In this paper, we propose a general robust probability classifier (RPC) based on $\varphi$-divergences to address classification problems with data uncertainty. Traditional classifiers make an exact distributional assumption about the class-conditional densities, which may be unavailable due to unavoidable observational noises and other uncertain factors. To address these issues, we propose a class-conditional probability distributional set based on $\varphi$-divergences to describe the data uncertainty. The optimal RPC is defined as the one with the minimal worst-case loss function value over all possible distributions in the proposed distributional set. For RPC models under an absolute error loss criterion (AE-RPC), we give general equivalent reformulations and show that the corresponding problems can be solved as polynomial-time-solvable second order cone programming or linear programming problems. For RPC models under a squared error loss criterion (SE-RPC), we define a modified distributional set based on the modified $\chi^2$-distance and show that it can be solved as an equivalent semi-definite programming problem.

Some experiments on binary and multiple classification problems validate that the proposed models are robust to the data uncertainty and can avoid the “over-learning” phenomenon.

Experimental results show that the proposed RPC models outperform RSVM for both binary and multiple classification problems on the tested data sets. For example, Figure 1 shows the performance of AE-RPC, LE-RPC and RSVM on Y5 data set when different training rates are selected. In general, SE-RPC provides the highest classification accuracy on the test sets and AE-RPC provides the best classification accuracy on the training sets among these models. Both AE-RPC and SE-RPC have the robustness to the data uncertainty and can avoid the “over-learning” phenomenon.

Figure 1: Performance of AE-RPC, SE-RPC and RSVM on Y5 data set.

References: