

Disciplina/módulo: Machine Learning para Bioinformática
Professor: Gilderlanio Santana de Araújo

Exercício I

1. Dados e tipos de dados

- a) Defina o conceito de "dados" e explique sua importância na era da informação.
- b) Descreva os diferentes tipos de dados existentes, como dados categóricos, numéricos e binários, e forneça exemplos de cada tipo.
- c) Discuta a relevância da qualidade dos dados para a obtenção de resultados confiáveis em análises e modelos de aprendizado de máquina.

2. Imputação de dados:

- a) Explique o que é imputação de dados e quais são seus objetivos principais.
- b) Descreva os diferentes métodos de imputação de dados. Pesquise na literatura e métodos estatísticos mais complexos multivariados.
- c) Discuta as vantagens e desvantagens de cada método de imputação de dados e quando cada um é mais adequado.
- d) Apresente exemplos práticos de situações em que a imputação de dados é necessária e quais os impactos que a falta de tratamento dos dados ausentes pode ter em análises e modelos de aprendizado de máquina.

3. Aprendizado supervisionado:

- a) Defina o conceito de aprendizado supervisionado e explique seu funcionamento em termos de treinamento e predição.
- b) Apresente exemplos de aplicações práticas do aprendizado supervisionado em diferentes áreas do conhecimento