

Research on the Soft Margin Support Vector Machine method and its application in classifying noisy linear data

Nguyen Thi Loan

Abstract: This study investigates the application of the Soft Margin Support Vector Machine (SVM) method for classifying linearly separable data in the presence of noise. While the standard hard-margin SVM is effective for perfectly separable datasets, its performance degrades significantly when outliers or overlapping class distributions are introduced. This research addresses this limitation by implementing a soft-margin approach that introduces a slack variable and a regularization parameter (C) to allow for a controlled degree of misclassification, thereby enhancing model robustness. The methodology involves formulating the primal optimization problem and deriving its dual representation using Lagrange multipliers. Experimental results on synthetic noisy linear datasets demonstrate that the soft-margin SVM effectively balances the trade-off between maximizing the margin and minimizing classification error. The findings confirm that this method provides superior generalization and noise tolerance compared to the hard-margin variant, making it a practical solution for real-world classification tasks where data imperfections are inevitable.

Keywords: Support vector machine, soft margin, data classification, noisy data