

PENERAPAN KLASIFIKASI *RANDOM FOREST* TERHADAP DATA GANGGUAN SPEKTRUM AUTISME (ASD) PADA ANAK-ANAK MENGGUNAKAN SELEKSI FITUR *PRINCIPAL COMPONENT ANALYSIS*

Luthfiyah Amatullah

Abstrak

Autistic Spectrum Disorder (ASD) merupakan gangguan perkembangan fungsi otak yang kompleks dan sangat bervariasi. Kelainan ini secara signifikan berpengaruh terhadap komunikasi verbal, non-verbal serta interaksi sosial. Gejala umum gangguan ini biasanya akan terlihat pada anak-anak mulai sejak usia dua tahun. Tujuan penelitian ini untuk menerapkan metode seleksi fitur *Principal Component Analysis* (PCA) pada data penelitian *Autistic Spectrum Disorder Screening Data for Children Data Set* yang diperoleh dari *University of California Irvine* (UCI) *Machine Learning Data Repository* dimana PCA berfungsi untuk mereduksi atau menyederhanakan dimensi data dari dataset serta mengklasifikasikannya menggunakan pemodelan klasifikasi *Random Forest*. Selain itu, untuk mengetahui bagaimana hasil evaluasi (*confusion matrix*) serta bagaimana perbedaan yang terdapat pada hasil evaluasi tersebut terhadap dataset yang menggunakan metode seleksi fitur *Principal Component Analysis* (PCA) dengan yang tidak menggunakan metode tersebut. Evaluasi dari hasil penelitian terhadap Data Gangguan Spektrum Autisme Pada Anak - anak yang menggunakan seleksi fitur PCA yaitu menghasilkan nilai akurasi sebesar 98%, *precision* sebesar 96%, *recall* sebesar 100% dan *specificity* sebesar 96%. Sedangkan hasil evaluasi yang tanpa melalui proses PCA terlebih dahulu menghasilkan nilai akurasi sebesar 91%, *precision* sebesar 92%, *recall* sebesar 84% dan *specificity* sebesar 100%.

Kata kunci : ASD, Gangguan Spektrum Autisme, Klasifikasi, *Random Forest*, *Principal Component Analysis*.

**APPLICATION OF RANDOM FOREST CLASSIFICATION TO AUTISM
SPECTRUM DISORDER (ASD) DATA IN CHILDREN USING
PRINCIPAL COMPONENT ANALYSIS FEATURE SELECTION**

Luthfiyah Amatullah

Abstract

Autistic Spectrum Disorder (ASD) is a complex and highly variable developmental disorder of brain function. This disorder significantly affects verbal, non-verbal communication and social interaction. General symptoms of this disorder will usually be seen in children starting from the age of two years. The purpose of this research is to apply the Principal Component Analysis (PCA) feature selection method to research data Autistic Spectrum Disorder Screening Data for Children Data Set obtained from the University of California Irvine (UCI) Machine Learning Data Repository where PCA serves to reduce or simplify dimensions data from the dataset and classify it using the Random Forest classification modeling. In addition, to find out how the results of the evaluation (confusion matrix) are and how the differences in the results of the evaluation are to datasets that use the Principal Component Analysis (PCA) feature selection method and those that do not use this method. Evaluation of the results of research on Autism Spectrum Disorder Data in Children using PCA feature selection, which resulted in an accuracy value of 98%, precision of 96%, recall of 100% and specificity of 96%. While the evaluation results without going through the PCA process first resulted in an accuracy value of 91%, precision of 92%, recall of 84% and specificity of 100%.

Keywords : ASD, Autistic Spectrum Disorder, Classification, Random Forest, Principal Component Analysis.