



UPPSALA  
UNIVERSITET

## Molecular Biotechnology Programme

Uppsala University School of Engineering

<b>UPTEC X 04 029</b>	<b>Date of issue 2004-01</b>	
Author <b>Björn Lindh</b>		
Title (English) <b>Feature Selection with Support Vector Machines in Functional Genomics</b>		
Title (Swedish) Egenskapsurval med supportvektormaskiner inom funktionell genomik		
Abstract There has been a rapid development of new measurement technologies to monitor the expression activity of the genome. The invention of micro-arrays permits simultaneous measurements of a large number of mRNA-molecules. This holds the promise to define diseases in molecular terms and could provide a basis for medical diagnoses. In this context it is crucial to develop computational analysing techniques that can classify and therefore differentiate between samples from normal and disease states. Support vector machines (SVM) is a learning system, which earlier have proved promising characteristics for this purpose. In this study we evaluated support vector machines for feature selection in gene expression data, and tried to give answer to the question why SVM seems to have other characteristics compared to less sophisticated classifiers. Also an alternative capacity evaluation method to "Leave One Out" (LOO) is discussed. The results showed how sparseness of data can affect the learning capacity and that either choice of margin softness or kernel seemed to affect the generalisation property of the system. This gives a hint of how to shape an SVM for classification of gene expression data.		
Keywords Lärande system, SVM, supportvektormaskin, cancerklassificering, microarray, genexpressionsdata		
Supervisor: <b>Jesper Tegnér</b>		
Scientific reviewer: <b>Bengt Persson</b>		
Language: <b>svenska</b>	Security	
<b>ISSN 1401-2138</b>	Classification	
Supplementary bibliographical information	Pages <b>40</b>	
<b>Biology Education Centre</b> Box 592 S-75124 Uppsala	<b>Biomedical Center</b> Tel +46 (0)18 4710000	<b>Husargatan 3 Uppsala</b> Fax +46 (0)18 555217

