111062118 江佩霖

Machine Learning
Lab3 Report

I. State the possible reason why the accuracy or F1-score change between Perceptron and LDA?

The Perceptron is a linear classifier that does not make specific distribution assumptions, making it sensitive to outliers, which can adversely affect its performance. In contrast, Linear Discriminant Analysis (LDA) assumes that the data follows a normal distribution with equal covariance among classes, which can enhance its accuracy when these assumptions hold true. Additionally, the Perceptron's iterative training process can lead to variability in results, while LDA provides a stable, closed-form solution that is less impacted by such fluctuations. Furthermore, the Perceptron is sensitive to feature scaling, which can significantly influence its performance, whereas LDA demonstrates greater resilience to variations in feature scaling.

II. Does MAP help? Why?

MAP estimation integrates prior distributions into the learning process, which is particularly beneficial for making predictions with small or imbalanced datasets, as it allows the model to leverage prior knowledge. In the context of the Perceptron, MAP serves as a regularization technique, effectively controlling overfitting by penalizing large weights and promoting a simpler model.

For LDA, MAP enhances robustness by refining class means and covariances, resulting in more reliable classifications, especially in challenging scenarios with overlapping classes.

III. Summarize how you solve the difficulty and your reflections

The Perceptron faced slow convergence in noisy data, requiring careful tuning of learning rates and iterations. Both models struggled with imbalanced datasets, prompting the use of class weighting or oversampling. Feature selection and preprocessing, like scaling, were crucial for performance.

I found the most challenging aspect to be addressing class imbalance, as it significantly impacted the models' performance metrics. Additionally, tuning the Perceptron for better convergence required a deeper understanding of its learning dynamics.