

Model Klasifikasi Berbasis Microarray pada Identifikasi Parkinson dengan Menggunakan Metode Firefly Algorithm-Support Vector Machine

Julizar Wiranto Harahap, Isman Kurniawan,
Fhira Nhita

Fakultas Informatika, Universitas Telkom, Bandung
julizarwh@students.telkomuniversity.ac.id, ismnkrn@telkomuniversity.ac.id,
fhiranhita@telkomuniversity.ac.id

Abstrak

Penyakit Parkinson adalah penyakit yang menyebabkan penderita tidak dapat menggunakan anggota tubuh motoriknya dengan baik dan benar. Sebelum hal tersebut terjadi, maka diperlukan sebuah mekanisme untuk mendeteksi penyakit tersebut agar dapat dilakukan upaya penanganan lebih awal. Salah satu cara untuk mendeteksi penyakit tersebut adalah dengan mengamati ekspresi darah yang berada pada tubuh manusia. Penelitian ini bertujuan untuk mengimplementasikan *Support vector machine* (SVM) dengan *hyperparameter tuning* yang didapatkan dari algoritma *firefly* (FA) untuk mendeteksi seseorang yang berkemungkinan menderita parkinson. Pengamatan yang dilakukan peneliti memiliki hasil bahwa SVM-FA dengan kernel polinomial memiliki performa tertinggi dibanding kernel linear dan kernel *radial basis function* (RBF) saat pengujian. Terjadi peningkatan performa setelah menggunakan *hyperparameter tuning* pada model SVM-FA linear dan model SVM-FA Polinomial. Model terbaik, yakni SVM-FA dengan kernel polinomial yang memiliki *accuracy* sebesar 65,63% serta *f1-score* sebesar 68,57%.

Kata kunci: *support vector machine, svm, firefly-algorithm, hyperparameter tuning, linear, polinomial, radial basis function*

Abstract

Parkinson's disease is a disease that causes sufferers to be unable to use their motor limbs properly and correctly. Before this happens, a mechanism is needed to detect the disease so that early treatment efforts can be made. One way to detect the disease is to observe the expression of blood in the human body. This study aims to implement *Support vector machine* (SVM) with *hyperparameter tuning* obtained from the *firefly* (FA) algorithm to detect someone who is likely to suffer from Parkinson's. Observations made by researchers have the result that SVM-FA with a polynomial kernel has the highest performance compared to the linear kernel and *radial basis function* (RBF) kernel during testing. There is an increase in performance after using *hyperparameter tuning* on the linear SVM-FA model and the SVM-FA Polynomial model. The best model is SVM-FA with a polynomial kernel which has an *accuracy* of 65.63% and an *f1-score* of 68.57%.

Keywords: *support vector machine, svm, firefly-algorithm, hyperparameter tuning, linear, polinomial, radial basis function*

1. Pendahuluan

1.1. Latar Belakang

Parkinson merupakan penyakit yang ditemukan oleh Dr. Jamse Parkinson pada tahun 1817. Penyakit ini berkaitan dengan gangguan motorik maupun non-motorik [1]. Parkinson belum memiliki pengobatan namun dapat dicegah dengan menerapkan terapi Levodopa (L-Dopa) [2]. Levodopa merupakan terapi untuk mencegah penyakit Parkinson, namun dibutuhkan alat yang berfungsi untuk mendeteksi seseorang sebelum divonis mengidap Parkinson. *Support vector machine* (SVM) yang menggunakan *hyperparameter tuning* yang dicari dengan algoritma *firefly* dapat menjadi salah satu alat untuk mendeteksi penderita Parkinson dengan mengamati keterkaitan ekspresi darah pada *dataset microarray* GSE 6613.

Salah satu algoritma yang telah banyak digunakan untuk mendeteksi data *microarray* yaitu SVM. Siyabend Turgut *et al.* pada tahun 2018 melakukan penelitian untuk mendeteksi penyakit kanker payudara dengan menggunakan SVM, *k nearest neighbor*, *multi layer perceptron*, *adaboost*, *gradientboost*, dan *linear regression* [3]. Devi Arocka Vanitha *et al* melakukan penelitian untuk mendeteksi *colon cancer* dan *lymphoma*. Scenario yang diterapkan dalam penelitian Devi Arocka Vanitha *et al* pada SVM yakni dengan menguji beberapa fungsi kernel yang dapat diterapkan pada SVM, yakni fungsi *linear*, fungsi *radial basis function*, fungsi *quadratic*, fungsi *polynomial* [4]. Peneliti Manjula C Belvagi *et al* pernah meneliti terkait evaluasi performa untuk mendeteksi *Intrusion* dengan menggunakan *logistic regression*, *Support vector machine*, *gaussian naive bayes* dan *random forest classifier* dengan *F1 Score support vector machine* sebesar 0.77 [5]. Zayrit Soumaya *et*