

# KLASIFIKASI STUDI MAHASISWA UNDIKSHA DENGAN *SUPPORT VECTOR MACHINE (SVM)* DAN *BACKPROPAGATION*

Oleh

Dwi Prima Handayani Putri, NIM 1915101012

Jurusan Teknik Informatika

## ABSTRAK

Kelulusan mahasiswa tepat waktu merupakan salah satu tujuan dalam setiap universitas, tidak terkecuali Universitas Pendidikan Ganesha. Permasalahan ketidاكلulusan mahasiswa tepat waktu tidak hanya akan menimbulkan permasalahan bagi mahasiswa tersebut, tetapi juga pada akreditasi kampus. Langkah awal untuk mencegah permasalahan tersebut adalah dengan membuat sebuah model *machine learning* yang dapat mengklasifikasikan apakah seorang mahasiswa terindikasi bermasalah atau tidak. Pada penelitian ini, algoritma *machine learning* yang digunakan adalah (*Support Vector Machine*) SVM dan *Backpropagation*. SVM dipilih karena SVM adalah salah satu metode terbaik untuk memecahkan masalah klasifikasi. Adapun *Backpropagation* dipilih karena algoritma ini dikatakan sangat *powerful* dan hampir selalu “cocok” dengan berbagai permasalahan *machine learning*. Dua buah non-linear kernel telah diaplikasikan pada SVM model, yaitu RBF dan *Polynomial Kernel*. Namun RBF kernel pada penelitian ini memberikan hasil berupa *overfitted model*. Adapun pada *Backpropagation* telah dicoba tiga buah *activation function*, yaitu *sigmoid*, *tanh*, dan *identity activation function*. *Sigmoid* dan *tanh activation function* merupakan non-linear *activation function*, sedangkan *identity activation function* merupakan linear *activation function*. Terdapat tiga buah percobaan untuk proporsi pembagian dataset. Peneliti memilih 90% : 10% sebagai proporsi pembagian data terbaik. Sebanyak 90% data digunakan untuk *training process*, dan 10% sisanya digunakan untuk *testing process*. Penelitian ini menggunakan *K-Fold Cross Validation* untuk membagi data pada *training process*. Adapun model terbaik diambil dari penetapan  $k=4$  pada *K-Fold Cross Validation*. SVM dengan *Polynomial Kernel* memberikan performa model terbaik dengan akurasi sebesar 97%. Adapun *Backpropagation* dengan *Sigmoid Activation Function* sebagai model terbaik juga memberikan hasil akurasi sebesar 97%. Sehubungan dengan perbandingan hasil *Confusion Matrix* serta kurva *AUC*, *Backpropagation* model menunjukkan performa yang sedikit lebih unggul daripada SVM model. Arsitektur serta parameter-parameter pada model harus diperhatikan agar model terbaik dapat dihasilkan.

**Kata Kunci:** *Klasifikasi, SVM, Backpropagation, Kernel, Activation Function*

# **CLASSIFICATION STUDY OF THE UNDIKSHA STUDENTS WITH SUPPORT VECTOR MACHINE (SVM) AND BACKPROPAGATION**

**By**

**Dwi Prima Handayani Putri, NIM 1915101012**

**Informatic Engineering Major**

## **ABSTRACT**

*Student graduation on time is one of the goals of every university, and Ganesha University of Education is no exception. The problem of not graduated on time will not only cause problems for these students, but also for campus accreditation. The first step to prevent this problem is to create a machine-learning model to classify whether a student is indicated as problematic or not. In this study, the machine learning algorithms used are SVM (Support Vector Machine) and Backpropagation. SVM was chosen because SVM is one of the best methods for solving classification problems. Backpropagation was chosen because this algorithm is preeminent and almost always "fits" with various machine learning problems. Non-linear kernels have been applied to the SVM model, namely RBF and Polynomial Kernel. Nonetheless, the RBF kernel gave an overfitted model result. As for Backpropagation, three activation functions have been tried, namely sigmoid, tanh, and identity activation function. Sigmoid and tanh activation functions are non-linear activation functions, while identity activation functions are linear activation functions. There are three experiments for the proportion of dataset distribution. The researcher chose 90% : 10% as the best proportion of dataset distribution. This study uses K-Fold Cross Validation to divide data on the training process. The best model is taken from the determination of  $k = 4$ . SVM with Polynomial Kernel provides the best model performance with an accuracy of 97%. The Backpropagation with the Sigmoid Activation Function as the best model also gives an accuracy of 97%. With the comparison of the results of the Confusion Matrix and the AUC curve, the Backpropagation model shows a slightly superior performance than the SVM model. The architecture and parameters of the model must be considered so that the best model can be produced.*

**Keywords:** *Classification, SVM, Backpropagation, Kernel, Activation Function*