

Trabalho para Aprendizado de Máquina

Prof. Marcos Prates

Data de Entrega: 30/11

1 O Projeto

Neste projeto, você irá aplicar algumas das técnicas vistas em aula para um problema real. O projeto DEVE ser feito individualmente.

2 Dados

O banco de dados (train.csv) consiste em característica de 4532 jogadores das ligas europeias. As variáveis `Contry_Club` e `Club_Position` representam o país do time do jogador e sua posição em campo respectivamente. Além dessas 2 variáveis principais existem várias outras características disponíveis dos atletas. O banco de dados test.csv possui a característica de 500 jogadores omitias as colunas `Contry_Club` e `Club_Position`.

3 Problema 1

Você deve treinar um classificador para determinar no teste a posição na qual o jogador joga.

4 Problema 2

Baseado nas características dos jogadores no teste você deve recomendar 3 países apropriados para o jogador jogar.

5 Relatório

O objetivo do relatório é apresentar os resultados obtidos no problema e obter um classificador e recomendador com boa performance. Assim espera-se de você:

- Uma análise descritiva dos dados.
- Uma descrição matemática dos problemas a serem resolvidos. Quem é X ? Quem é Y ? Que função de risco será utilizado? Qual recomendador?
- Quais classificadores você implementou? É esperado que você compare ao menos 4 métodos distintos.
- Quais recomendador você implementou?
- Qual método forneceu melhores resultados? Como você mediu isso? Qual a performance de cada um deles?

- Utilize alguma técnica de redução de dimensionalidade. Plote as duas ou três primeiros componentes encontrados. Você consegue encontrar alguma padrão/cluster nos dados?
- Descreva quais as dificuldades encontradas neste projeto e o que foi feito para superá-las.
- Se você tivesse mais tempo, o que faria de diferente? Como acha que poderia obter resultados ainda melhores?

6 Considerações Finais

- O banco de treino pode ser repartido como achar apropriado em cada problema para responder as perguntas do trabalho.
- Além do relatório em pdf todos os códigos usados para gerar todas as análises deve ser anexados.
- Um código em separado para o problema 1 e problema 2 deve ser entregue aplicando o classificador/recomendador criado para rodar no banco de teste. A pessoa que tiver maior acertividade em cada um dos problemas ganhará **5 pontos extra!!!**
- as entregas devem ser feitas pelo Moodle até 23:55 do dia 30/11/2018.