



Challenge ML

Análise e previsão de vendas e descontos

Michel Arruda
05/06/2015

Contextualização

- Vendas de 45 lojas
- Cada loja com vários departamentos e em diferentes regiões
- A rede de lojas organiza eventos promocionais de descontos ao longo do ano geralmente antes de datas comemorativas importantes como o Super Bowl, o Dia do Trabalho, Ação de Graças e o Natal.

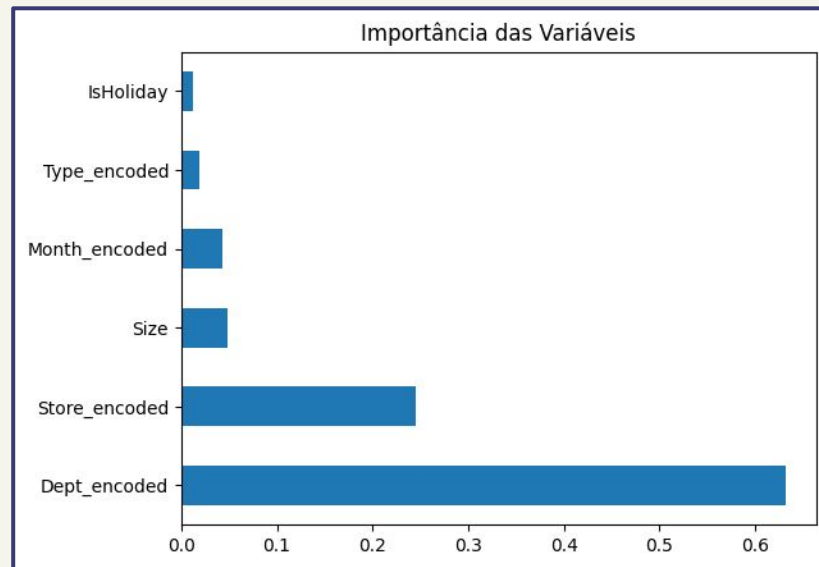
Contextualização

Objetivos:

- Prever as vendas de cada departamento em cada loja para o ano seguinte.
- Propor ações recomendadas com base nos insights obtidos, priorizando aquelas que tenham maior impacto no negócio.
- Modelar os efeitos dos descontos durante as semanas festivas.
- Criar uma API que permita ao sistema da loja consultar, por meio de um endpoint, a previsão de vendas para as próximas quatro semanas.

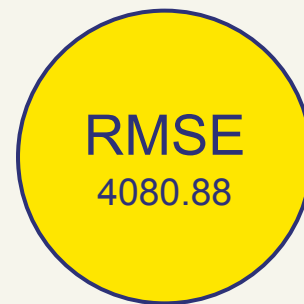
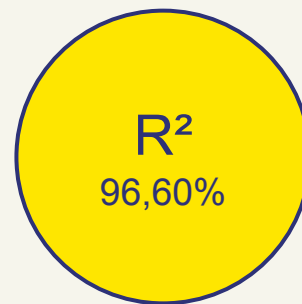
Modelo de previsão de vendas

- Foram testadas algumas técnicas e RandomForest teve a melhor performance.
- Foram selecionadas e utilizadas as colunas:
 - ['Store', 'Dept', 'IsHoliday', 'Type', 'Size', 'Month']
- A variável Dept representa 63,23% da importância para previsão da venda semanal.
- As demais variáveis também apresentaram importância relevante.



Modelo de previsão de vendas

- Avaliação:
 - Treino: dados de 2010 a 2011
 - Teste: 2012
- Resultados de avaliação:
 - O modelo consegue explicar 96,60% das variações nos dados do ano seguinte. (R^2)
 - Indica uma excelente capacidade preditiva
 - O erro médio (RMSE) das previsões é de cerca de 4 mil dólares. Esse valor é relativamente pequeno comparado à escala dos dados e compatível com o alto R^2 observado.



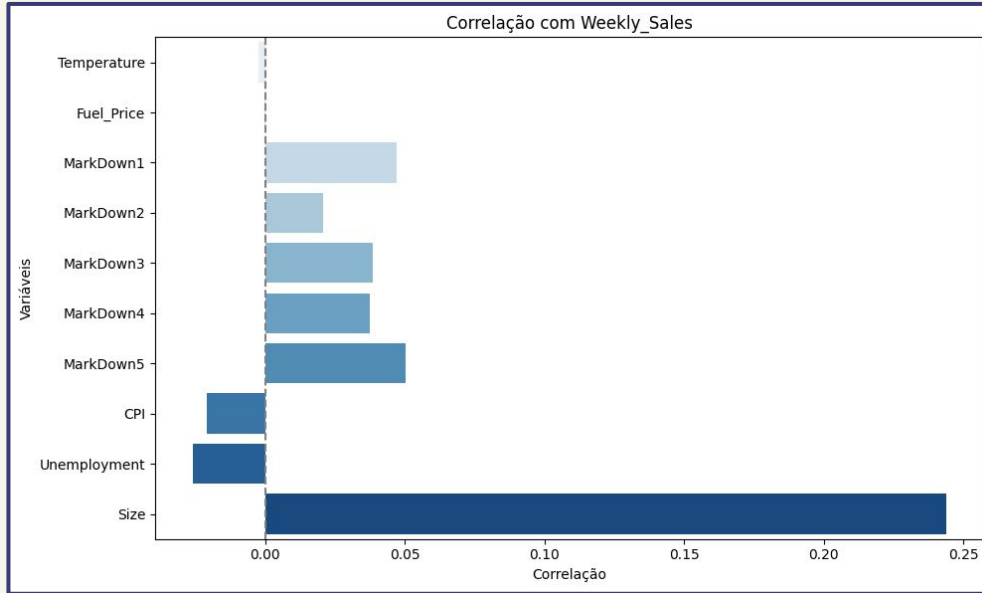
Modelo de previsão de vendas - Previsão 2013

- Foram previstas todas as vendas semanais para cada loja e departamento e salvo no arquivo *'data/previsao_2013.csv'*
- A tabela apresenta um exemplo de previsão do primeiro trimestre para o ano seguinte (2013) para a loja 1 e departamento 1.
- O gráfico apresenta a previsão de vendas para loja 1 no ano de 2013

Store	Dept	Type	Size	Date	IsHoliday	Year	Month	Week	Predicted_Sales
1	1	0	151315	2013-01-04	1	2013	1	1	41963.33
1	1	0	151315	2013-01-11	0	2013	1	2	17286.6475
1	1	0	151315	2013-01-18	0	2013	1	3	17286.6475
1	1	0	151315	2013-01-25	1	2013	1	4	41963.33
1	1	0	151315	2013-02-01	0	2013	2	5	28966.5083333333
1	1	0	151315	2013-02-08	0	2013	2	6	28966.5083333333
1	1	0	151315	2013-02-15	0	2013	2	7	28966.5083333333
1	1	0	151315	2013-02-22	1	2013	2	8	41963.33
1	1	0	151315	2013-03-01	0	2013	3	9	21757.56
1	1	0	151315	2013-03-08	0	2013	3	10	21757.56
1	1	0	151315	2013-03-15	0	2013	3	11	21757.56
1	1	0	151315	2013-03-22	0	2013	3	12	21757.56
1	1	0	151315	2013-03-29	0	2013	3	13	21757.56



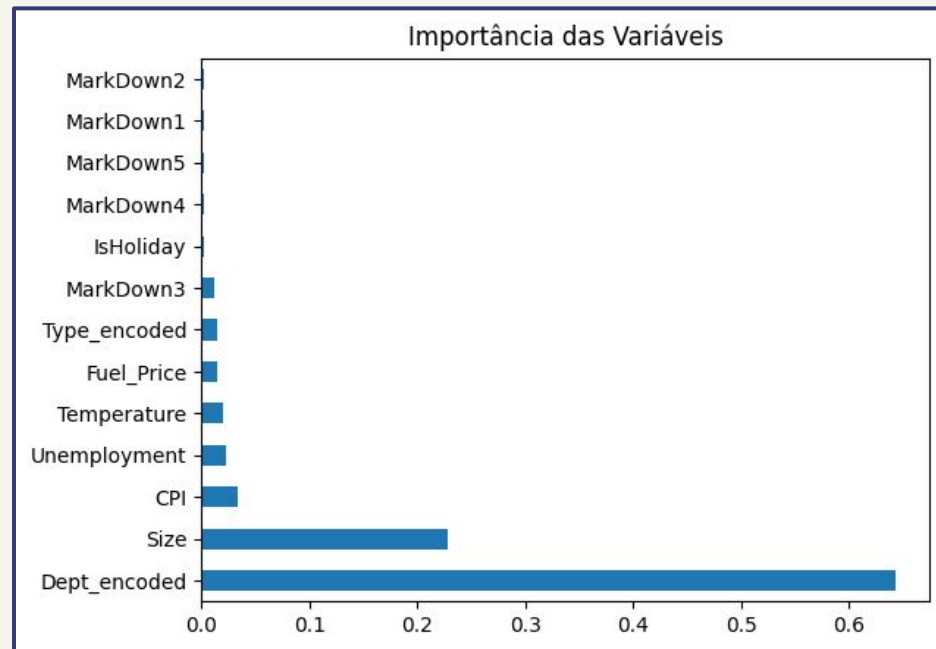
Insights e recomendações - Análise de correlação



Variável	Correlação	Interpretação
Size	+0.24	Lojas maiores tendem a vender mais.
MarkDown5	+0.05	Promoções aumentam levemente vendas.
CPI	-0.02	Vendas caem levemente com aumento de CPI.
Unemployment	-0.02	Desemprego afeta negativamente as vendas.

Insights e recomendações - Importância das variáveis

- Gráfico apresenta a importância das variáveis no modelo.
 - Importância de Markdown 3 → Investir nesse tipo de promoção
 - Size (tamanho da loja) → Focar em lojas maiores para campanhas nacionais.
 - Fuel_Price ou Temperature → Sempre avaliar oscilação para ajustar preço
 - CPI ou Unemployment → Ajustar precificação em tempos de crise.



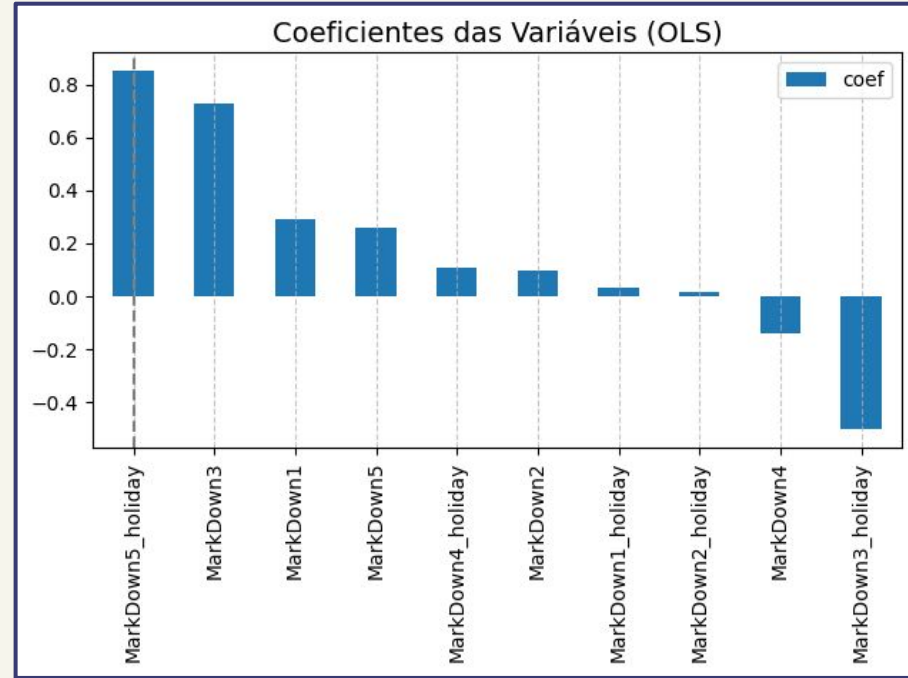
Efeitos de descontos em datas festivas

- Foi treinado um modelo de regressão para analisar os descontos em datas festivas.
- O modelo explica apenas 2,1% da variabilidade das vendas semanais. Isso é muito baixo, indicando que os descontos e feriados capturam apenas uma pequena parte do que afeta as vendas.

Efeitos de descontos em datas festivas

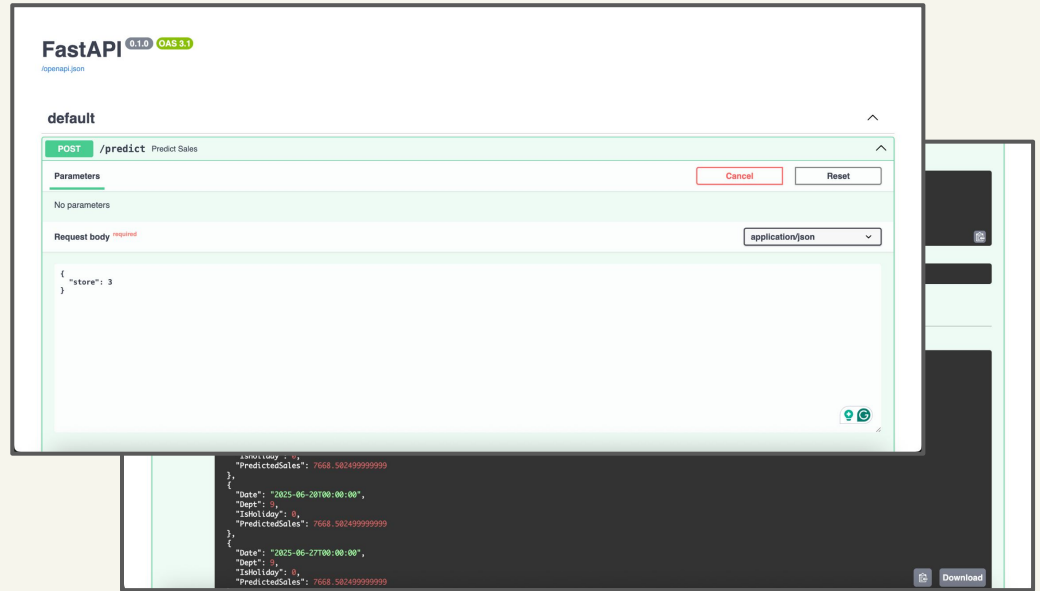
Com relação a descontos:

- Markdown3 é um desconto que impacta muito durante o ano todo, cerca de 73%, e perde força em feriados, cerca de 50%.
- Markdown5 tem o maior ganho durante semanas de feriado, cerca de 85%: estratégia de promoção forte.



API para previsão de vendas

- Previsão de 4 semanas
- Input: Store
- Output lista com:
 - Date
 - Dept
 - IsHoliday
 - PredictedSales



Trabalhos futuros

- Modelo preditivo:
 - Experimentar outras técnicas de machine learning para treinar o modelo
 - Realizar Hyperparameter tuning no melhor modelo para ajustar parâmetros
 - Criar novas features para uso no modelo
 - Explorar outros tipos de encode para variáveis categorias (e uso de redutor de dimensionalidade)
 - Explorar técnicas de análise temporal para previsão e análise do impacto das variáveis
- Análises e insights:
 - Criar novas visualizações afim de buscar novos insights
 - Análises profundamente o impacto dos índices nas vendas (CPI, desemprego, preço do combustível, temperatura)