IST707 Applied Machine Learning

HW9: SVMs, kNN, and Random Forest for handwriting recognition

In this homework, you will use SVMs, kNN, Random Forest, and Gradient Boosting algorithms for handwriting recognition, and compare their performance with the naïve Bayes and decision tree models you built in previous week.

Steps:

1. Describe data pre-processing steps and the chosen evaluation method and measure(s)

2. Use the train set to build kNN, SVM, Random Forest, and Gradient Boosting models. Report test performance, compare them, and use your theoretical knowledge to explain whether the algorithm performance difference makes sense or not.

3. Write a report to describe what you did, including the data preparation, transformation, algorithm tuning, the generated models and their performance. In the end, summarize which model works the best and why.

4. Submit your report as a Microsoft Word document {.doc, .docx, etc.}

Grading rubrics:

1. Are the models constructed correctly?
2. Is the result analysis conclusion convincing?
3. Are sufficient details provided for others to repeat the analysis?
4. Does the analysis include irrelevant content?