

Implementasi Algoritma Optimasi Spiral pada Metode Klasifikasi *Support Vector Machine* (SVM) (Studi Kasus : Prediksi Diabetes)

Made Adi Widyananda¹, Irma Palupi², Annisa Aditsania³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung
¹adiwid@students.telkomuniversity.ac.id, ²irmapalupi@telkomuniversity.ac.id,
³aaditsania@telkomuniversity.ac.id

Abstrak

Klasifikasi merupakan metode data mining yang diformulasikan untuk membuat perkiraan keanggotaan dalam grup untuk contoh data, proses ini digunakan untuk menganalisis koneksi antara data dalam kumpulan data yang besar. Salah satu metode klasifikasi yang sering digunakan adalah *Support Vector Machine* (SVM), dalam metode SVM terdapat fungsi kernel yang membantu dalam menyelesaikan permasalahan klasifikasi yang tidak dapat dipisahkan secara linear salah satunya yaitu *Radial Basis Function* (RBF) Kernel. Dalam menggunakan metode SVM dengan fungsi kernel RBF, terdapat parameter *Gamma* dan *C* yang nilainya dapat mempengaruhi bentuk hyperplane dalam menghasilkan model klasifikasi yang baik, sehingga dibutuhkan nilai parameter *Gamma* dan *C* yang optimal dalam menghasilkan klasifikasi yang baik. Pada penelitian kali ini menggunakan *Spiral Optimization Algorithm* dalam mengoptimasi parameter *Gamma* dan *C*, dengan melakukan beberapa tahapan percobaan dalam menentukan parameter terbaik dari *Spiral Optimization Algorithm* untuk menentukan parameter *Gamma* dan *C*, metode klasifikasi SVM dengan fungsi kernel RBF dapat menghasilkan akurasi tertinggi yaitu 86,15 % dengan rata-rata akurasi 80,12 % berdasarkan dataset *Pima Indians Diabetes*.

Kata kunci : Klasifikasi, *Support Vector Machine*, *Spiral Optimization Algorithm*, *Pima Indians Diabetes*.

Abstract

Classification is a data mining method that is formulated to estimate group membership for data samples, this process is used to analyze the connections between data in a large data set. One of the classification methods that are often used is Support Vector Machine (SVM), in the SVM method there is a kernel function that helps in solving classification problems that cannot be separated linearly, one of which is the Radial Basis Function (RBF) Kernel. In using the SVM method with the RBF kernel function, Gamma and C parameters can affect the shape of the hyperplane in producing a good classification model, so that optimal Gamma and C parameter values are needed to produce a good classification. This study using the Spiral Optimization Algorithm in optimizing Gamma and C parameters, by conducting several experimental stages in determining the best parameters of the Spiral Optimization Algorithm to determine the Gamma and C parameters, SVM classification method with RBF kernel function can produce the highest accuracy is 86.15% with an average accuracy is 80.12% based on Pima Indians Diabetes dataset.

Keywords: Classification, Support Vector Machine, Spiral Optimization Algorithm, Pima Indians Diabetes.

1. Pendahuluan

Latar Belakang

Data mining, juga disebut *Knowledge Discovery in Databases* (KDD), adalah bidang yang menemukan informasi baru yang berpotensi berguna yang berasal dari data yang besar. Data mining telah diterapkan di banyak bidang, termasuk bidang bioinformatika [1]. Dalam data mining, klasifikasi diformulasikan untuk membuat perkiraan keanggotaan dalam grup untuk contoh data. Proses ini menggunakan analisis data yang kompleks untuk menentukan koneksi data dalam kumpulan data yang sangat besar [2].

Support Vector Machine merupakan metode yang termasuk kedalam *supervised learning* yang menganalisa data dan mengenali pola untuk klasifikasi yang dimana svm bekerja untuk menemukan permukaan keputusan (*Hyperplane*) yang maksimal dari setiap titik data [3]. SVM merupakan metode klasifikasi yang memiliki performa yang sangat baik dan sudah digunakan beberapa kali seperti pada penelitian yang dilakukan oleh Noble W untuk *computational biology* [4], dan Shujun Huang untuk klasifikasi genom kanker [5]. Pada metode SVM, terdapat fungsi kernel yang membantu menyelesaikan permasalahan klasifikasi yang tidak dapat dipisahkan secara linear. Dalam menggunakan metode SVM dengan fungsi kernel, terdapat beberapa parameter yang mempengaruhi