

Nome: Maiko Douglas Santos Campos
RGM: 27637891

Projeto e Aplicação de Mineração de Dados

O dataset usado para o experimento é o AUTOMOBILE PRICE DATA, se trata de um dataset de Amostra da base do próprio Azure Machine Learning Studio. O projeto consiste no uso do algoritmo de machine learning REGRESSÃO LINEAR, para prever os valores dos automóveis contidos no dataset com base nas informações fornecidas..

Extração, Tratamento e Transformação.

Importação do Dataset.

Dataset de Amostra, contida no Azure ML Studio, AUTOMOBILE PRICE DATA.

Transformação de Dados

Selecionado colunas do dataset para realizar o experimento, removendo apenas a coluna normalized-losses, pois não agrega no resultado do modelo.

Realizado limpeza de campos com informações ausentes, limpando os campos de todas as colunas do dataset.

Divisão do Dataset em duas partes sendo elas em 75% usado para treinar o algoritmo e 25% usado como referência.

Principais Atributos do Dataset.,

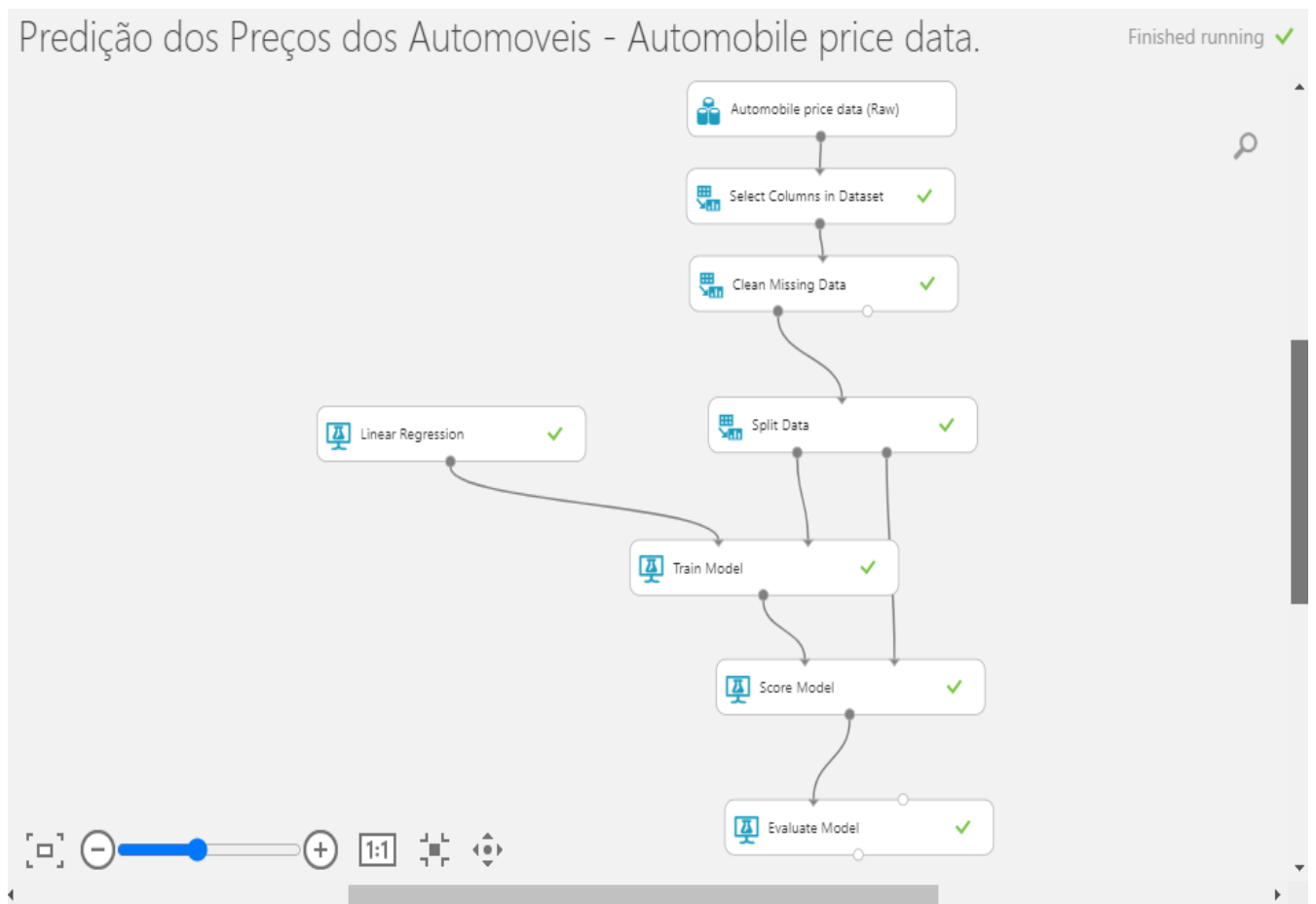
- Tipo de Combustível
- Número de Portas
- Modelo do Carro
- Modelo das Rodas
- Tamanho do Veículo
- Tipo de Motor
- Potência do Veículo
- Cavalo de Potência
- RPM
- Preço

Algoritmo utilizado

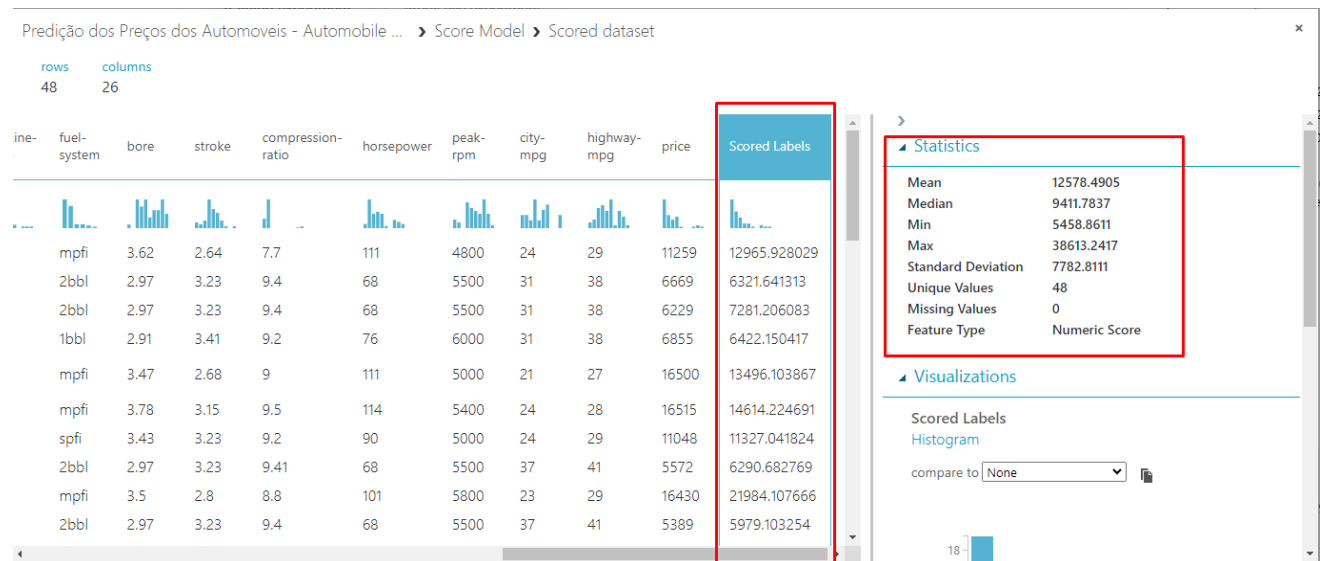
- Tipo: Regressão
- Algoritmo específico: Regressão Linear

Etapas do Processo

- Carregamento do Dataset
- Select Columns Data - Selecionar as colunas do dataset.
- Clean Missing Data - Limpeza dos Dados
- Split Data - Divisão do Dataset
- Train Model - Treinamento do Modelo
- Linear Regression - Uso do algoritmo Regressão Linear
- Score Model - Modelo de Referência
- Evaluate Model - Avaliação do Modelo



Resultado



Valores previstos dos automóveis e as informações estatísticas.

Metrics

Mean Absolute Error	1707.69803
Root Mean Squared Error	2560.753088
Relative Absolute Error	0.285216
Relative Squared Error	0.097337
Coefficient of Determination	0.902663

Métricas.

O Coeficiente de determinação representa o quão próximo o algoritmo acertou a predição dos valores do automóveis contidos na base.Ou seja, 90% de assertividade..

Considerações:

O modelo utiliza dados históricos da tabela de preços dos produtos para prever o valor de novos produtos. Isso é valioso, pois utiliza informações existentes para prever futuros valores.

Previsão de Novos Produtos: O modelo pode ser aplicado não apenas para analisar os preços dos produtos existentes, mas também para estimar o valor de novos produtos que ainda não foram lançados no mercado.

Tomada de Decisão de Negócios: Ao prever o valor de novos produtos, a empresa pode tomar decisões informadas sobre preços de venda, planejamento financeiro e estratégias de marketing, o que é crucial para o sucesso nos negócios.

Conclusão:

Em resumo, o texto sugere que o modelo de previsão de preços gerado pode ser uma ferramenta valiosa para os negócios no ramo de vendas. Ao utilizar dados históricos, o modelo oferece uma oportunidade para prever com mais precisão os valores de novos produtos, auxiliando na tomada de decisões estratégicas relacionadas a preços e posicionamento no mercado. No entanto, a validação contínua e a consideração de outros fatores do mercado são essenciais para maximizar a eficácia desse modelo na prática comercial.