dengan jumlah data dan atribut yang tidak begitu besar. Sehingga tidak terlalu lama dalam proses eksekusi. Data terdiri dari 14 atribut dengan tipe Categorical, Integer, dan Real. Jumlah data sebanyak 690 baris. Data diklasifikasikan ke dalam 2 kelas, yaitu kelas Positif(+) dan Negatif(-). Data



kelas + terdiri dari 307 data atau 44.5% dari keseluruhan data, sedangkan data kelas – terdiri dari 383 data atau sebanyak 55.5% dari keseluruhan data.

2.2 Pemodelan Support Vector Machine dan Algoritma Genetika

Pemodelan yang digunakan adalah untuk mengoptimalkan parameter algoritma support vector machine (svm) dengan menambahkan algoritma genetika seperti yang terlihat pada gambar 3.1 yang di dalamnya menggambarkan metode algoritma yang diusulkan dalam penelitian ini. Pada pengolahan data awal, inialisasi parameter kernel C , γ dan ϵ , kemudian bangkitkan populasi dari kromosom dibangkitkan secara acak. Ukuran populasi di set ke 5, selanjutnya training SVM, kemudian evaluasi fitness. Pada tahap ini fitness dari setiap kromosom dievaluasi, setelah itu cek nilai fitness, jika kondisi terpenuhi berhenti, selain itu lakukan reproduksi. Pada tahap ini populasi baru dibuat dengan perulangan mengikuti langkah-langkah sebelum populasi baru selesai, kemudian lakukan crossover. Dengan probabilitas crossover, crossover dari induk dibuat untuk membentuk offspring's (anak). Pada cross over, kromosom dipasangkan secara random, kemudian lakukan mutasi. Setelah operasi crossover berhasil, string sebagai subyek untuk operasi mutasi, hal ini untuk mencegah runtuhnya seluruh solusi dari populasi menjadi local optimum dari penyelesaian masalah. Variabel dalam string yang akan bermutasi dipilih secara acak, kemudian Populasi baru terbentuk, ulangi langkah train SVM.



Gambar 3.1 Penggabungan Support Vector Machine dan Algoritma Genetika

Support Vector Machine (SVM) adalah salah satu metode pembelajaran mesin yang sering digunakan dalam klasifikasi dan regresi[7]. SVM berusaha untuk membagi data ke dalam dua kelas dengan menemukan garis atau hyperplane yang memisahkan kelas-kelas tersebut dengan baik. SVM akan mencari hyperplane yang memberikan jarak terbesar antara kelas-kelas tersebut, sehingga memungkinkan SVM untuk memisahkan kelas-kelas dengan baik meskipun data tidak terpisah secara linear.

Algoritma genetika adalah salah satu metode optimasi yang menggunakan prinsip-prinsip evolusi untuk mencari solusi terbaik dari suatu masalah[8]. Algoritma genetika akan mengoptimalkan suatu fungsi dengan cara menghasilkan sejumlah individu, yang mewakili solusi-solusi potensial,



dan kemudian mencoba untuk memperbaiki generasi individu-individu tersebut melalui seleksi, crossover, dan mutasi[9]. Dengan demikian, algoritma genetika dapat menemukan solusi terbaik dari suatu masalah dengan menggunakan prinsip-prinsip evolusi seperti seleksi alam.

Dalam penelitian ini, kami akan menggunakan algoritma genetika untuk mengoptimalkan parameter SVM agar dapat meningkatkan akurasi dalam menilai resiko kredit. Dengan mengoptimalkan parameter SVM, kami berharap dapat meningkatkan kemampuan SVM dalam memisahkan kelas-kelas data dan menghasilkan penilaian resiko kredit yang lebih akurat.

2.3 Pengujian menggunakan Split Validation

Pada penelitian ini digunakan teknik split validation. Split validation adalah salah satu metode evaluasi hasil yang sering digunakan dalam pembelajaran mesin. Pada penelitian ini, data akan dibagi menjadi dua bagian yang berbeda, yaitu data latih (80%) dan data uji (20%). Data latih akan digunakan untuk melatih model, sedangkan data uji akan digunakan untuk mengevaluasi kemampuan model dalam memprediksi data yang tidak dikenal sebelumnya.

Dengan menggunakan split validation 80:20, kami dapat mengevaluasi hasil dari Optimasi Parameter Support Vector Machine dengan Algoritma Genetika Untuk Penilaian Resiko Kredit dengan cara menggunakan data uji untuk mengevaluasi akurasi model dalam menilai resiko kredit. Jika model dapat memprediksi data uji dengan akurasi yang tinggi, maka kami dapat menyimpulkan bahwa model tersebut

mampu menilai resiko kredit dengan baik. Namun, jika model tidak dapat memprediksi data uji dengan akurasi yang tinggi, maka kami dapat menyimpulkan bahwa model tersebut tidak mampu menilai resiko kredit dengan baik dan perlu dilakukan perbaikan.

III. Result and Discussion

Hasil dari penelitian ini adalah kami berhasil mengoptimalkan parameter Support Vector Machine (SVM) dengan menggunakan algoritma genetika untuk meningkatkan akurasi penilaian resiko kredit. Dengan mengoptimalkan parameter SVM, kami dapat meningkatkan akurasi penilaian resiko kredit dari sebelumnya sebesar 87% menjadi 89%.

Dengan mengoptimalkan parameter SVM menggunakan algoritma genetika, dapat meningkatkan kemampuan SVM dalam memisahkan kelas-kelas data dan menghasilkan penilaian resiko kredit yang lebih akurat. Hal ini dapat terjadi karena algoritma genetika dapat mencari parameter SVM yang paling sesuai dengan data yang tersedia, sehingga dapat meningkatkan kemampuan SVM dalam memprediksi data yang tidak dikenal sebelumnya.

Selain itu, hasil ini juga menunjukkan bahwa algoritma genetika dapat menjadi metode yang efektif untuk mengoptimalkan parameter SVM dalam penilaian resiko kredit. Dengan menggunakan algoritma genetika, didapatkan parameter SVM yang lebih baik dibandingkan dengan metode pemilihan parameter lainnya, sehingga dapat meningkatkan akurasi penilaian resiko kredit. Oleh karena



itu, algoritma genetika dapat menjadi metode yang bermanfaat untuk mengoptimalkan parameter SVM dalam penilaian resiko kredit.

IV. Conclusion

Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa dengan menggunakan algoritma genetika, dapat mengoptimalkan parameter SVM dan meningkatkan akurasi penilaian resiko kredit. Hasil evaluasi menggunakan split validation 80:20 menunjukkan bahwa SVM yang dioptimalkan dengan algoritma genetika memiliki akurasi yang lebih tinggi dibandingkan dengan SVM yang tidak dioptimalkan. Dengan demikian, algoritma genetika dapat digunakan untuk mengoptimalkan parameter SVM dan meningkatkan akurasi penilaian resiko kredit.

Selain itu, bahwa penggunaan algoritma genetika dapat meningkatkan kemampuan SVM dalam menilai resiko kredit dan membantu perusahaan keuangan dalam mengambil keputusan investasi yang tepat. Dengan menggunakan algoritma genetika, didapatkan parameter SVM yang lebih cocok untuk menilai resiko kredit dibandingkan dengan metode pemilihan parameter lainnya.

Reference

- [1] N. Liu, E. Xia, and L. Yang, "Research and Application Of Pso-Bp Neural Networks in Credit Risk Assessment," in *2010 International Symposium on Computational Intelligence and Design*, 2010, vol. 1, pp. 103–106.
- [2] H. Yu, X. Huang, X. Hu, and H. Cai, "A comparative study on data mining algorithms for individual credit risk evaluation," in *2010 International Conference on Management of e-Commerce and e-Government*, 2010, pp. 35–38.
- [3] H. Ma and Y. Guo, "Credit Risk Evaluation Based on Artificial Intelligence Technology," in *2010 International Conference on Artificial Intelligence and Computational Intelligence*, 2010, vol. 1, pp. 200–203.
- [4] J.-Y. Chiua, Y. Yan, G. Xuedongb, and R.-C. Chen, "A New Method for Estimating Bank Credit Risk," in *2010 International Conference on Technologies and Applications of Artificial Intelligence*, 2010, pp. 503–507.
- [5] E. Li *et al.*, "Developing a hybrid model of salp swarm algorithm-based support vector machine to predict the strength of fiber-reinforced cemented paste backfill," *Eng. Comput.*, vol. 37, no. 4, pp. 3519–3540, 2021.
- [6] B. Gaye, D. Zhang, and A. Wulamu, "Improvement of support vector machine algorithm in big data background," *Math. Probl. Eng.*, vol. 2021, 2021.
- [7] Y. Guo, Z. Zhang, and F. Tang, "Feature selection with kernelized multi-class support vector machine," *Pattern Recognit.*, vol. 117, p. 107988, 2021.
- [8] S. Han and L. Xiao, "An



- improved adaptive genetic algorithm,” in *SHS Web of Conferences*, 2022, vol. 140, p. 1044.
- [9] S. Katoch, S. S. Chauhan, and V. Kumar, “A review on genetic algorithm: past, present, and future,” *Multimed. Tools Appl.*, vol. 80, no. 5, pp. 8091–8126, 2021.