

PROJETO APLICADO II



INFORMAÇÕES DO GRUPO

Nome do projeto:

Análise do comportamento de defeitos apresentados por alternadores automotivos em período de garantia

Integrantes:

Carlos Oliveira

Erik Isidoro

Felipe Ferraz



01

INTRODUÇÃO

Apresentação da empresa

03

O MÉTODO ANALÍTICO



02

DESAFIOS

Descrição do problema e
Proposta analítica

04

MEDIDAS DE ACURÁCIA



The background is a textured, light brown surface. On the left, there are vertical brown lines representing pipes, with orange elbow joints and a yellow T-junction. A small circular gauge with a needle is attached to the yellow T-junction. Several gears of different sizes and colors (orange, grey, yellow) are scattered across the background, some partially visible on the edges.

01

INTRODUÇÃO

Apresentação da empresa

A EMPRESA

Valeo Sistemas Automotivos



A Valeo desenvolve soluções e produtos de alta tecnologia que colocam o Grupo no epicentro das três revoluções que afetam a indústria automotiva atual: eletrificação, veículos autônomos e mobilidade digital.

PRODUTOS

- Alternadores
- Radiadores
- Electroventiladores
- Intercoolers
- Ar-condicionado
- Compressores
- Evaporadores
- Faróis
- Lanternas
- Palhetas
- Atuadores
- Embreagens

VALEO EM NÚMEROS

PRINCIPAIS DADOS



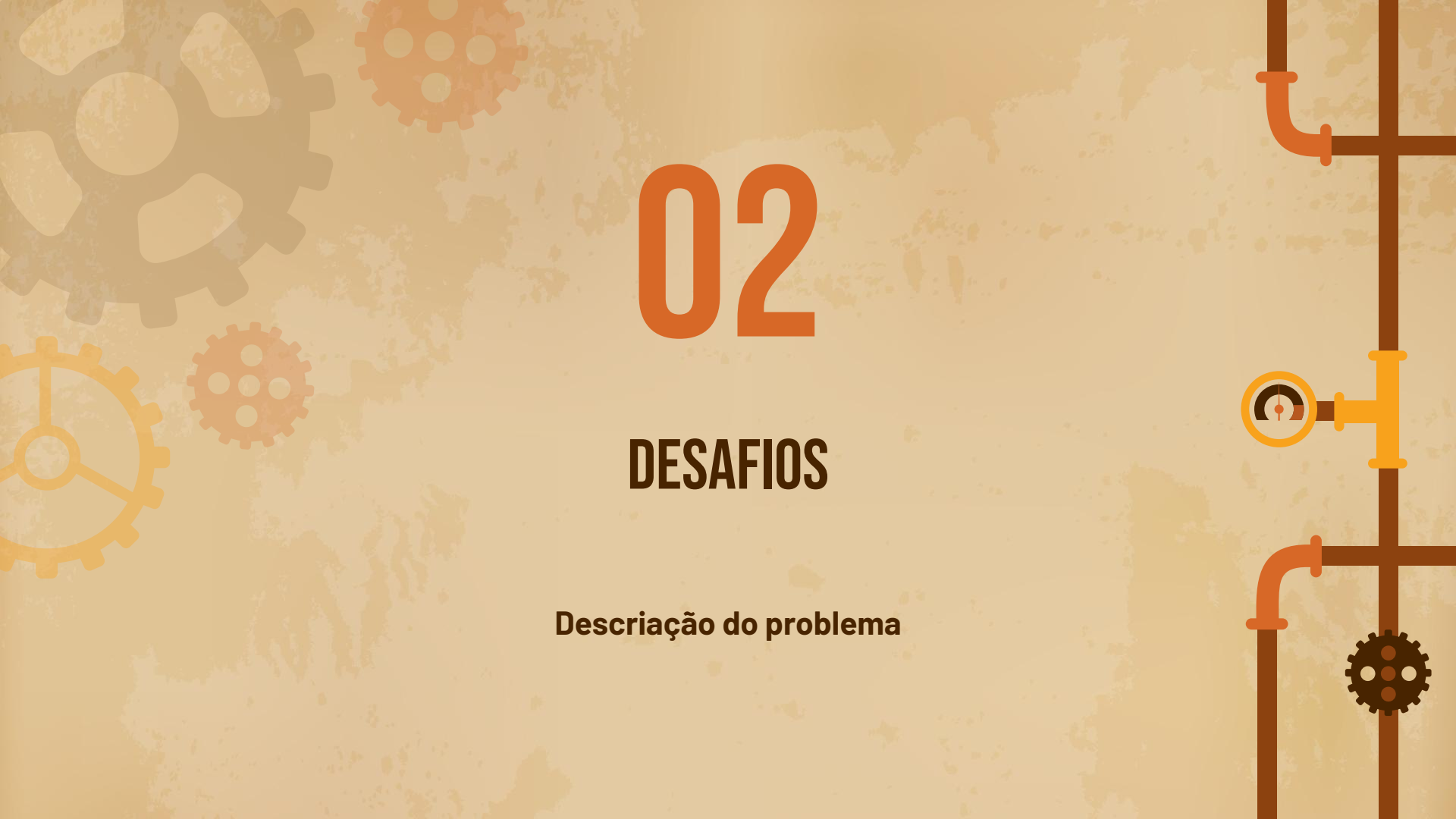
MISSÃO

"A Valeo é uma fornecedora automotiva e parceira de montadoras em todo o mundo. Como empresa de tecnologia, concebemos soluções inovadoras para a mobilidade inteligente, com especial enfoque na condução intuitiva e na redução das emissões de CO2. O Grupo também fornece e distribui peças de reposição para montadoras e operadores independentes de reposição"

02

DESAFIOS

Descrição do problema



DESCRIÇÃO DO PROBLEMA



A garantia das peças da Valeo passa a vigorar junto com a do veículo vendido por uma montadora. Neste período, o defeito apresentado por uma peça é identificado em uma concessionária, no Brasil ou no exterior, e reportado à Valeo. Esses dados provenientes das autorizadas são inseridos no banco de dados da Valeo e utilizados pela área de Qualidade, que por sua vez, não possui iniciativas de Data Science. Diante desse cenário, a Valeo tem a necessidade de analisar o comportamento dos casos de falha em campo de suas peças de forma mais eficiente, para conseguir melhorar a qualidade de atendimento aos seus clientes.

PROPOSTA ANALÍTICA



Dentre as diversas peças produzidas pela empresa, o grupo escolheu trabalhar apenas com os alternadores, pois seria difícil fazer uma análise para todos os produtos. Sendo assim, a análise proposta foi definida com o objetivo de entender o comportamento de defeitos apresentados por alternadores automotivos em período de garantia, utilizando-se para isso de um dataset extraído de um banco de dados da empresa e de um algoritmo construído pelo grupo. Por fim, essa análise poderá servir de modelo para os demais tipos de peças da empresa.

DATASET

Tamanho: 3111 linhas com 21 colunas

Período: 02/2023 até 22/02/2024

Tipos de dados:

- VARCHAR: **cliente (customer)**, **planta do cliente (customer plant)**, **modelo do veículo (vehicle name)**, família do alternador (sub family), sobressalente ou original (spare or original), nr motor (engine nr), país do reparo (repair country), clima (climate);
- INT: pn cliente (customer pn), quilometragem (kms), fator técnico (technical factor), tempo em serviço da peça em meses (mis) e km/ano (km/year);
- DATE: data de produção da peça (production date), início da garantia da peça (inservice date), data do reparo (repair date), data de gravação do cliente (customer recording date), data de carga (upload date);
- FLOAT: custo total de garantia (total warranty cost €), custo de mão de obra (labor cost €), custo de material (material cost €).

Tratamento: Exclusão de dados/linhas inconsistentes (NA) e **anonimização**

03

O MÉTODO ANALÍTICO

Resultados

O MÉTODO ANALÍTICO

Para aplicação das análises, utilizamos um método de aprendizado de máquina com três versões diferentes. No primeiro modelo, aplicamos a análise com o algoritmo de k-vizinhos com $k=3$, o que resultou em uma acurácia de 86%. No segundo modelo, ajustamos o valor de k para 11, o que trouxe uma melhora nos resultados, alcançando uma acurácia de 89%. No terceiro modelo, utilizamos uma árvore de decisão com os três melhores atributos identificados na análise anterior, aplicando uma matriz de confusão e separando a base em conjunto de treino e teste. Configuramos a profundidade máxima (`max_depth`) como indefinida (`None`) e utilizamos o critério de Gini para o cálculo das divisões, o que nos levou a uma acurácia de 100%.

04

MEDIDAS DE ACURÁCIA

Resultados

MEDIDAS DE ACURÁCIA

Esta imagem ilustra o aumento da acurácia conforme os diferentes modelos foram aplicados, mostrando visualmente a diferença entre os modelos KNN com $k=3$, $k=11$ e o modelo de árvore de decisão.

Portanto, o terceiro modelo mostrou-se mais satisfatório, com uma acurácia de 100% para novas previsões, demonstrando um potencial promissor para aplicações práticas.

Socore de acuracidade(2):

1.00

Classification Report(2):

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 1.00 | 1.00 | 1.00 | 18 |
| 1 | 1.00 | 1.00 | 1.00 | 58 |
| 2 | 1.00 | 1.00 | 1.00 | 44 |
| 3 | 1.00 | 1.00 | 1.00 | 814 |
| accuracy | | | 1.00 | 934 |
| macro avg | 1.00 | 1.00 | 1.00 | 934 |
| weighted avg | 1.00 | 1.00 | 1.00 | 934 |

MEDIDAS DE ACURÁCIA

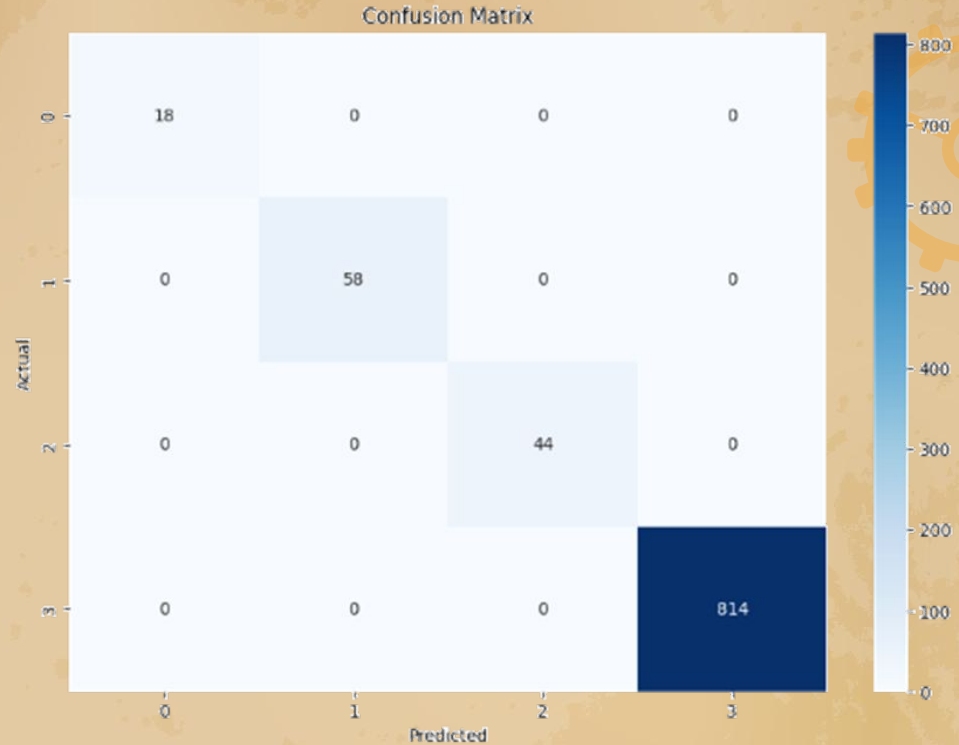
O projeto envolveu uma série de decisões metodológicas para garantir a eficácia dos modelos testados. Para o algoritmo de K-vizinhos mais próximos (KNN), realizamos testes com diferentes valores de k para observar como o ajuste desse parâmetro impactaria a precisão. Optamos por testar $k=3$ e $k=11$, observando que a mudança proporcionou uma melhora na acurácia sem aumentar significativamente o custo computacional. No caso do modelo de árvore de decisão, optamos por limitar o critério ao Gini, permitindo uma classificação robusta e, ao mesmo tempo, configurando a profundidade máxima (`max_depth`) como `None` para capturar a complexidade completa dos dados. Essa configuração mostrou-se essencial para identificar corretamente as classes, considerando as variações de atributos dos dados

ANÁLISE COMPARATIVA DOS MODELOS

O aumento do valor de k no algoritmo KNN de 3 para 11 mostrou-se eficaz para melhorar a acurácia do modelo, uma vez que um k maior reduziu o impacto de ruídos ao considerar um grupo maior de vizinhos para a classificação. No entanto, a superioridade do modelo de árvore de decisão foi evidente, pois ele permitiu uma análise mais estruturada e hierárquica dos dados, resultando em 100% de acurácia. Essa acurácia elevada pode sugerir um ajuste excessivo (overfitting), que pode limitar a capacidade de generalização do modelo para novos dados. Portanto, seria recomendável realizar uma validação cruzada em iterações futuras para avaliar melhor essa possibilidade.

MEDIDAS DE ACURÁCIA

A matriz de confusão do terceiro modelo pode ser inserida aqui para ilustrar visualmente os resultados detalhados do modelo de árvore de decisão.



AVALIAÇÃO E MÉTRICAS

Utilizamos a matriz de confusão para avaliar detalhadamente os resultados do terceiro modelo. A matriz indicou uma baixa incidência de erros de classificação, com valores significativos nos verdadeiros positivos e negativos, o que reforça a precisão do modelo. Essa análise foi essencial para verificar a taxa de erros falsos positivos e negativos, fornecendo insights claros sobre a confiabilidade do modelo em previsões futuras

APLICAÇÃO PRÁTICA E RESULTADOS

Os estudos ressaltam a importância do tratamento do fornecedor MakeD na prevenção de problemas na fabricação de suas peças. A previsão realizada com o modelo indica que, sem ações corretivas, a empresa continuará a registrar problemas significativos, com uma previsão de 5000 novos casos. Esse alto índice sugere que é crucial para a empresa implementar controles de qualidade rigorosos para reduzir a incidência de falhas. A aplicação do modelo de aprendizado de máquina permite, portanto, uma previsão precisa de incidentes futuros e auxilia na tomada de decisões estratégicas para a manutenção da qualidade e satisfação do cliente.

The background is a textured, light beige surface. It is decorated with various mechanical elements: a large grey gear in the top left, a yellow gear in the bottom left, a red gear in the bottom center, and a large yellow gear in the bottom right. A network of brown pipes with yellow and red fittings runs along the left and right edges. A small circular gauge with a red needle is attached to a pipe on the right side.

OBRIGADO!