RANDOM PROJECTIONS FOR SUPPORT VECTOR MACHINES

By

Saurabh Paul

A Thesis Submitted to the Graduate
Faculty of Rensselaer Polytechnic Institute
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
Major Subject: COMPUTER SCIENCE

Approved:

_____
Petros Drineas, Thesis Adviser

_____
Malik Magdon-Ismail, Thesis Adviser

_____
Mohammed J. Zaki, Thesis Adviser

Rensselaer Polytechnic Institute
Troy, New York

November 2012
(For Graduation December 2012)

ABSTRACT

Let $\mathbf{X} \in \mathbb{R}^{n \times d}$ be a data matrix of rank ρ, representing n points in $\mathbb{R}^d$. The linear support vector machine constructs a hyperplane separator that maximizes the 1-norm soft margin. We develop a new *oblivious* dimension reduction technique which is precomputed and can be applied to any input matrix $\mathbf{X}$. We prove that, with high probability, the margin and minimum enclosing ball in the feature space are preserved to within ϵ-*relative error*, ensuring comparable generalization as in the original space. We present extensive experiments on synthetic and real-world datasets in support of the theory.