

Projeto Modelo Preditivo de Suprimentos

BRF

Organismo Embalagens

Entendimento do negócio, dos dados e recomendações para a solução

5 de março de 2021

Erick Damasceno

Qubo | Consultor em Estatística e Inteligência Artificial

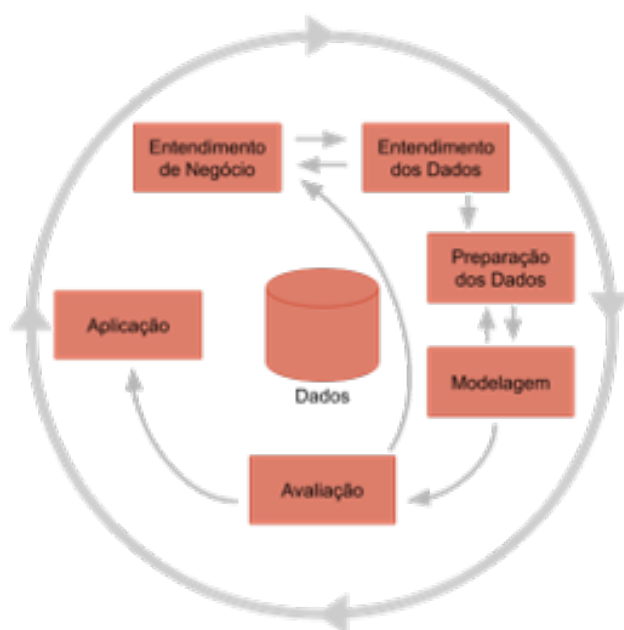
erick@qubo.uno

+55 (61) 998 113 154

1. Contexto e expectativas

1.1. Contexto

Esse relatório consolida a etapa de entendimento do negócio e dos dados após rodadas de reuniões com a área de Embalagens. O documento servirá de guia para o desenvolvimento das etapas de preparação dos dados (ETL) e, posteriormente, desenvolvimento dos organismos.



1.2. Pessoas envolvidas

Para se chegar ao presente entendimento de negócio, diagnóstico de dados e recomendações, foram levadas em consideração as informações de negócios e dados compartilhados entre as seguintes pessoas:

| Organização | Nome | Área ou cargo |
|-------------|------------------|------------------|
| BRF | Bruno Marques | CIEX Suprimentos |
| BRF | Karina Macagnam | CIEX Suprimentos |
| BRF | Willians Almeida | Embalagens |
| BRF | Rafael Silva | Gestão TI |

| | | |
|------|-----------------|--------------------|
| Qubo | Erick Damasceno | Consultor |
| Qubo | Eugênio Pessoa | Cientista de dados |
| Qubo | Pedro Drummond | Gestão do Programa |
| Qubo | Rafael Ferreira | CTO |
| Qubo | Rodrigo Paiva | Comercial |

1.3. Alinhamento de expectativas

Por se tratar de um projeto com foco em inteligência artificial (aprofundada em modelagem preditiva, nesse caso) é importante alinhar as expectativas das pessoas envolvidas sobre o que a inteligência artificial “**É**” e o que ela “**NÃO É**” e sobre o que ela “**FAZ BEM**” e o que ela “**NÃO FAZ TÃO BEM**”.

A inteligência artificial (IA) é uma área multidisciplinar, concentrada em tópicos computacionais, matemáticos e estatísticos, porém não está limitada a eles. Ela não é uma técnica em si, mas um conjunto de técnicas que evoluem com o tempo. Atualmente, técnicas como Aprendizagem de Máquina (*Machine Learning*), Visão Computacional, Processamento de Linguagem Natural (*NLP*) dominam a geração atual dessas soluções. Entretanto, não invalidam outras técnicas de gerações anteriores como Reconhecimento de Texto ou Automações de Tarefas, por exemplo. Dessa forma, IA pode representar um conjunto extenso de técnicas que podem inclusive ser combinadas entre si.

Por outro lado, a IA não é uma ferramenta mágica ou replicável em qualquer contexto. Existem limites técnicos para o que essas soluções podem fazer bem. Para tentar desenvolver uma intuição de quando uma tarefa pode ser ensinada com sucesso para uma máquina, considere a seguinte regra *imperfeita*:

Qualquer coisa que você pode fazer com 1 segundo de raciocínio provavelmente pode ser ensinada para uma máquina.

Por exemplo, ao olhar para um e-mail qualquer, você consegue perceber em aproximadamente 1 segundo que se trata ou não de um spam. Ao olhar para uma foto de um animal, você consegue dizer que animal é aquele instantaneamente.

Já no outro lado do espectro, existem tarefas que você geralmente precisa investir uma grande quantidade de tempo e raciocínio, como realizar uma pesquisa para cura do câncer ou escrever um romance. Tarefas nesse lado do espectro tendem a ser incrivelmente mais difíceis de serem ensinadas para uma máquina.

Outra variável importante para incorporar nessa intuição é a quantidade de informação disponível sobre a tarefa de interesse. Os algoritmos de aprendizagem de máquina precisam consumir centenas, milhares ou até milhões de exemplos (dados) para aprender a fazer boas previsões, previsões e recomendações. Além disso, esse aprendizado inicial pode não ser suficiente para o funcionamento perpétuo da solução, sendo necessário continuar alimentando os algoritmos para que esses estejam sempre calibrados.

Com essas limitações, é importante ter em mente que nem sempre será viável construir uma solução utilizando inteligência artificial. Nesses casos, caminhos alternativos poderão ser sugeridos para minimizar as dores da equipe de Embalagens.

2. Problema identificado

2.1. Declaração do problema

Os produtos da BRF são embalados por plásticos de diferentes formatos e densidades, feitos sob medida por seus fornecedores para acomodar cada linha de produtos de suas marcas. Entretanto, os preços dessas embalagens oscilam bastante devido aos custos das principais resinas envolvidas no processo de fabricação. Então, para entender a formação de preços das embalagens finais é necessário compreender a formação de preços das suas principais matérias-primas: **polietileno e polipropileno**.

O Brasil produz as duas resinas, porém a Braskem (principal produtora) detém praticamente um monopólio do mercado nacional. Tipicamente, seus preços obedecem ao mercado internacional, direcionados pelo mercado dos EUA. Então, embora haja uma paridade com os preços de outros fornecedores externos, existem algumas barreiras alfandegárias que dificultam a compra dessas resinas de outros países, sem contar a longa duração logística da operação de importação.

Atualmente, a equipe de compras de embalagens analisa o mercado olhando diversos fatores como disponibilidade de resinas, reservas de petróleo, custos de gás natural, entre outros. Todas essas informações encontram-se espalhadas em diversas fontes, tornando a análise em si uma “colcha de retalhos”. Essa análise é feita com 2 ou 3 semanas de antecedência para tentar

prever o que vai acontecer com os preços de mercado no Brasil, servindo assim como base para o seu processo de planejamento de compras. Caso consigam antever um aumento de preços, compra-se um volume maior com antecedência. Em caso de previsão de redução de preços, a equipe pode aguardar alguns dias para realizar uma compra melhor.

Por esses motivos, a equipe está interessada idealmente em:

“Saber os custos das resinas no mercado futuro”.

Mais especificamente, saber os preços de:

- Polietileno de Alta densidade (PEAD)
- Polietileno de Baixa densidade (PEBD)
- Polietileno de Baixa densidade linear (PEBDL)
- Polipropileno (PP)

2.2. Informações adicionais

- Até antes da pandemia de COVID-19, os contratos de compras de embalagens tinham vigência entre um ano e um ano e meio. Com a pandemia, as incertezas do mercado afetaram os preços, forçando uma redução da vigência desses contratos para 3 ou 4 meses;
- A cadeia produtiva nacional opera a base de Nafta (petróleo), então oscilações de preço de petróleo afetam diretamente o preço das resinas.
- Alguns eventos extremos como explosão de fábricas, fechamento de fábricas e falta de estoque podem fazer os preços subirem de 30% a 40%.

3. Bases de dados e outras fontes de informação

Diversas bases de dados foram mapeadas para a formulação da solução. Essas bases fornecerão elementos de natureza quantitativa para alimentar *dashboards* e modelos preditivos. Outras informações de natureza qualitativa também serão incorporadas na solução para ajudar a contornar as limitações dos modelos.

3.1. Dados do SAP (contratos de compras de embalagens)

→ Dados devem ser disponibilizados pela BRF

O SAP da BRF possui registros sobre os preços pagos nas embalagens dos produtos. Esse histórico de preços poderá ser utilizado para avaliar as diferenças em relação aos valores das resinas plásticas.

Esses dados compreendem informações de:

- Data da compra
- Material comprado (qual embalagem)
- Quantidade total
- Duração do contrato (vigência)
- Valor total da compra (não confundir com gastos mensais)
- Moeda de negociação (real, euro, dólar, rmb, etc)
- Fornecedor
- País do fornecedor

Filtros para seleção:

Solicitar filtros específicos para a equipe de compra de embalagens para que a extração seja feita no SAP.

3.2. ICIS

→ Dados devem ser disponibilizados pela BRF

Acompanhamento de preços de mercado das commodities de plástico.

- Disponibiliza API;
- Licença paga, mas a BRF possui a licença.
- Além dos dados quantitativos, os dados de insights da ICIS também podem ser exportados.

URL de referência:

<https://www.icis.com/explore/commodities/chemicals/plastics/>

3.3. BR Investing

→ Dados gratuitos disponibilizados online

Informações complementares sobre o mercado. Especificamente:

- Custo Petróleo WTI
- Custo Gás Natural EUA
- Custo Petróleo Brent

A BR Investing não fornece uma API, portanto os dados deverão ser extraídos via robôs (crawlers/scrapers).

URL de referência:

<https://br.investing.com/>

3.4. OPEC

→ **Dados gratuitos disponibilizados online**

Informações complementares sobre o mercado.

- Disponibilidade de barris de petróleo mundial
- Relatórios mensais (Monthly Oil Market Report) em Excel
- Esses relatórios disponibilizam dados do Brasil também

URL de referência:

https://www.opec.org/opec_web/en/publications/338.htm

Ver opções na coluna da direita para selecionar o ano das publicações (PDFs e planilhas em Excel).

3.5. EIA (Energy Information Administration - USA)

→ **Dados gratuitos disponibilizados online**

Informações complementares sobre o mercado.

- Dados sobre petróleo e gás natural
- Preços, estoques, produção, consumo, vendas
- Útil para benchmarking internacional
- **Disponibiliza API.**

URLs de referência:

<https://www.eia.gov/opendata/qb.php?category=371>

Existe uma vasta gama de possibilidades. Navegar pelas categorias junto com a equipe de compras de embalagens para refinar quais os dados específicos serão usados.

3.6. census.gov (U.S. Exports/Imports of Merchandise)

→ Dados gratuitos disponibilizados online

Informações complementares sobre o mercado.

- Preços registrados nas atividades portuárias de importações e exportações

URLs de referência:

- <https://www.census.gov/foreign-trade/reference/products/catalog/ftdproducts.html#Port>
- <https://www.census.gov/data/developers/updates/new-discovery-tool.html>
- <https://www.census.gov/foreign-trade/schedules/b/2021/c39.html>

3.7. Referências qualitativas

→ Notícias, Releases, etc., de fontes online gratuitas

→ A BRF será consultada posteriormente sobre os links dos itens específicos que serão considerados na solução final.

Compilação de fontes de notícias sobre:

- *Compliance*;
- Incrementos de capacidade produtiva (Release Petroquímicas / ABIEF e Mercado em geral);
- Alterações fiscais;
- Manutenções e paradas petroquímicas;
- Fatores ambientais;

- Disponibilidade nacional de resinas (Braskem);
- Informações de clima (ex.: “furacões”).

4. Solução proposta

4.1. Painel web com dados quantitativos

O painel quantitativo fornecerá uma compilação dos dados já utilizados pela equipe de compras de embalagens e que atualmente encontram-se espalhados em diversas fontes. Além disso, trará uma expansão de funcionalidades com a inclusão de modelos estatísticos de previsões de séries temporais, análises detalhadas das relações entre as variáveis disponíveis e uma calculadora preditiva para simular cenários. De modo detalhado, o painel contemplará:

- Compilação das principais séries históricas que já são consultadas, mas estão espalhadas nas diversas fontes de dados mapeadas nesse relatório;
- Previsões automáticas para o mercado futuro (curto prazo).
- Calculadora preditora para mercados futuros (inserir dados de diferentes cenários e estimar o resultado para cada um). Alguns cenários podem ser pré-definidos para já gerar automaticamente algumas respostas;
- Marcos históricos nos gráficos para auxiliar na interpretação dos eventos;
- Comparação dos preços pagos pela BRF nas embalagens com os preços das resinas.

4.2. Painel web com informações qualitativas

- Extração automática de notícias de sites selecionados e relatórios em PDF.

4.3. Direcionamento estatístico

Mineração de dados

- Explorar relacionamento entre os dados coletados para tentar encontrar padrões relevantes.

Modelos clássicos de previsão de séries temporais

- Possível para dados de todas as bases quantitativas mapeadas.
- Permite estimar o custo das resinas no mercado futuro

Decomposição gráfica das séries em tendências e sazonalidades

- Possível para dados de todas as bases quantitativas mapeadas.
- Permite entender o comportamento histórico das séries e ter insights para inserir nos modelos futuros.

Cointegração de séries

- Tentar encontrar uma outra série que ajude a prever a tendência dos preços. Precisa testar entre potenciais séries candidatas.

Modelos multivariados

- Criar modelos preditivos multivariados para tentar prever o valor das resinas com base em outras variáveis.
- Levar em consideração que os modelos lineares tradicionais vão sofrer com a autocorrelação. Precisa tirar a diferença de 1ª ordem das séries antes dessa modelagem.
- Testar metodologia de “diferenças-em-diferenças”.