

Trabalho 01- Classificação

1001524 – Aprendizado de Máquina I
2022/1 - Turmas A, B e C
Prof. Dr. Murilo Naldi
PESCD: Bruno Silva Sette

naldi@ufscar.br

Descrição

- Primeiro Trabalho
 - Tarefa de classificação
 - Duplas ou trios
 - Composto de três etapas:
 - Escolha do conjunto de dados
 - Estudo do conjunto de dados
 - Classificação

Escolha do conjunto de dados

- Cada dupla formada deve ir até o fórum do AVA da disciplina
 - Olhar todos os *posts* dos outros grupos
 - Escolher um conjunto de dados para trabalhar
 - Revelante
 - Desafiador
 - Complexo
- Postar no AVA o nome dos membros da dupla e descrição do conjunto
 - Grupos que usem conjuntos já citados em *posts* terão notas penalizadas

Estudo do conjunto de dados

- Explorar o conjunto de dados é tão ou mais importante do que a classificação em si
 - Solucionar problemas de atributos
 - Visualização
 - Seleção e redução de atributos
 - **Justificar a necessidade de cada técnica aplicada!**
- É precisa deixar o conjunto em condições ideais para aplicação

Classificação

- Aplicação de diferentes métodos de classificação
 - Justificar a escolha
 - Ajuste de parâmetros
 - Comparar adequadamente os resultados
 - Gráficos ilustrando resultados

Material e entrega

- Entrega e apresentação dos trabalhos no dias:
 - Turma A : 26 de julho às 14 horas!
 - Turma B : 29 de julho às 14 horas!
 - Entrega material no AVA
 - Apresentação em sala de aula
 - Aproximadamente 15 minutos com entrevista
- Grupos podem escolher
 - Todo material/trabalho seja feito em forma de relatório + apresentação (*slides*)

Notas

- Cada fase do trabalho vale 1/3 da nota
 - Escolha, estudo e classificação
- Nota do trabalho
 - Nota do Relatório (NR) = 0 – 10
 - Nota da Apresentação (NA) = 0 – 1
 - $NF = NA * NR$
- Avaliação individual das apresentações

Exemplos repositórios de dados

- <https://www.kaggle.com/>
- <https://www.nature.com/sdata/policies/repositories>
- http://oad.simmons.edu/oadwiki/Data_repositories
- <https://archive.ics.uci.edu/ml/datasets.html>