The Support Vector Machine (SVM) is a popular binary classification model due to its superior generalization performance, relative ease-of-use, and applicability of kernel methods. SVM training entails solving an associated quadratic programming (QP) that presents significant challenges in terms of speed and memory constraints for very large datasets. Slow training times are especially of concern when one considers that re-training is often necessary at several values of the model's regularization parameter, $C$, as well as associated kernel parameters.

The active set method is an important method for SVM training, especially since it supports incremental training, and has gained recent significance as a method for solving the entire regularization path. The regularization path-following algorithms have been shown to find the SVM solution at all values of $C$ in not much more time than it takes, on average, to perform training at a single value of $C$ with traditional methods. Unfortunately, the majority of active set implementations used for SVM training only allow positive definite kernels, and those implementations that do allow semi-definite kernels tend to be complex and can exhibit instability and, worse, lack of convergence. This severely limits applicability since it precludes the use of the linear kernel, can be an issue when duplicate data points exist, and doesn't allow use of low-rank kernel approximations to improve tractability for large datasets.

This research introduces practical active set implementations for SVM training with semi-definite kernels. Methods are given for both conventional SVM training as well as for computing the regularization path that are simple and numerically stable. The proposed regularization path-following algorithm is shown to be orders of magnitude faster, more accurate, and significantly less complex than competing methods and does not require the use of external solvers. A method for computing the approximate regularization path and approximate kernel path with a conventional active set method using warm-starts is also given and shown to provide significant, orders of magnitude, speed-ups relative to the traditional “grid search” where re-training is performed at each parameter value. Theoretical analysis reveals new insights into the nature of the path-following algorithms.