

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
INSTITUTO DE CIÊNCIAS EXATAS  
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO  
Cálculo Numérico – Primeiro Semestre 2020

Mineração de dados utilizando álgebra linear  
Trabalhos Práticos 1 e 2

Observações

- 1) O trabalho é **estritamente** individual. Trabalhos julgados iguais terão nota zero sem prejuízo das demais sanções.
- 2) Data limite para entrega: 23/09/2020; valor 25 pontos no total. Gerar um arquivo em formato **pdf** e aguardar instruções do repositório para o qual estes trabalhos deverão ser postados.
- 3) A documentação é parte fundamental do trabalho; favor inteirar-se do formato sugerido para o relatório.
- 4) Sinta-se à vontade em procurar-me para esclarecer dúvidas e/ou resolver problemas. Estarei à disposição às terças e quintas via *e\_mail* institucional.
- 5) Assistir no meu canal do *YouTube* a apresentação destes trabalhos presentes

**Trabalho Prático 1**

Refazer a Aula Prática de Mineração de Dados utilizando Álgebra Linear (texto distribuído) e produzir um relatório nos moldes apresentados

**Trabalho Prático 2**

Seja o conjunto de sequências '*primates*' presentes no toolbox de Bioinformática do **MatLab** (basta executar '*load primates*'). Caso você disponha de um outro conjunto de sequências (  $> 20$  ), sinta-se à vontade de usá-lo neste trabalho em substituição ao conjunto *primates*.

- a) Usar um *vector space model* para representar essas sequências; o resultado será uma matriz  $A$  de dimensões adequadas aos nossos propósitos delineados a seguir. Proceder à decomposição por valores singulares e visualizar as entidades no espaço  $R^3$ .
- b) Usando os recursos do Matlab, produzir árvores hierárquicas associadas .
- c) Usar o *Kmeans* para separar os grupo

Marcos Augusto dos Santos