

Resultados

Dados descritivos inferidos sobre cada coluna

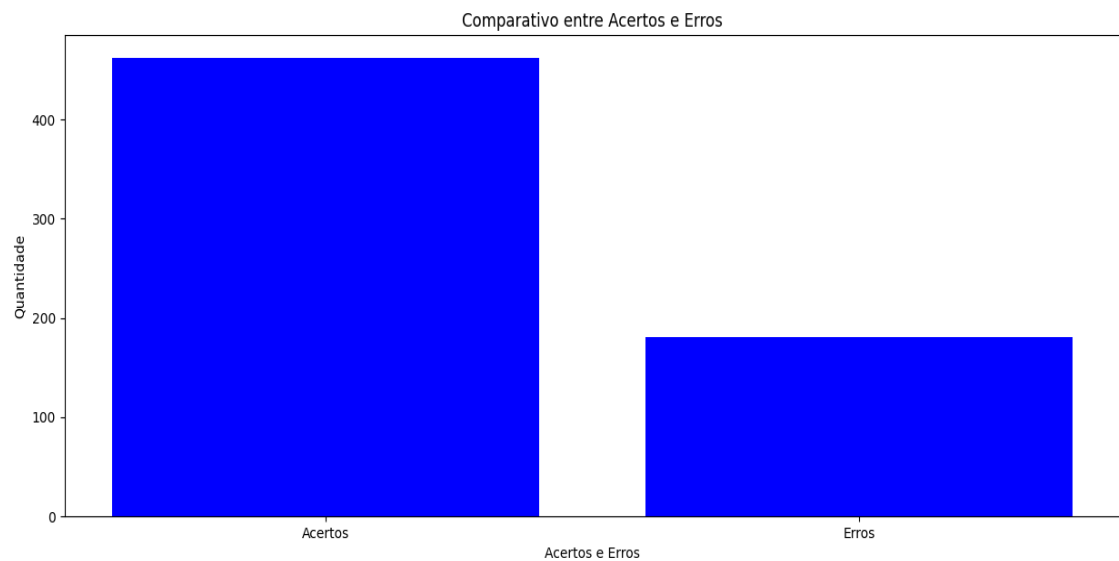
- Coluna True_class (antes da substituição dos valores nulos)
 - Média: 38.57
 - Desvio Padrão: 39.58
 - Valor mais comum: 0.00
 - Variância: 1566.66
- Coluna True_class (após a substituição dos valores nulos)
 - Média: 48.25
 - Desvio Padrão: 38.54
 - Valor mais comum: 74.00
 - Variância: 1485.51
- Coluna Pred_class
 - Média: 52.71
 - Desvio Padrão: 37.60
 - Valor mais comum: 3.00
 - Variância: 1413.92
- Coluna probabilidade
 - Média: 0.62
 - Desvio Padrão: 0.27
 - Valor mais comum: 1.00
 - Variância: 0.07

Dados descritivos inferidos sobre o modelo

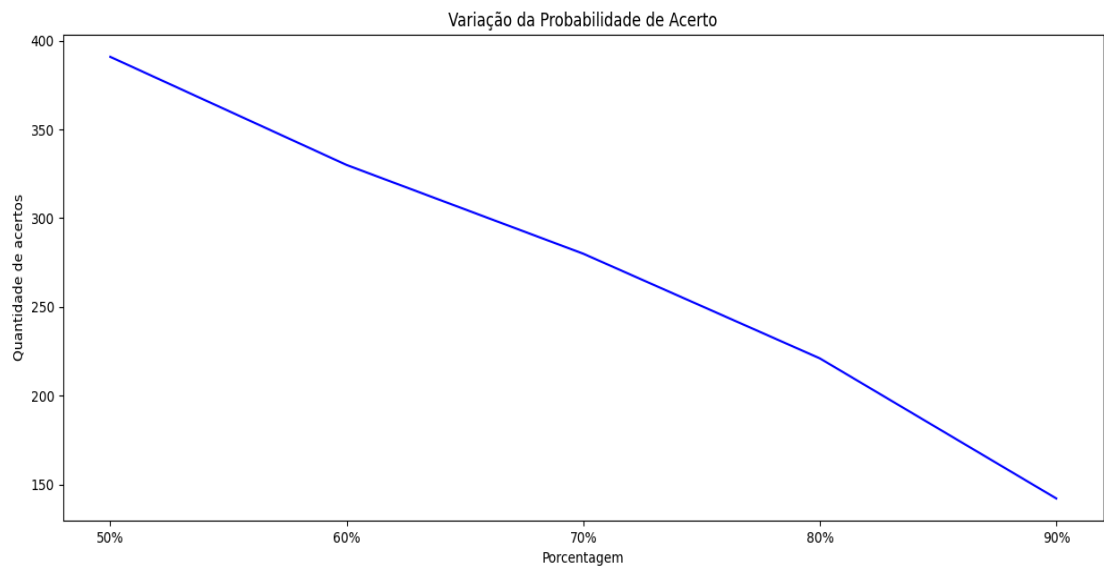
- Valor que houve mais predições corretas: 74 (56x)
- Valor que houve menos predições corretas: 48 (1x)
- Valor que houve mais predições incorretas: 2 (14x)
- Valor que houve menos predições incorretas: 62 (1x)
- Predições corretas feitas pelo modelo original: 0

- Predições corretas feitas pelo modelo após substituição dos valores nulos: 462
- Predições corretas dos valores de 'revision': 0 do total de 43

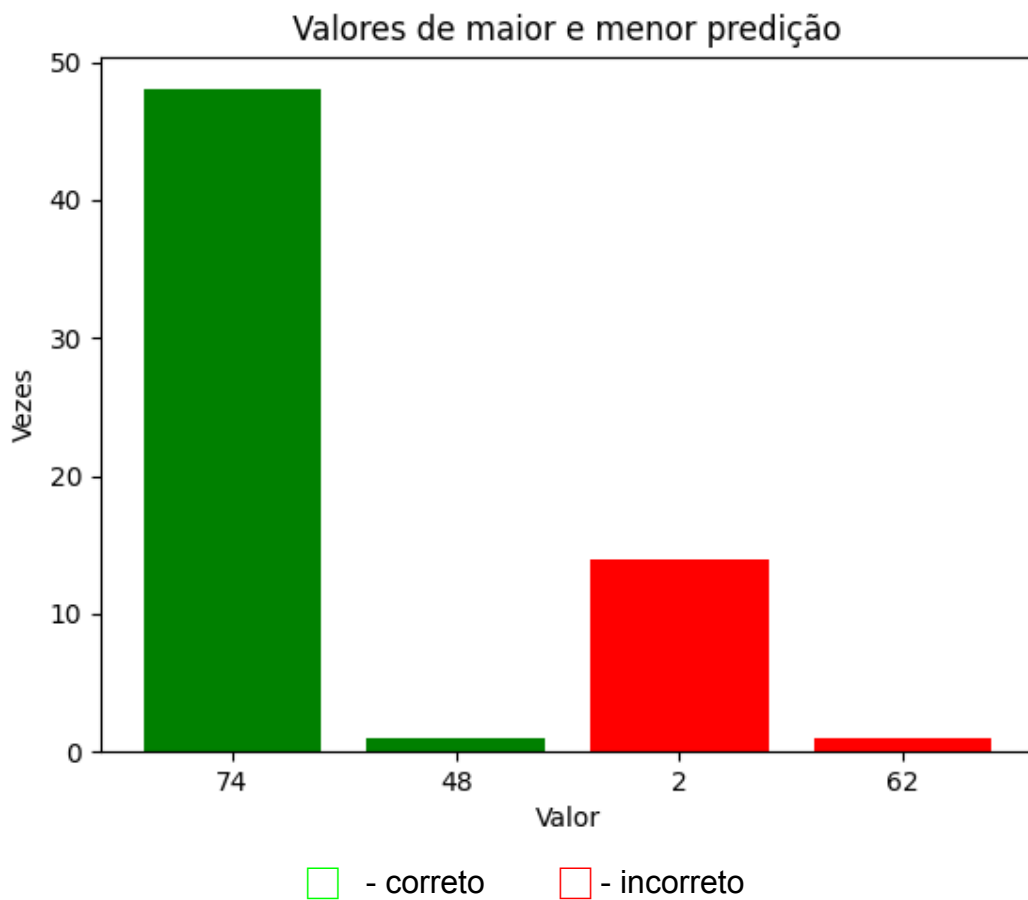
Gráficos gerados



O gráfico acima mostra a diferença entre os acertos e erros do modelo.



O gráfico acima mostra a variação da quantidade de valores por probabilidade de acerto, sendo mostrado os valores a partir de 50% e chegando a 90%.



O gráfico acima mostra um comparativo entre quais valores obtiveram uma predição correta mais e menos frequente, e quais obtiveram uma predição incorreta mais e menos frequentemente.

Métricas de Cálculo de Desempenho

No geral foi alcançada uma precisão que pode ser considerada satisfatória, dado o tamanho reduzido de dados disponíveis no modelo. Segundo a precisão e a acurácia foi obtido um desempenho de 71.85% de acerto nas predições, sua igualdade se deve ao fato de o modelo possivelmente não possuir divergências com falsos positivos e falsos negativos, que poderiam interferir no valor da acurácia. Então pode-se assumir que o modelo tem um grau de confusão baixo. Enquanto que o Mean Absolute Error aponta uma taxa de acerto de 85.65%, assim sendo possível inferir que o modelo mesmo errando ainda prediz valores próximos aos esperados. Quanto ao que pode ter interferido no alcance de um resultado melhor, deve ser considerado o tamanho da base de dados e também os campos vazios.

- Precisão: (71.85%)
 - A precisão é a métrica mais simples de se utilizar para medir o desempenho, sendo simplesmente feita pela divisão da quantidade de acertos pela quantidade total de elementos.
- Mean Absolute Error: (85.65%)
 - Outro método utilizado para calcular a precisão foi o Mean Absolute Error (Erro Médio Absoluto). Seu cálculo se dá por: Erro Absoluto = |Valor Esperado - Valor Obtido| e ao final é calculada a média entre todos erros.
- Acurácia: (71.85%)
 - A acurácia pode ser confundida com a precisão, porém seus cálculos diferem, enquanto a precisão apenas divide os acertos pelo total, a acurácia leva em conta para realizar os cálculos os positivos, negativos, falsos positivos que são predições dadas como corretas quando estavam erradas e falsos negativos que são tidas como erradas quando na verdade estão corretas. Segue a fórmula:

$$\frac{P + N}{P + N + FP + FN}$$

Sendo P os valores de predições corretas, N de predições incorretas, FP os Falsos Positivos e FN os Falsos Negativos.