

Universidade do Sul de Santa Catarina – Unisul
UC – Inteligência Artificial
Nome: Allan dos Santos
RA: 107223124680

Relatório de Modelagem Preditiva

Introdução

Este documento apresenta um modelo preditivo para prever o turnover de funcionários, utilizando diversos algoritmos de classificação. Os modelos incluem Árvore de Decisão, K-Nearest Neighbors, Naive Bayes, Regressão Logística e Redes Neurais. Cada modelo foi avaliado com base na acurácia e em suas respectivas matrizes de confusão para comparação de desempenho.

Descrição dos Algoritmos Utilizados

Árvore de Decisão

A Árvore de Decisão é um modelo que divide os dados em ramos para tomada de decisão. Ela é fácil de interpretar e pode capturar interações complexas entre variáveis.

K-Nearest Neighbors

K-Nearest Neighbors (K-NN) classifica dados com base nos rótulos dos k vizinhos mais próximos. É um algoritmo simples, mas pode ser sensível a dados ruidosos.

Naive Bayes

O Naive Bayes é um modelo probabilístico baseado no Teorema de Bayes. Ele assume independência condicional entre as variáveis preditoras.

Regressão Logística

A Regressão Logística é um modelo de classificação que usa uma função logística para modelar a probabilidade de uma amostra pertencer a uma classe específica.

Rede Neural

Redes Neurais são modelos compostos por neurônios artificiais, que podem aprender padrões complexos em grandes conjuntos de dados.

Resultados e Avaliação

Acurácia dos Modelos:

Decision Tree: 82.00%

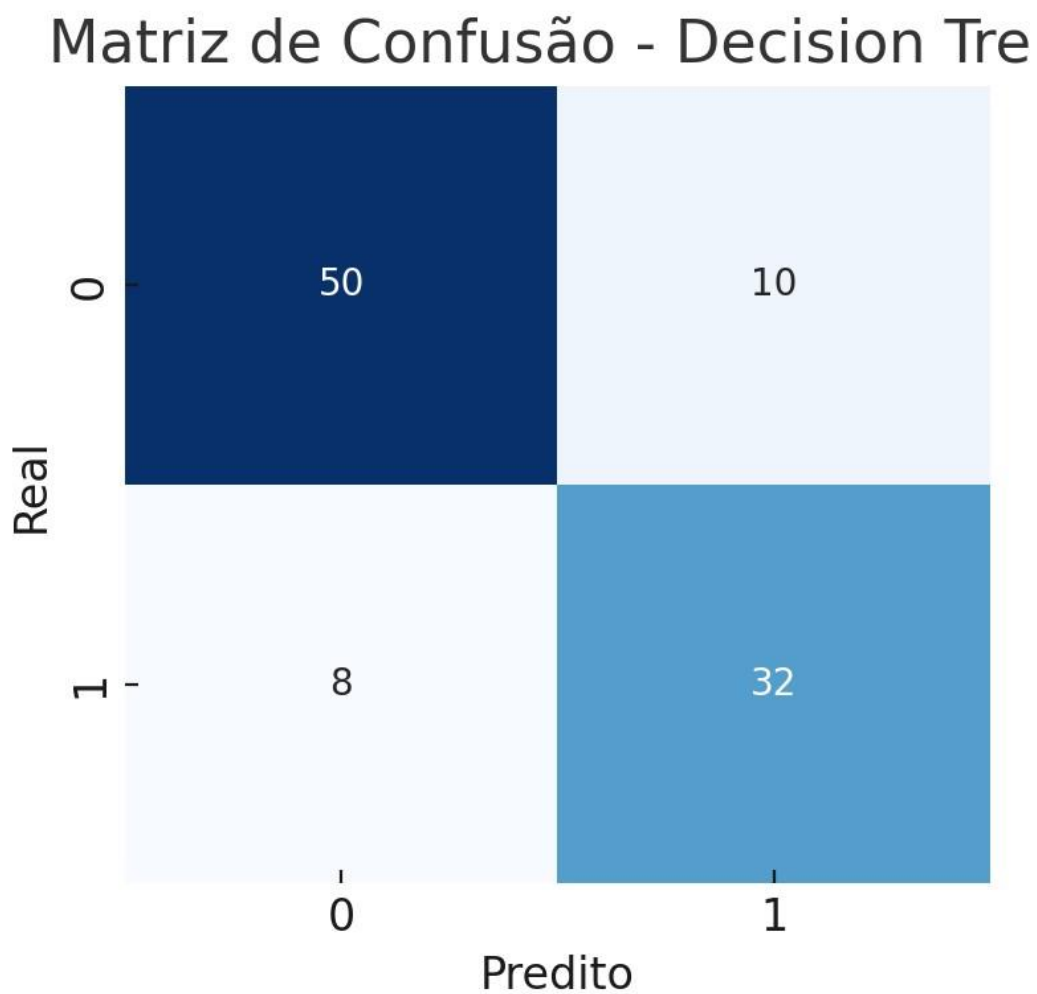
K-Nearest Neighbors: 78.00%

Naive Bayes: 79.00%

Logistic Regression: 82.00%

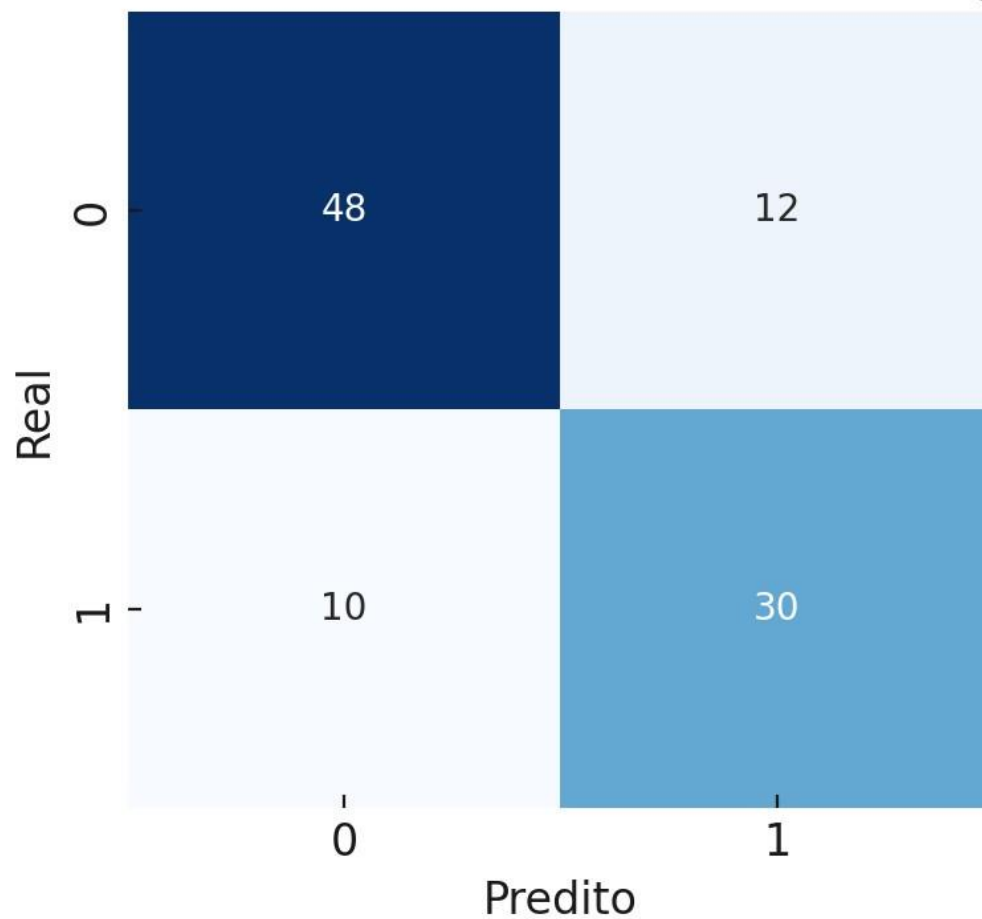
Neural Network: 86.00%

Matriz de Confusão - Decision Tree



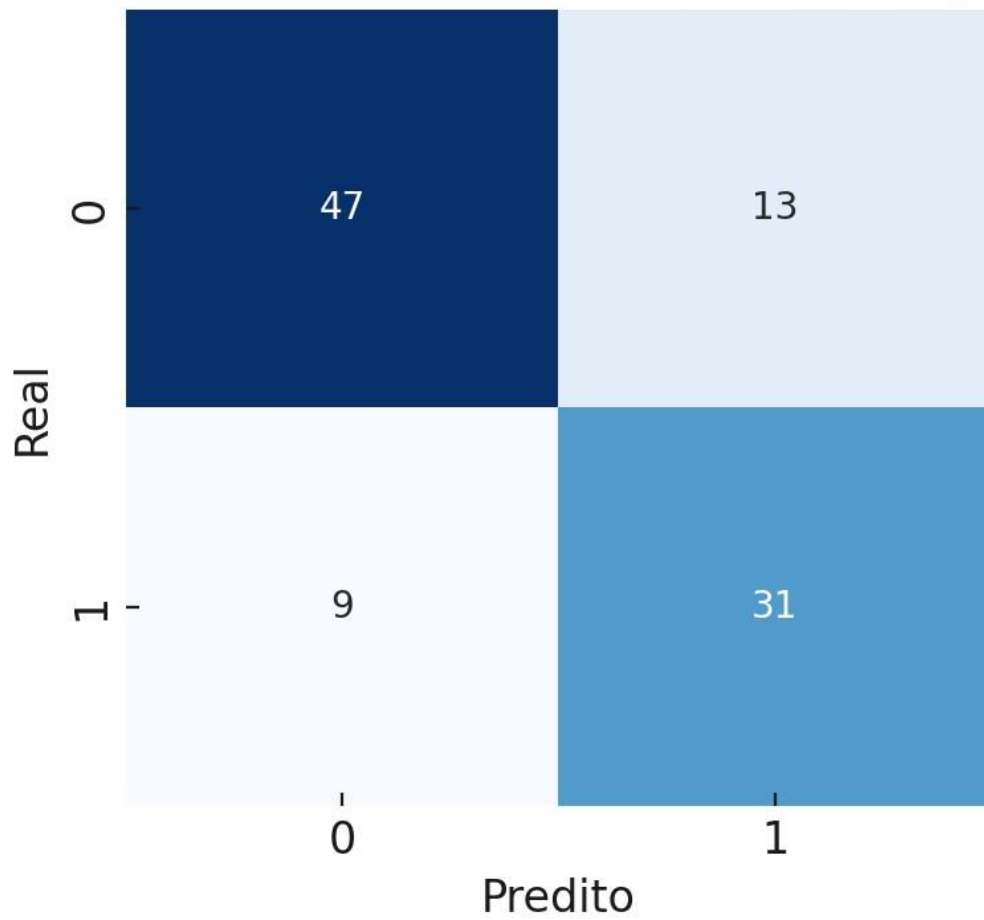
Matriz de Confusão - K-Nearest Neighbors

Matriz de Confusão - K-Nearest Neigh



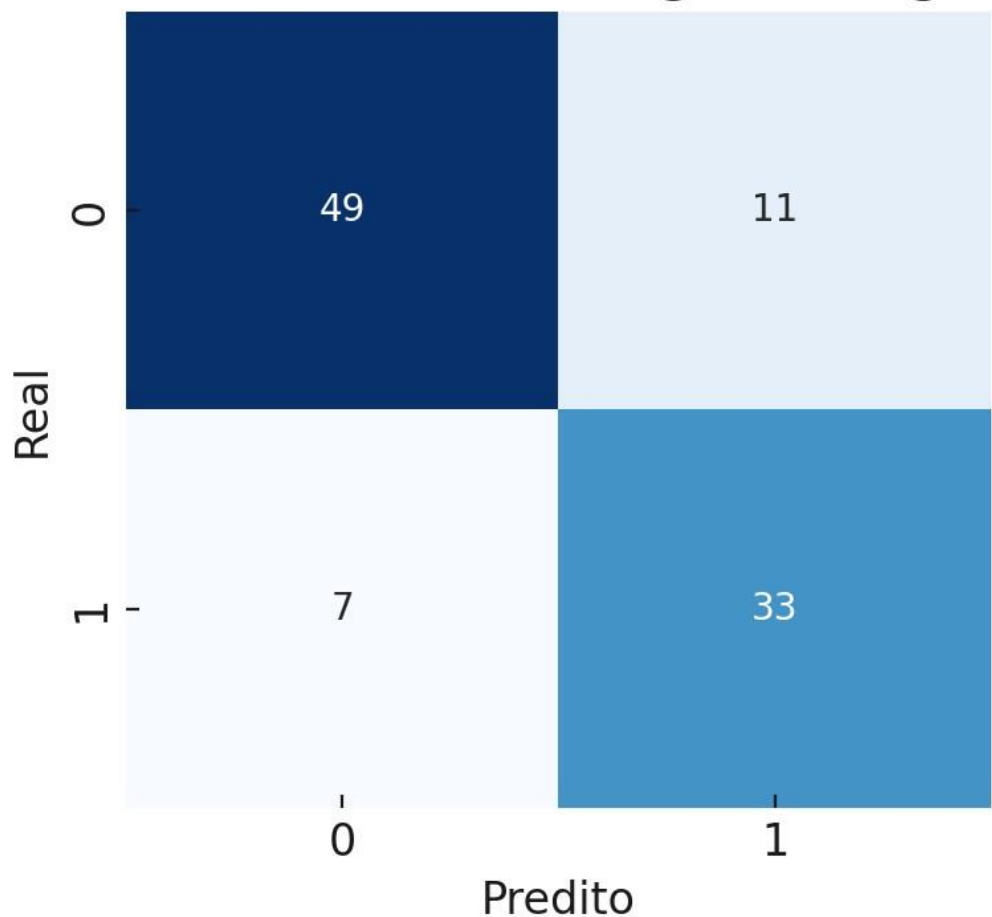
Matriz de Confusão - Naive Bayes

Matriz de Confusão - Naive Bayes



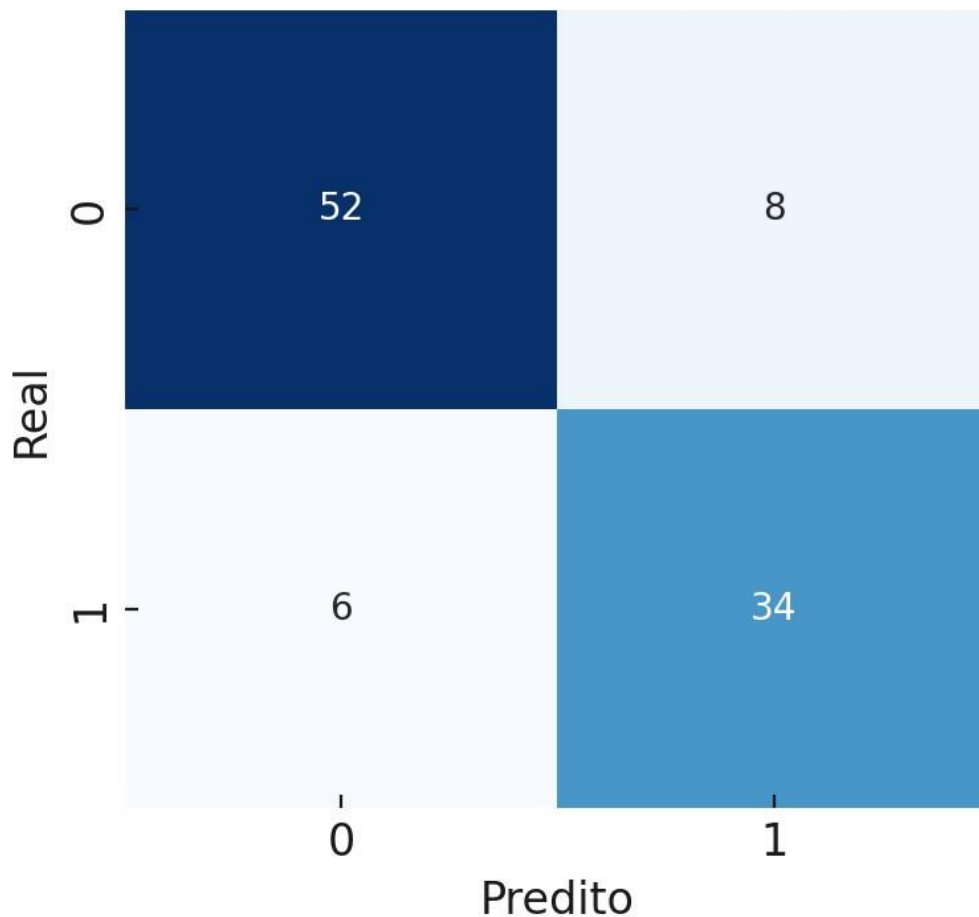
Matriz de Confusão - Logistic Regression

Matriz de Confusão - Logistic Regres



Matriz de Confusão - Neural Network

Matriz de Confusão - Neural Netwc



Conclusão

Este relatório apresentou e comparou o desempenho de vários algoritmos de classificação para prever o turnover de funcionários. A análise das métricas de acurácia e das matrizes de confusão permite identificar que, neste caso, o modelo de Redes Neurais apresentou o melhor desempenho em termos de acurácia. A escolha do melhor modelo, no entanto, também pode depender de fatores como interpretabilidade e velocidade de processamento.