



Universidade Estadual de Maringá

Departamento de Informática

Prof. Dr. Wagner Igarashi

wigarshi@din.uem.br



Trabalho 1

Equipes de 2 a 3 participantes.

Cada equipe deverá escolher um tema não repetido dentre os listados a seguir e procurar um database no Kaggle:

Nro Tema	Tema
1	Previsão da pobreza doméstica na Costa Rica
2	Problema de previsão de AVC
3	Previsão do preço do carro
4	previsão de vendas de Mercado
5	Análise de Previsão de Empréstimos
6	Prevendo a perda de funcionários
7	Previsão de reserva de hotel
8	Previsão de temperatura aparente
9	Classificação de Reclamações do Consumidor
10	Projeto de esboço ao vivo
11	Dança Tradicional
12	Detector de doenças foliares
13	Identificação da Marca do Carro
14	Projeto de Análise de Tweet
15	Análise de Emoções usando PNL
16	Sistema Previsor de Desastres
17	sistema de revisão de filme IMDB
18	Previsão de mortalidade em UTI usando RNA

Encontrar e informar de qual site foi extraído o notebook python e qual foi a técnica originalmente utilizada (chamada aqui de técnica A).

Cada aluno deverá aplicar uma técnica de machine learning diferente da originalmente aplicada (técnica A). Explicar: qual é o contexto dos dados, qual tarefa de aprendizagem de máquina foi realizada (classificação, regressão, clusterização, predição, etc), como foi realizada a extração de características, quais foram os atributos selecionados, mostrar algumas estatísticas descritivas dos atributos, gráficos, utilizar crossvalidation com diferentes k-folds e mostrar qual foi a melhor configuração para tentar ter um resultado melhor do que a técnica A (usar grid search). Apresentar uma tabela final comparando e rankeando os resultados da avaliação das técnicas utilizadas com a técnica A.

Caso o estudo do Kaggle já apresente a aplicação de diversas técnicas, as técnicas que cada membro da equipe aplicar devem ser diferentes das técnicas originais.



Universidade Estadual de Maringá

Departamento de Informática

Prof. Dr. Wagner Igarashi

wigarshi@din.uem.br



Os resultados devem ser apresentados em formato de slides, os quais devem ser encaminhados juntos com o código fonte (preferencialmente utilizando python ou python notebook).

- **trabalhos entregues em atraso terão nota 0.**

Obs: A avaliação do trabalho será realizada utilizando os seguintes critérios: originalidade, complexidade do problema abordado, aparência da interface de interação com o usuário, apresentação dos resultados, completude dos itens especificados, português correto, código compilável e executável. **Será utilizado software de verificação de similaridade, assim se houver indícios de cópia de trabalho, os trabalhos envolvidos terão nota 0.**