

Introdução a Aprendizagem de máquina para Bioinformática

Aula 2

Msc. Amanda Araújo Serrão de Andrade

Biomédica, Mestre em Modelagem Computacional e Doutoranda em Genética

Orientadora: Dra. Ana Tereza Ribeiro de Vasconcelos

Sumário da aula

- Classificação supervisionada (teórica e prática)
- Regressão (prática complementar)
- Clusterização hierárquica (prática complementar)

Overview do minicurso

Aula 2

- Classificação
- Clusterização e Regressão (prática complementar)

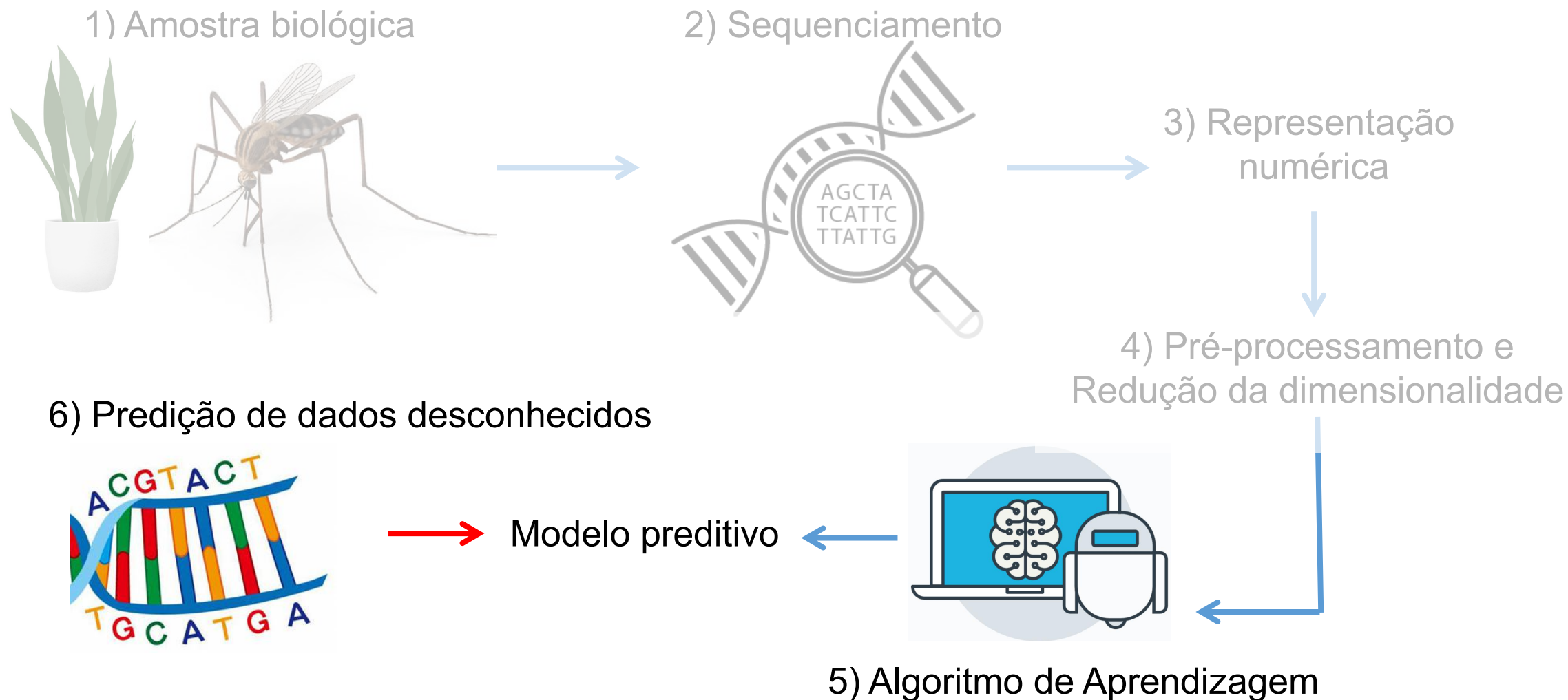
Aula 3

- Principais métricas de avaliação
- Resultados
- Comparação de performance

Overview do minicurso

Objetivo da aula: executar nossa primeira classificação supervisionada
<ul style="list-style-type: none">• Classificação• Clusterização e Regressão (prática complementar)
Aula 3
<ul style="list-style-type: none">• Principais métricas de avaliação• Resultados• Comparação de performance

Exemplo prático: Como identificar se um vírus desconhecido infecta plantas ou mosquitos?



A escolha dos algoritmos de Aprendizagem de máquina

- Para cada um dos problemas — Classificação, Regressão, Clusterização e Agrupamento — existem diversos algoritmos que podem ser utilizados.
- A cada problema, os algoritmos disponíveis devem ser experimentados a fim de identificar aqueles que obtêm melhor desempenho.

5) Algoritmos de Aprendizagem de máquina

Classificação supervisionada

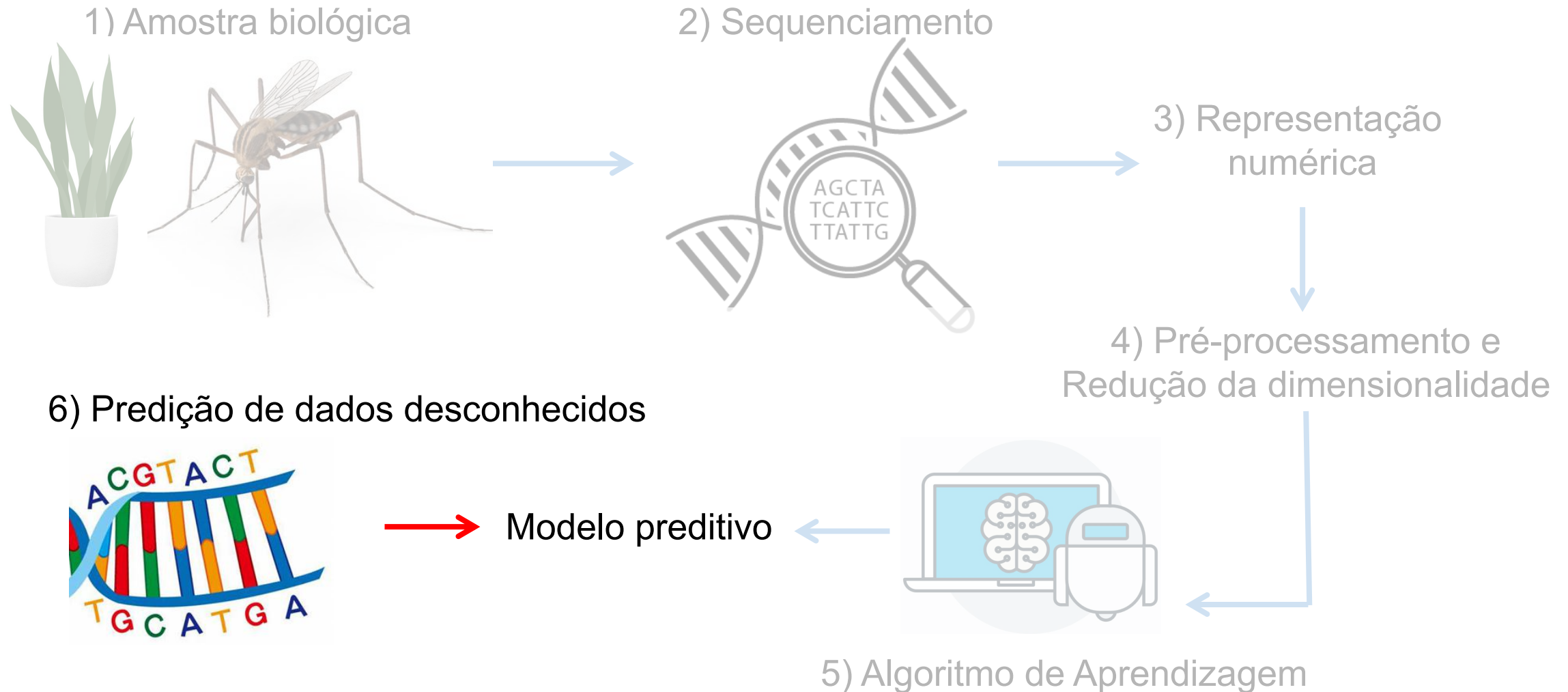
Principais algoritmos:

- randomForest;
- AdaBoost;
- k-NN;

Classificadores *ensemble*: Associação de múltiplos algoritmos de classificação.

Lista completa dos algoritmos disponíveis: <https://topepo.github.io/caret/available-models.html>

Exemplo prático: Como identificar se um vírus desconhecido infecta plantas ou mosquitos?



Práticas da aula

- `classification_Aula2.R`
- `hierar_clustering_Aula2.R` (prática complementar)
- `regression_Aula2.R` (prática complementar)

Referências e material de apoio

- <https://ilumeo.com.br/todos-posts/2020/06/22/um-tour-pelos-10-principais-algoritmos-de-machine-learning>
- https://awari.com.br/algoritmos-de-classificacao/?utm_source=blog
- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. An Introduction to Statistical Learning : with Applications in R. New York :Springer, 2013.
- machinelearningmastery.com
- Youtube: StatQuest



Dúvidas?

aandradebio@gmail.com

- As dúvidas mais específicas devem vir acompanhadas com código e um conjunto de dados de teste para agilizar a resposta



Obrigada!!

aandradebio@gmail.com

atr@Incc.br