MODELOS QUANTITATIVOS APLICADOS A PREVISÃO DE DEMANDA DE VAREJO ON-LINE UTILIZANDO ABORDAGENS DE MACHINE LEARNING

autores

Introdução

No varejo on-line a operação demanda ações praticamente em tempo real a fim de compreender e atingir resultados tangíveis para a organização. A evolução recente nos últimos anos e grande disseminação de técnicas de Data Science e Machine Learning permitiram novas aplicações em diversas áreas, e com o varejo não está sendo diferente. Dado o volume e velocidade na geração de informações pela interação constante com os consumidores através de interfaces digitais, o varejo on-line torna-se um campo rico para pesquisas e aplicações de modernas técnicas de análise de dados.

Método

Através do acesso as bases de dados da organização foi obtida o conjunto de dados amostral inicial com cerca de 2 anos de operação histórico do site de vendas, inicialmente trabalhando com variáveis de volumes de vendas diretas ao consumidor (B2C) e vendas através de terceiros (MarketPlace). É possível, então, analisar o comportamento diferenciado para cada tipo de operação.

A organização possui atualmente uma meta de vendas que foi estipulada através de cálculos aplicados em planilhas, buscando abordagens usuais como médias e observações de tendências. A tarefa passou a ser a avaliação do erro de previsão atual, buscando entender os diversos comportamentos de vendas afetados por eventos exógenos, tais como Black Friday, feriados e sazonalidades semanais, greves de caminhoneiros entre outros.

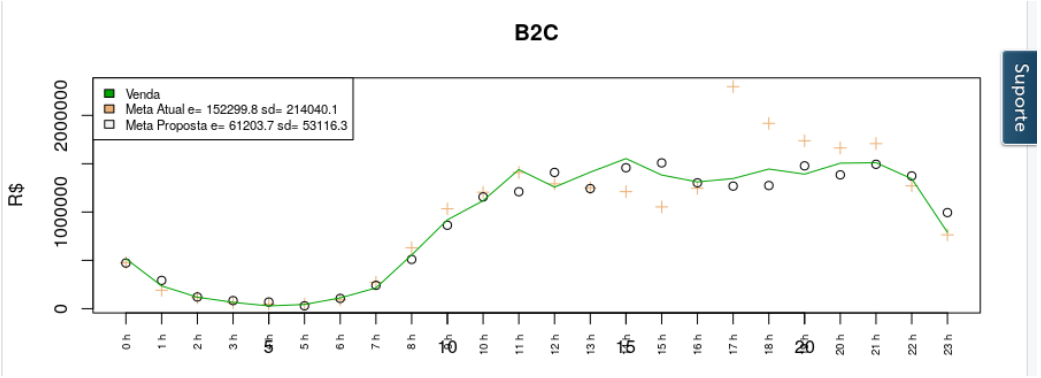
Realizado o tratamento e enriquecimento de dados originais, foram aplicadas técnicas tradicionais de análises de séries temporais, com abordagens de modelos auto regressivos. Posteriormente foram aplicadas técnicas de machine learning (algoritmos) referenciadas como adequadas para buscar resultados melhores em modelos preditivos de séries temporais. A utilização do SVM (support vector machines) apresentou resultados adequados.

Dados Coletados

Podemos observar na figura abaixo 3 séries de dados:

1. A linha sólida demonstrando a venda em um determinado dia

2. A meta projetada atualmente através de planilhas, representada por um marcador “+”
3. A meta proposta, neste caso utilizando o algoritmo SVM, através do marcador “o”



Além da visualização gráfica, procuramos gerar uma métrica de comparação para avaliação dos modelos atuais e propostos. Utilizaram-se duas métricas:

1. Erro de previsão: diz respeito a média das diferenças entre valor real e valor previsto, representado no gráfico acima como e
2. Desvio padrão das diferenças: a operação do negócio diz respeito a tomar ações quando ocorrem variações não esperadas ocorrem. Sendo assim, busca-se na meta uma maior linearidade nas diferenças. Por isto, utilizou-se inicialmente o cálculo do desvio padrão, demonstrando que quanto menor, melhor será a previsão.

Resultados obtidos até o momento

Até o momento os resultados obtidos através do modelo auto regressivo demonstram uma melhoria de cerca de 24% em relação ao erro de previsão utilizando-se o cálculo atual. Ao se utilizar o SVM atingimos 41% de redução de erro de previsão. Os desvios padrão demonstraram uma redução equivalente.

Conclusões

Os resultados obtidos até o momento nos demonstram que temos espaço para um maior refinamento e continuidade do estudo para aprimoram o modelo. Algumas técnicas que podem ser aplicadas:

- Janelas deslizantes no modelo autoregressivo

- Aplicar a variação com média móvel (ARIMA)
- Aplicar a técnica de tunneling no SVM (similar a janelas deslizantes)
- Utilizar algoritmos de tunning para o SVM
- Agregar mais variáveis exógenas (clima, fatores econômicos, entre outros).