

# Trabalho de Introdução à Ciência de Dados

Trabalho referente à A2 do segundo semestre de 2021

Valor: 10 pontos

Trabalho pode ser feito sozinho ou em dupla

**Entregar um .pdf, incluindo a justificativa da representação gráfica, o código e a imagem final produzida. Apresentação ultrarrápida (ênfase no seu principal resultado).**

Atenção ao nome dos arquivos: <nome\_do\_aluno1>\_<nome\_do\_aluno2>\_<tema\_trabalho>

Questão 1) Escolha um dos métodos a seguir:

- Análise de Componentes Principais
- Técnica de Clusterização de dados
- Modelos de Regressão
- Modelos de Classificação
- Modelos de Predição

Questão 2) Faça um breve ensaio explicando o que é método escolhido e indique quais os modelos mais utilizados

*Nessa questão é esperado que o aluno faça um breve ensaio sobre o método escolhido na sessão anterior. Nesse ensaio, é interessante destacar história do método, modelos comumente utilizados, diferença entre os modelos, limitações dos modelos, entre outros pontos.*

Questão 3) Explique um exemplo de uso dos modelos

*Nessa questão é esperado que o aluno explique pelo menos um exemplo de uso de pelo menos um dos modelos mais comumente utilizados destacados na sessão anterior. O exemplo de uso deve conter o contexto no qual o modelo é usado, qual o objetivo do uso do modelo e explicar os resultados encontrados.*

Questão 4) Escolha um modelo dentre os indicados na questão 3 explique como ele funciona utilizando um pseudocódigo

*Nessa questão é esperado que o aluno explique por meio de um pseudocódigo pelo menos um dos modelos ressaltados na questão 3. É importante destacar que um pseudocódigo não é a mesma coisa que um código: a ideia é explicar o funcionamento do modelo e a importância de cada um dos outputs das funções que executam o modelo. O uso de um código não isenta o aluno de escrever um pseudocódigo que explique o funcionamento do modelo escolhido.*

Questão 5) Indique quais as principais bibliotecas e funções utilizadas de acordo com a linguagem Python (ou a linguagem R) para o uso do método escolhido

*Nessa questão é esperado que o aluno indique as principais bibliotecas e funções utilizadas caso queira aplicar o modelo escolhido na questão 4. Caso na questão anterior, o aluno tenha escrito um código além de um pseudocódigo, basta ele comentar o código indicando com qual fim cada uma das bibliotecas foi utilizada e comentar as funções utilizadas indicando de qual biblioteca elas fazem parte e qual o objetivo de cada uma das funções.*

Questão 6) Indique como os modelos do método escolhido podem ter a performance avaliada

*Nessa questão é esperado que o aluno discuta os principais métodos de comparação de performance de diferentes modelos. Para além de indicar como cada método é usado, é importante explicar o método de comparação em si, a intuição por trás dele, indicar como se calcula, quais as suas limitações, indicar métodos alternativos, entre outros pontos.*

Ao final, é importante fazer comentários gerais sobre o seu aprendizado no processo. E, caso o trabalho tenha sido feito em dupla, explicitar qual foi a dinâmica de trabalho.