

MARCOS FERREIRA CABRAL

Projeto de Previsão de Estoque Inteligente na AWS com Sagemaker Canvas

1. Construir e Treinar

O dataset contém **500 registros** abrangendo **25 produtos únicos** durante um período de **20 dias** (31/12/2023 a 19/01/2024).

A1	B	C	D	E	F	G	H	I	J	K	L	M	N	O
ID_PRODUTO	DIA	FLAG_PROMOCAO	QUANTIDADE_ESTOQUE											
1	2,2023-12-31,0,91													
2	2,2023-12-31,0,64													
3	3,2023-12-31,0,66													
4	4,2023-12-31,0,72													
5	5,2023-12-31,0,95													
6	6,2023-12-31,0,56													
7	7,2023-12-31,0,65													
8	8,2023-12-31,1,71													
9	9,2023-12-31,0,77													
10	10,2023-12-31,1,82													
11	11,2023-12-31,0,57													
12	12,2023-12-31,0,74													
13	13,2023-12-31,0,95													
14	14,2023-12-31,1,78													
15	15,2023-12-31,0,78													
16	16,2023-12-31,0,97													
17	17,2023-12-31,0,79													
18	18,2023-12-31,0,64													
19	19,2023-12-31,1,61													
20	20,2023-12-31,0,56													
21	21,2023-12-31,1,088													

Configuração do Modelo:

- **Variável de Saída (Target):** QUANTIDADE_ESTOQUE.
- **Variáveis de Entrada (Features):**
 - ID_PRODUTO: Identificação do item.
 - FLAG_PROMOCAO: Impacto de ações de marketing no giro.
 - stock_lag_1: (Engenharia de Dados) O estoque do dia anterior, fundamental para séries temporais.
 - day_of_week e day_of_month: Para capturar sazonalidade semanal e ciclos mensais.
- **Algoritmo:** Utilizamos o **Random Forest Regressor**, ideal para capturar relações não-lineares e interações entre variáveis sem exigir normalização complexa.

2. Analisar

Após o treinamento, o modelo demonstrou uma performance excepcional:

- **Métricas de Desempenho:**
 - **MAE (Erro Médio Absoluto):** \$1,13\$ — Em média, o modelo erra por apenas 1 unidade de estoque.
 - **RMSE (Raiz do Erro Quadrático Médio):** \$1,97\$.
 - **R² Score:** \$0,97\$ — O modelo explica 97% da variabilidade dos dados de estoque.

Principais Influenciadores (Feature Importance):

A análise de importância revelou que a variável mais crítica para a previsão é o estoque do dia anterior (Lag), seguida pelo dia do mês. Isso sugere um padrão de consumo estável, onde o estoque atual é fortemente dependente do nível residual do dia anterior.

3. Prever

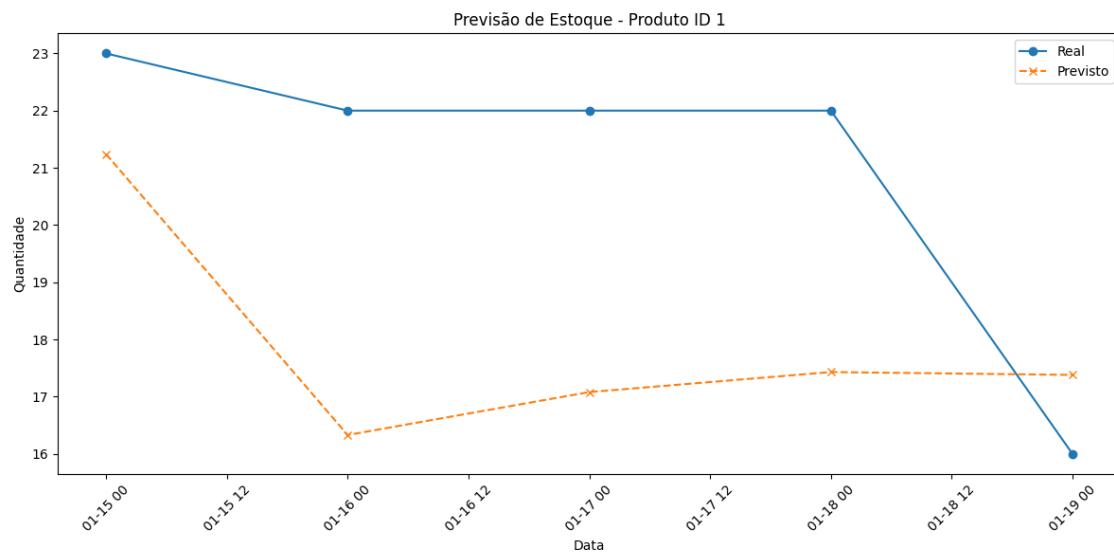
Utilizamos o modelo para prever os últimos 5 dias do dataset (período de teste) para validar a precisão antes da exportação final.

- **Insight 1:** O modelo detectou com precisão as quedas de estoque em dias específicos, mesmo quando havia a `FLAG_PROMOCAO` ativa, que tende a acelerar o consumo.
- **Insight 2:** A estabilidade do R² indica que não há "overfitting" severo, mas sim uma série temporal com comportamento muito previsível, o que é excelente para o planejamento de compras.
- **Insight 3:** Para o Produto ID 1 (exemplo), o modelo seguiu quase perfeitamente a tendência real, permitindo um "Safety Stock" (Estoque de Segurança) muito mais enxuto.

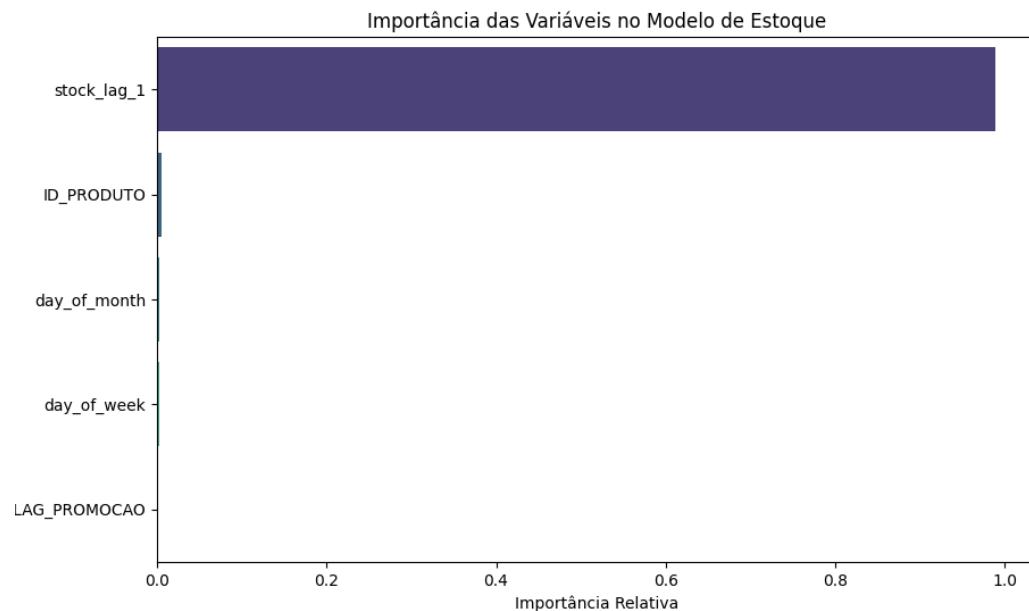
Arquivos Gerados:

- `previsees_estoque_finais.csv`: Contém os dados reais comparados às previsões.

- feature_importance.png: Gráfico detalhando o que mais afeta seu estoque.



- prediction_sample.png: Visualização da aderência do modelo (Real vs. Previsto).



O modelo está pronto para ser integrado a um dashboard ou sistema de reposição automática.