

Lista de Exercícios #1

Mineração de Dados - QS 2020

Instruções

1. A lista deve ser feita individualmente.
2. Desconsiderando as questões de implementação, **todas** as respostas devem ser manuscritas. O material entregue deve consistir da cópia digitalizada da folha, por exemplo, uma foto **legível** da folha de soluções.
3. Os exercícios de implementação podem ser feitos na linguagem de programação que quiser, desde que para executar o seu código não seja necessária a instalação de **nenhum** software com custo maior que zero, isso inclui o Windows :).
4. Qualquer tentativa de fraude identificada será punida de acordo com o código de honra.
5. Além das respostas de cada pergunta no formato textual, você deve escolher um dos exercícios marcados com \star e enviar o link para *download* de um vídeo descrevendo a solução do exercício. O vídeo deve ter no máximo 5 minutos e no início deve aparecer o rosto e carteirinha do aluno. A falha no recebimento do vídeo fará a lista ser desconsiderada.

Exercício 1 Manipulando dados reais

Escolha qualquer base de dados do repositório da UCI. Você deve analisar as estatísticas básicas da base de dados e gerar pelo menos três gráficos de tipos diferentes. Utilize aquilo que foi visto na Prática 01 e na Prática 02. Note que embora as aulas práticas sejam em Python, você é livre para implementar em qualquer linguagem que queira, conforme já apresentado e discutido no plano de ensino.

Exercício 2 \star Comparando LDA e QDA

Implemente, na linguagem de sua preferência, os algoritmos LDA e QDA vistos nos vídeos. Considere a base de dados disponível aqui, gere os gráficos da fronteira de decisão similar aos dos slides. Você não deve gerar as linhas das fronteiras, apenas os pontos coloridos de acordo com a classificação obtida pelo modelo é suficiente, como feito no slide 20 do material de LDA. Você pode utilizar bibliotecas para computar a inversa das matrizes e os determinantes.

Obs: caso tenha problemas com o link no PDF: https://www.dropbox.com/s/r8kxr8pl2awprie/uma_base.csv?dl=0