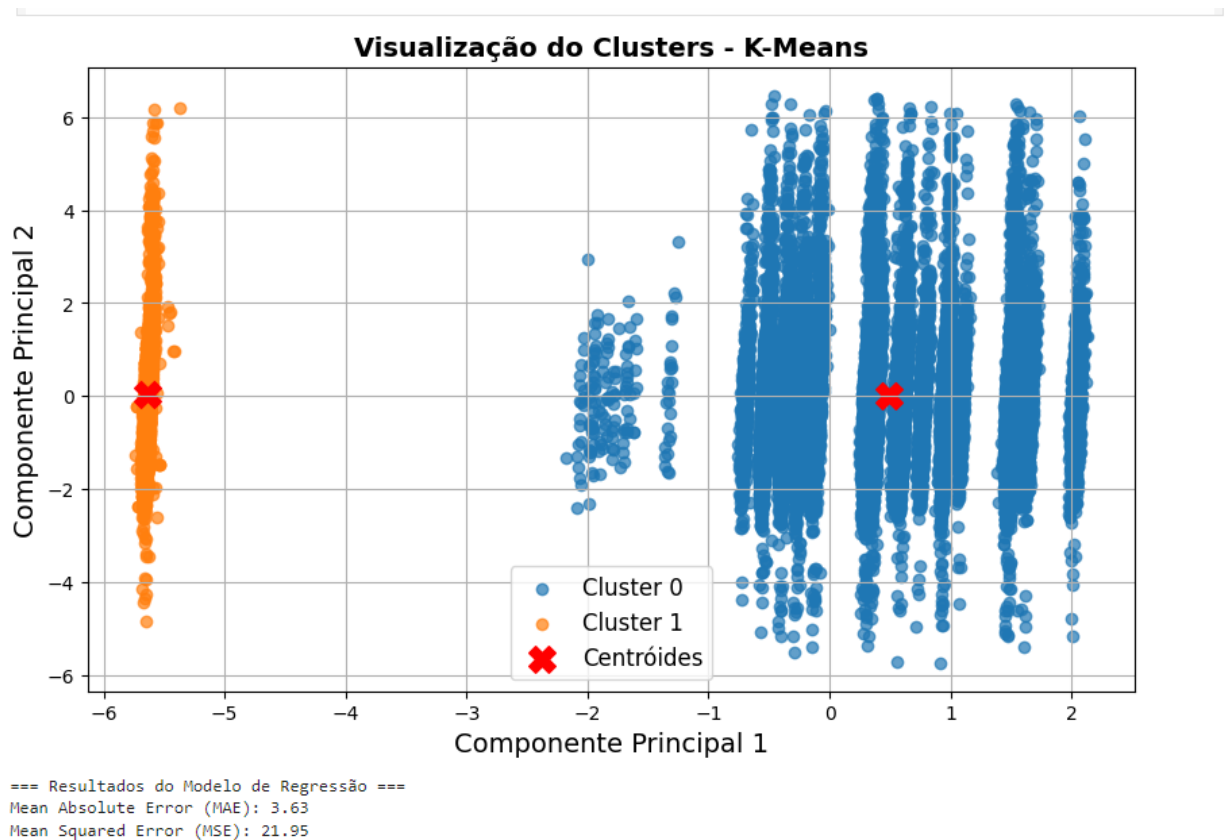


Visualização dos Clusters - K-Means

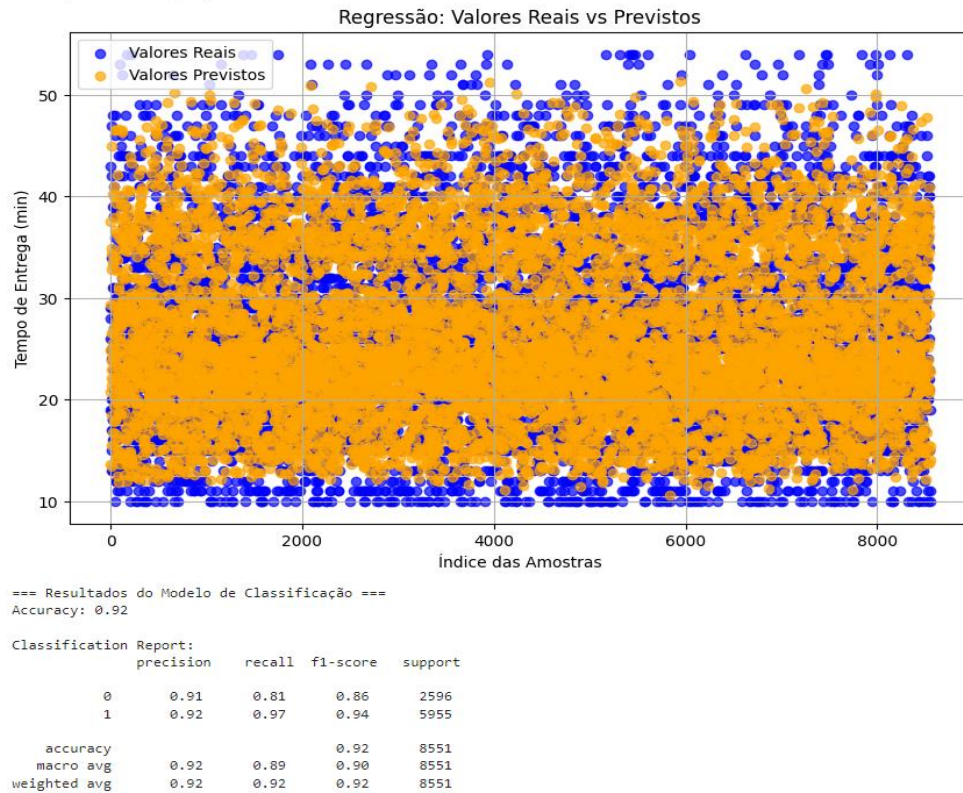


Objetivo do Gráfico: A visualização ajuda a entender como os dados foram agrupados em dois clusters e a distribuição das amostras em relação a essas divisões. A presença dos centróides indica os pontos médios de cada grupo.

- **Pontos:** Cada ponto no gráfico representa uma amostra dos dados, e as cores indicam a qual cluster a amostra pertence.
- **Centróides (marcados com X vermelho):** São os pontos que representam o centro de cada cluster. Eles são calculados pelo K-Means como a média dos dados em cada cluster.

O Gráfico mostra a segmentação dos dados em clusters, o que pode ajudar a identificar padrões subjacentes nos dados.

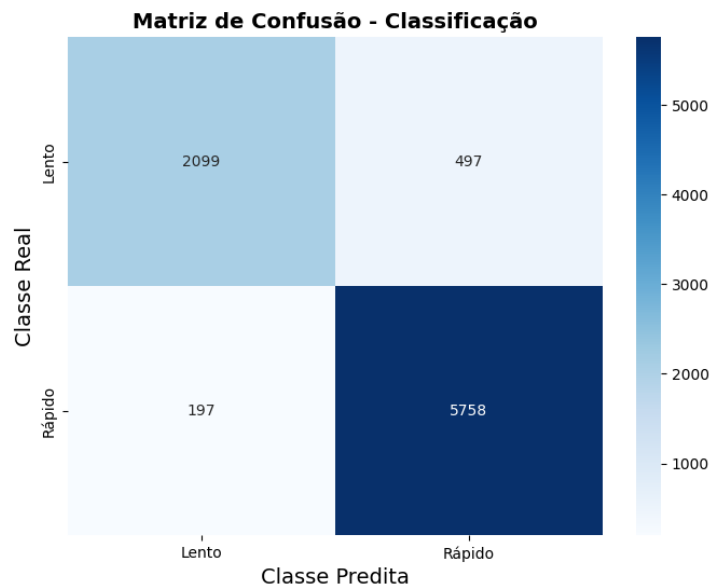
Regressão: Valores Reais vs Previstos



- **Descrição:** Esse gráfico mostra uma comparação entre os valores reais e os valores previstos do tempo de entrega, **Eixo Y:** O tempo de entrega em minutos (tanto os valores reais quanto os previstos).
- **Eixo X:** O índice das amostras no conjunto de teste.
- **Pontos azuis:** Representam os valores reais do tempo de entrega.
- **Pontos laranja:** Representam os valores previstos pelo modelo de regressão (Random Forest Regressor).

Objetivo do Gráfico: Este gráfico ajuda a visualizar a precisão do modelo de regressão. Quando os pontos laranja estão próximos dos pontos azuis, isso indica que as previsões do modelo estão bem ajustadas aos valores reais. Se houver uma grande discrepância entre eles, isso indica que o modelo não está fazendo boas previsões.

Matriz de Confusão



Descrição: A matriz de confusão é uma ferramenta utilizada para avaliar a performance de um modelo de classificação.

- **Linhas (eixo Y):** Representam as classes reais (Lento ou Rápido).
- **Colunas (eixo X):** Representam as classes previstas pelo modelo.
- **Valores dentro da matriz:** Indicam a quantidade de previsões para cada combinação de classe real vs prevista. Cada célula da matriz mostra o número de exemplos classificados em uma determinada categoria.

Legendas:

- **Verdadeiro Positivo (TP):** Casos em que a classe real foi "Rápido" e o modelo também previu "Rápido".
- **Verdadeiro Negativo (TN):** Casos em que a classe real foi "Lento" e o modelo previu "Lento".
- **Falso Positivo (FP):** Casos em que a classe real foi "Lento", mas o modelo previu "Rápido".

- **Falso Negativo (FN):** Casos em que a classe real foi "Rápido", mas o modelo previu "Lento".

Objetivo do Gráfico: A matriz de confusão ajuda a visualizar o desempenho do modelo de classificação, mostrando onde o modelo está cometendo erros. Para uma boa performance, os valores de TP e TN devem ser altos, enquanto os valores de FP e FN devem ser baixos. Quanto maior o valor em cada célula de "True", melhor o modelo está realizando a tarefa de classificação.