Chapter 7

Conclusion and Future Work

Most of the real-world applications such as text categorization, optical character recognition, speech recognition, and intrusion detection are essentially multi-class classification problems. Support Vector Machine (SVM) developed by Vapnik based on statistical learning theory was originally designed for binary classification. It is based on the idea of Structural Risk Minimization (SRM) which minimizes the generalization error. Extending SVM for multi-class classification is a challenging problem. Two types of approaches for multi-class SVM has been suggested in literature. One is considering all data in one optimization. However, it requires very large computation time. The other is decomposing multi-class into a series of binary SVMs, such as "One-Against-All" (OAA) and "One-versus-One" (OvO). It has been pointed out in literature that both conventional OvO and OAA SVMs suffer from the problem of unclassifiable region. To resolve unclassifiable region in conventional OvO and conventional OAA, decision tree OvO SVM and decision tree OAA SVM formulation were proposed respectively.

In this thesis, we have designed few multiple classifier systems to improve the performance of multi-class classification problems.

We investigated the role of kernel function and its parameters and trade-off parameter C to achieve better classification accuracy of support vector machine. It is noted from experimental results that the classification accuracy of support vector machine depends on the choice of kernel function and its parameters, and trade-off parameter C. We have proposed few new kernel functions which satisfy the Mercer’s conditions. We have also proposed AdaBoost, one of multiple classifier system, which consider SVM as base classifier. It takes into account the suitable kernel function and its parameters, and trade-off parameter to improve the performance of support vector machine. The experimental results show that with the proper choice of kernel function and its parameters using proposed algorithm, it is possible to achieve better classification accuracy for all datasets. We have also noticed that proposed new kernel function
provides better classification accuracy on few datasets in comparison to other kernel functions.

We also proposed new formulations of OvO ODT-SVM and OAA ODT-SVM using the statistical measures i.e. information gain, gini index, chi-square and scatter-matrix-based class separability in kernel-induced space. We have also shown theoretically that the computation time of training and testing of both the ODT-SVMs using statistical measures is better in comparison to conventional SVMs. The performance of our proposed frameworks is evaluated in terms of classification accuracy and computation time of training and testing phase. The experimental results on UCI repository datasets indicate that classification accuracy of our proposed frameworks are comparable or better than conventional OvO, SVM-BDT, conventional OAA and two widely used multi-class classifiers e.g. C4.5 and MLP for most of the datasets. Our experimental results also demonstrate that the computation time of proposed ODT-SVMs formulation is significantly less in comparison to conventional SVM and SVM-BDT models.

Intrusion detection is a difficult multi-class classification problem. We have investigated the performance of our proposed both ODT-SVM classifiers using all measures in terms of (i) classification accuracy (ii) Training time and (iii) Testing time. We have also investigated the performance of multiple classifier system using two different strategies such as simple voting, weighted majority of combining classifiers to take final decisions. The results of the experiment on KDD’99 intrusion dataset indicate that average classification accuracy of our proposed OvO ODT-SVM and OAA ODT-SVM is better than conventional SVMs and significantly better than widely used multi-class classifiers e.g. C4.5 and MLP. The average classification accuracy improves with multiple classifier systems using simple voting and weighted majority schemes to combine the classifiers (one of OvO ODT-SVM, C4.5, and MLP) to take final decision. Our experimental results also demonstrate that the computation time of proposed ODT-SVMs formulation is significantly less in comparison to conventional SVM and other multi-class classifiers. However, the training and testing phase of multiple classifier systems requires significantly large computation time as we have to build and evaluate individual models in multiple classifier systems.
It will be both interesting and challenging to determine the association between the kernel function and the underlying distribution of data. This will enable us to choose suitable kernel function for a given dataset to achieve maximum class separability between data points of two different classes. There is also need to define some statistical measures to determine better class separability between data points corresponding to two different classes to enhance the performance of ODT-SVMs.