

Machine Learning I

Instruções para o projeto final

O projeto. Ao longo do módulo, discutimos extensivamente sobre diversas etapas do processo de modelagem quando desenvolvemos um modelo de Machine Learning (ML), seja para fins de classificação, ou de regressão. Neste projeto, em grupos, vocês deverão escolher uma base de dados do interesse de vocês, como as disponíveis no *kaggle* ou em outros repositórios similares, para trabalhar em um problema de predição. O problema pode ser de regressão ou classificação, ficando esta escolha a critério do grupo. Inclusive, sintam-se livres para trabalhar com problemas em que as duas abordagens possam ser possíveis (por exemplo: uma base de dados originalmente pensada para regressão, mas a partir da qual podem ser criadas categorias de classificação).

Entregável. O grupo deverá entregar um notebook contendo todas as etapas de modelagem, incluindo:

- Introdução/contextualização do problema, com o objetivo/perguntas que o grupo se propõe a responder;
- análise exploratória de dados, com todas as etapas necessárias de processamento dos dados;
- etapas de modelagem (treinamento do modelo, validação cruzada, otimização de hiperparâmetros e predição).

O notebook deverá ser hospedado em um repositório no github, sendo a entrega feita a partir do link correspondente.

Apresentação do projeto. Na última aula do módulo (27/10) acontecerão as apresentações do projeto. Apesar de uma parte da avaliação do projeto consistir na análise do notebook, os integrantes também serão avaliados pela sua capacidade de comunicação e síntese dos resultados. Cada grupo terá até 15 min para sua apresentação. O formato é livre, ou seja: não é necessário criar qualquer tipo de apresentação mais formal (embora seja uma possibilidade, caso o grupo prefira).

Quaisquer dúvidas e outros problemas, estou à disposição!