

Aprendizaod de Máquina

Lista 3 - Entrega 05/10

Marcos Prates

Lista em Duplas. Se você usou R ou outro programa orientado a programação para resolver os exercícios práticos, submeta junto com as respostas o código utilizado para resolvê-las (**os envios devem ser feito pelo Moodle**). Se você usou algum programa visual, explique qual programa foi usado, e como o problema foi resolvido nele (que funcionalidades foram usadas etc).

Exercício 1. Seu objetivo é criar classificadores para prever o dígito correspondente a uma imagem. O conjunto de dados está disponível em <https://www.kaggle.com/c/digit-recognizer>. Você deve usar o arquivo train.csv para criar os seus classificadores (incluindo validação), e deve fornecer ao site as previsões encontradas para o conjunto test.csv. Note que os dígitos correspondentes no conjunto teste não estão indicados! O site irá rankear seu grupo de acordo com as previsões fornecidas. Como parte do exercício, você deverá:

(a) [0.25 pts] Inscrever seu time no Kaggle. Qual o nome dele?

(b) [1.0 pts] Plotar 5 imagens deste banco.

Você também implementar os seguintes classificadores, assim como estimar seus riscos via conjunto de teste (usando o conjunto de validação para estimar seus erros), mostrar o resultado de cada um:

(c) [2.25 pts] Bagging

(d) [2.25 pts] Árvores de Classificação

(e) [2.0 pts] Boosting

(f) [2.25 pts] SVM

Você também *poderá* implementar outros classificadores que achar razoável. Escolha um dos classificadores ajustados (o que achar melhor) e submeta suas respostas ao site, assim vamos ver qual dupla fica melhor rankeada na sala (a dupla com melhor posição ganhará 3 pontos extra).