

ABSTRACT

PERFORMANCE ANALYSIS OF SEVERAL KERNEL FUNCTIONS IN SUPPORT VECTOR MACHINE CLASSIFICATION OF DATA ON LIVER DISEASE PATIENTS

BY

CITRA MARIA MAGDALENA NAIBAH

Support Vector Machine (SVM) is one of the machine learning methods used for classification by dividing data into two different classes. The principle of the SVM method is to find the best separating function called hyperplane. If the data cannot be separated linearly then the data is nonlinear data. One method to overcome this is to use a kernel function. The purpose of this research is to apply the SVM method to determine the performance of the best kernel function based on the highest accuracy value for the classification of liver disease patients. Based on the analysis results, the best kernel function for this case is the Radial Basis Function (RBF) kernel function with a cost value = 12 and gamma = 2.5 on a split dataset of 90% training and 10% testing, with an accuracy value of 84.63%. By using the kernel function and the split dataset proportion, the parameters $\mathbf{w}$ and $\mathbf{b}$ are obtained, as follows:

$w_{Age} = -5,1951$, $w_{TB} = -7,3708$, $w_{DB} = -9,1687$, $w_{AP} = -9,2968$, $w_{AA} = -10,6287$,
 $w_{AspA} = -10,3948$, $w_{TP} = 0,1057$, $w_{Albumin} = 9,3402$, $w_{AGR} = 2,5066$, $\mathbf{b} = 0,8868$.

Keywords: Support Vector Machine; Radial Basis Function Kernel;
Classification; Liver Disease

ABSTRAK

ANALISIS KINERJA BEBERAPA FUNGSI *KERNEL* PADA KLASIFIKASI *SUPPORT VECTOR MACHINE* TERHADAP DATA PENDERITA PENYAKIT *LIVER*

OLEH

CITRA MARIA MAGDALENA NAIBAHO

Support Vector Machine (SVM) adalah salah satu metode *machine learning* yang digunakan untuk pengklasifikasian dengan membagi data menjadi dua kelas yang berbeda. Prinsip metode SVM adalah mencari fungsi pemisah disebut *hyperplane* yang terbaik. Apabila data tidak dapat dipisahkan secara linear maka data tersebut merupakan data nonlinear. Salah satu metode untuk mengatasi hal tersebut adalah dengan menggunakan fungsi *kernel*. Tujuan penelitian ini adalah menerapkan metode SVM untuk mengetahui kinerja fungsi *kernel* terbaik berdasarkan nilai akurasi tertinggi terhadap klasifikasi penderita penyakit *liver*. Berdasarkan hasil analisis diperoleh, bahwa fungsi *kernel* terbaik untuk kasus ini adalah fungsi *kernel Radial Basis Function* (RBF) dengan nilai $cost = 12$ dan $gamma = 2,5$ pada *split* dataset 90% *training* dan 10% *testing*, dengan nilai akurasi sebesar 84,63%. Dengan menggunakan fungsi *kernel* dan proporsi *split* dataset tersebut diperoleh parameter $\mathbf{w}$ dan b , yakni sebagai berikut:

$w_{Age} = -5,1951$, $w_{TB} = -7,3708$, $w_{DB} = -9,1687$, $w_{AP} = -9,2968$, $w_{AA} = -10,6287$,
 $w_{AspA} = -10,3948$, $w_{TP} = 0,1057$, $w_{Albumin} = 9,3402$, $w_{AGR} = 2,5066$, $b = 0,8868$.

Kata Kunci : *Support Vector Machine; Kernel Radial Basis Function; Klasifikasi; Penyakit liver*