Sumário

[Visão Geral 1](#_Toc181803552)

[T-student - Performance 1](#_Toc181803553)

[Informações 1](#_Toc181803554)

[Hipóteses 1](#_Toc181803555)

[Resultados 2](#_Toc181803556)

[T-student – Tempo de Execução 2](#_Toc181803557)

[Informações 2](#_Toc181803558)

[Hipóteses 2](#_Toc181803559)

[Resultados 2](#_Toc181803560)

[Conclusões 3](#_Toc181803561)

## Visão Geral

Durante a produção de uma machine learning, que tem o objetivo de prever quais seriam os possíveis clientes do aplicativo produzido durante o interdiciplinar, foram realizados vários testes para analisar a performance de diferentes modelos em diversas condições. Ao final, foi chegado em um empasse, os dois melhores modelos apresentavam boas métricas e bons tempos de execução, entretanto o modelo de K-Nearest Neighbors (KNN) apresenta o melhor tempo de execução e o modelo de Árvore de Decisão apresenta a melhor performance. Para analisar qual modelo seria mais vantajoso para o projeto, utilizara os dados captados durante a validação cruzada realizada com eles.

# T-student - Performance

## Informações

O objetivo deste teste é entender se a performance do modelo de Árvore de Decisão (representado por B) seria significativamente maior que a performace do modelo de KNN (representado por A). Essa busca serve para entender se teria um grande impacto no desempenho caso seja escolhido o modelo de KNN.

Amostras de A (KNN) = [

0.8861538461538462,

0.9177215189873418,

0.9150326797385621,

0.9074074074074074,

0.88]

Amostras de B (Árvore de Decisão):

[0.9245283018867925,

0.9316770186335404,

0.9174311926605505,

0.91875,

0.9382716049382716]

## Hipóteses

A hipótese alternativa carrega a afirmação que a média do tempo de execução do modelo A é menor que a média do tempo de execução do modelo B, ela seria apresentada da seguinte forma:

**H1 = µB > µA**

A hipótese nula consequentemente carrega a afirmação que a média do tempo de execução do modelo A é maior ou igual a média do tempo de execução do modelo B, ela seria apresentada da seguinte forma:

**H0 = µB ≤ µA**

## Resultados

Tem fortes indícios de que a performance do modelo B é significativamente maior que a performance do modelo B.

Não há indícios de que a performance do modelo B é significativamente maior que a performance do modelo B.

# T-student – Tempo de Execução

## Informações

O objetivo deste teste é entender se o tempo de execução do modelo de KNN (representado por A) seria significativamente menor que o tempo de execução do modelo de Árvore de Decisão (representado por B). Essa busca serve para entender se teria um grande impacto no tempo de execução caso seja escolhido o modelo de Árvore de Decisão.

Amostras de A (KNN) : [

0,

0.03315179999844986,

0.028251200001250254,

0.019164100001944462,

0.01984499999889522,

0.03355329999976675,

0.017156900001282338]

Amostras de B (Árvore de Decisão) : [

0.027779800000644173,

0.02602599999954691,

0.021471400001246366,

0.018629399999554153,

0.022458800000094925,

0.021002900000894442]

## Hipóteses

A hipótese alternativa carrega a afirmação que a média do tempo de execução do modelo A é menor que a média do tempo de execução do modelo B, ela seria apresentada da seguinte forma:

**H1 = µA < µB**

A hipótese nula consequentemente carrega a afirmação que a média do tempo de execução do modelo A é maior ou igual a média do tempo de execução do modelo B, ela seria apresentada da seguinte forma:

**H0 = µA ⩾ µB**

## Resultados

Tem fortes indícios de que o tempo de execução do modelo A é significativamente menor que o tempo de execução do modelo B.

Não há indícios de que o tempo de execução do modelo A é significativamente menor que o tempo de execução do modelo B.

# Conclusões

Analisando os resultados, ambos os testes são favoráveis ao modelo de Árvore de Decisão, já que indicam que ele tem uma performance significativamente maior e o tempo de execução que o mais interessante para o projeto seria utilizar o modelo de Árvore de Decisão.