**Abstract No: 39**                            ***ICT, Mathematics and Statistics***

# A PERFORMANCE EVALUATION OF VARIOUS SVM TOOLS AND DIFFERENT FEATURE SELECTION STRATEGIES

## T.Kokul[*]

*Postgraduate Institute of Science, University of Peradeniya, Sri Lanka*
[*] *kokul1984@gmail.com*

Support vector machine (SVM) is an efficient classification technique which is widely used in many machine learning applications due to the property of outperforming other classifiers in its generalization performance. There are several tools available for SVM but they differ in their implementation and efficiency. This study aims to evaluate three popular SVM tools: LIBSVM, SVM[light] and MATLAB packaged SVM, and four wrapper based feature selection techniques: Sequential forward selection (SFS), sequential backward selection (SBS), sequential forward floating selection (SFFS) and sequential backward floating selection (SBFS), in classification. The evaluvation was performed on five benchmark numerical datasets: *Segment*, *Vehicle* and *Satimage* from the UCI machine learning repository, and *Madelon* and *Gisette* from the NIPS 2003 feature selection challenge. The former subset of data is of multiclass whereas the latter is of binary class classification tasks. Each dataset was scaled to be in [-1, 1] and were classified using one-against-all (OVA) SVMs by comparing linear and RBF kernels. The performance evaluvation among SVM tools were tested for statistical significance using ANOVA test. Testing results show that LIBSVM and SVM[light] outperform MATLAB packaged SVM in classification. LIBSVM is of near performance to SVM[light]. Moreover, LIBSVM is faster when training data is of dense format. The kernel evaluation of sparse vector is slower in LIBSVM so the total training time is at least twice of that using the dense format. This issue has been well tackled by SVM[light] that is faster in training when using the sparse format. In addition to this, the feature selection techniques were evaluvated using LIBSVM. The testing results show that SFFS technique yields more compact feature sets while maintaining comparable performance to other feature selection techniques in classification. Based on the experimental results, LIBSVM can be considered as a better tool and SFFS technique gives more compact feature set which can be classified using the implementation of LIBSVM.

**Keywords**: Support vector machine, feature selection, LIBSVM, SVM[light].