

# Nanodegree Engenheiro de Machine Learning

## Proposta de projeto final

---

### Smart BackOrders

#### Análise e Predição de Pedidos Pendentes (BackOrders) utilizando Modelos de Inteligência Artificial

João Luiz de Souza Torres  
Novembro de 2018

---

#### Pedidos Pendentes (BackOrders)

BackOrders de produtos é um problema comum da cadeia de suprimentos, sendo uma preocupação importante na Gestão de Inventário de Produtos.

Definição de BackOrder: Pedido ou compromisso não efetivado ou cumprido. Um pedido pendente é uma demanda imediata de determinado item cujo estoque é insuficiente para satisfazer tal demanda.

Existem diversos fatores que contribuem para a ocorrência de Backorders. Eis alguns:

- Falha de gestão
- Falta de comunicação
- Incapacidade de fabricar itens suficientes para atender a demanda
- Falta de dados para prever com precisão a demanda
- Ausência de sistema informatizado que auxilie
- Estratégia da empresa.

O problema é amplamente estudado e debatido, dentro e fora do Brasil, visando a criação de soluções que possam minimizá-lo. [1] [2] [3] [6]

O assunto também é bastante requisitado em anúncios de vagas de trabalho relacionados à área de logística. [4] [5]

Motivação pessoal: atualmente exerço atividades profissionais como analista de desenvolvimento de software em uma empresa Multinacional Brasileira. Estou alocado em um Sistema de Software para Atacadista Distribuidor e componho uma equipe responsável pela criação de novas *features* bem como também pela manutenção das *features* existentes. Neste contexto (Atacado Distribuidor), é de alta relevância a utilização de mecanismos de análise de pedidos pendentes (Backorders) visando atuar na mitigação do problema, melhorando assim o controle de estoque de materiais diversos, bem como a otimização das compras e do atendimento de pedidos. Tais medidas impactam diretamente na redução de custos e no consequente aumento de satisfação dos clientes. Assim, o projeto Smart Backorders é bastante alinhado à minha atividade de trabalho, já tendo sido apresentado à empresa como uma ideia inicial de solução, passível de ser agregada ao software da empresa.

## Descrição do problema

Contextualização:

<b>Conceito de BackOrders</b>	Produtos que estão sem estoque temporariamente (*), mas é permitido ao cliente gerar um pedido a ser atendido com estoque futuro, ou ainda, venda com estoque atual negativo.
<b>(*) % BackOrders</b>	Importante KPI ( <i>Key Performance Indicator</i> ) na Logística.
<b>Dilema</b>	Forte demanda <i>versus</i> planejamento de estoque subotimizado.
<b>Algumas Questões</b>	O cliente pode esperar? (pedidos em atraso, pedidos pendentes no fornecedor, perda de vendas, baixa no índice de satisfação do cliente (NPS), dentre outros fatores).

Áreas de Negócio onde é relevante:

<b>Produção</b>	Programação, controle de estoque, planejamento agregado.
<b>Comercial e Marketing</b>	Alocação da força de vendas, promoções, lançamento de novos produtos.
<b>Finanças</b>	Investimento na fábrica, em equipamentos, planejamento orçamentário.
<b>Pessoal</b>	Planejamento de mão de obra, contratações, demissões.

Alguns números sobre Backorders no Brasil:



### "Análise de rupturas na cadeia farmacêutica"

"A análise dos mais de 3.000.000 de pedidos das farmácias para os distribuidores mostrou que 14% dos pedidos apresentavam algum tipo de ruptura, fazendo com que 8,2% do valor do pedido não fosse faturado, gerando com isso uma perda anual no faturamento na ordem de R\$ 2 bilhões." - Pesquisa ILOS - Aché, ABRAFARMA, BD, Fidelize e Plannera (2015)).

Percepção do risco da ocorrência de Backorders nas operações:

Tabela 1 - Risco total da probabilidade – Gerentes versus funcionários

Fatores de risco	Probabilidade de provocar o desabastecimento (%)						Risco total
	Gerentes			Funcionários			
	Baixa	Média	Alta	Baixa	Média	Alta	
Back order	0,0	14,3	85,7	0,0	11,1	88,9	38,6
Transporte	28,6	14,3	57,1	5,6	19,4	75,0	22,6
Logística	0,0	14,3	85,7	0,0	22,2	77,8	34,4
Tempo de entrega	14,3	57,1	28,6	33,3	38,9	27,8	12,2
Sistema fora do ar	71,4	28,6	0,0	88,9	11,1	0,0	11,6
Fatores totais percebidos	22,9	25,7	51,4	25,6	20,6	53,9	16,6

Tabela 2 - Risco total do impacto – Gerentes versus funcionários

Fatores de risco	Impacto ou dano quando ocorre o desabastecimento (%)						Risco total
	Gerentes			Funcionários			
	Baixo	Médio	Alto	Baixo	Médio	Alto	
Back order	0,0	14,3	85,7	0,0	11,1	88,9	38,6
Transporte	42,9	14,3	42,8	5,6	25,0	69,4	16,5
Logística	0,0	28,6	71,4	2,8	11,1	86,1	31,8
Tempo de entrega	57,1	14,3	28,6	44,4	25,0	30,6	9,8
Sistema fora do ar	57,1	42,9	0,0	91,7	8,3	0,0	9,9
Fatores totais percebidos	31,4	22,9	45,7	28,9	16,1	55,0	15,3

"Inovando a gestão do risco de ruptura na cadeia logística automotiva"

Anais do V SINGEP - São Paulo - SP - Brasil - 20, 21 e 22/11/2016

VI Simpósio Internacional de Gestão de Projetos, Inovação e Sustentabilidade (VI SINGEP)

## Método Reativo: Exemplo de Sistema de Registro de BackOrder visando Sugestão de Compras posterior

### Sugestão de Compras Backorder

1. A funcionalidade backorder tem como objetivo fazer uma sugestão de compras para peças que estão em pedido e não possuem estoque. O diferencial desta sugestão é que as peças serão reservadas automaticamente no momento em que forem lançadas no estoque, para garantir que sejam utilizadas, exclusivamente, no pedido que deu origem à sugestão.
2. Para utilizar esta funcionalidade insira um novo pedido com uma peça sem estoque.
3. Clique no botão Confirmar para finalizar a inclusão do pedido.
4. Em seguida, altere o mesmo pedido, pois as opções de Backorder só estão disponíveis na alteração.
5. Acesse a opção de **sugestão de compras backorder** em Ações Relacionadas\Sug. Compras BO.
6. Selecione a forma de pedido referente ao backorder para a sugestão e confirme.

Como a Inteligência Artificial pode auxiliar na resolução do problema?

A partir dos fatos já ocorridos, utilizando os dados disponíveis, aplicar modelos e algoritmos que possam estabelecer uma “previsão” sobre a probabilidade de nova ocorrência.

A proposta deste trabalho é contribuir com a análise preditiva de Backorders, apresentando-a como uma ferramenta auxiliar ao processo de tomada de decisão nas empresas.

Visa especialmente, uma mudança de postura: REATIVA x PRÓ-ATIVA, atuando diretamente nos índices de satisfação de clientes, com a consequente melhoria nas vendas.

Trabalhando para identificar peças em risco de atraso antes do evento, então o negócio tem tempo para reagir.

Este é o objetivo do projeto “Smart BackOrders: Análise e Predição de Pedidos Pendentes (BackOrders) utilizando Modelos de Inteligência Artificial”.

## Conjuntos de dados e entradas

DESAFIO: “Can You Predict Product Backorders?”(desafio Kaggle)

Kaggle: É uma plataforma para aprender estatísticas, aprendizado de máquina e outros conceitos de ciência de dados, com um forte foco na aplicação prática dessas habilidades através de desafios em um ambiente competitivo. (foi adquirida pelo Google em Março/2017)  
[www.kaggle.com](http://www.kaggle.com)

Desafio Kaggle : “Can You Predict Product Backorders?” (disponibilizado em Abril / 2017)

[www.kaggle.com/tiredgeek/predict-bo-trial](http://www.kaggle.com/tiredgeek/predict-bo-trial) (\*)

O arquivo de dados (Dataset) de treinamento contém os dados históricos para as 8 semanas anteriores à semana que estamos tentando prever.

Os dados disponibilizados no Dataset foram coletados como instantâneos semanais no início de cada semana.

Dataset de Treinamento: “Kaggle\_Training\_Dataset.csv” (1.6 milhões de registros).

Dataset de Testes/Validação: “Kaggle\_Test\_Dataset.csv” (241.8 mil registros).

sku	chave de identificação de produto (Random ID for the product)
min_bank	quantidade de estoque mínima recomendada (Minimum recommend amount to stock)
potential_issue	problema de origem para parte identificada (Source issue for part identified)
pieces_past_due	partes atrasadas da fonte (Parts overdue from source)
perf_6_month_avg	desempenho da fonte para o período anterior de 6 meses (Source performance for prior 6 month period)
perf_12_month_avg	desempenho da fonte para o período anterior de 12 meses (Source performance for prior 12 month period)
local_bo_qty	quantidade de pedidos atrasados (Amount of stock orders overdue)
deck_risk	indicativo de risco (Part risk flag)

<b>oe_constraint</b>	<b>indicativo de risco (Part risk flag)</b>
<b>ppap_risk</b>	<b>indicativo de risco (Part risk flag)</b>
<b>stop_auto_buy</b>	<b>indicativo de risco (paralisar compra automática) (Part risk flag)</b>
<b>rev_stop</b>	<b>indicativo de risco (Part risk flag)</b>
<b>went_on_backorder</b>	<b>o produto entrou em “backorder”. Este é o valor alvo. (Product actually went on backorder. This is the target value.)</b>

(\*) Obs: visando preservar os dados em caso de mudanças na estrutura do desafio, armazenada na plataforma Kaggle, os dados foram movidos e estão disponíveis no repositório Github do projeto:

## Descrição da solução

Para a construção solução utilizaremos algoritmos de aprendizado supervisionado em machine learning, visando comparar seus desempenhos na realização da tarefa.

Visando também ampliar a comparação, adicionaremos modelos simples de redes neurais artificiais.

Modelos utilizados:

- Decision Trees (Decision Tree Classifier)
- Random Forests (Random Forest Classifier)
- Gradient Tree Boosting (Gradient Boosting Classifier),
- Bagging meta-estimator (Bagging Classifier)
- Under-sampling (Random Under Sampler Classifier)
- MultiLayer Perceptron (Rede Neural Multicamadas) (MLP Classifier)

Na implementação serão utilizados os pacotes:

- scikit-learn - Machine Learning in Python [14]
- imbalanced-learn - SMOTE, under-sampling and over-sampling methods [15]

## Modelos de referência (benchmark)

Embora o desafio tenha sido retirado da plataforma Kaggle (sem justificativas), algumas submissões ainda podem ser encontradas. [7] [8]

Em pesquisa, localizamos também alguns trabalhos realizados, dentro e fora do Brasil, que utilizam a mesma base de dados, em trabalhos acadêmicos ou artigos em sites especializados. [9] [10] [11] [12] [13]

Os trabalhos encontrados nos possibilitam conhecer e comparar outras abordagens, tecnologias, estratégias, linguagens de programação e frameworks, enriquecendo em muito a nossa experiência.

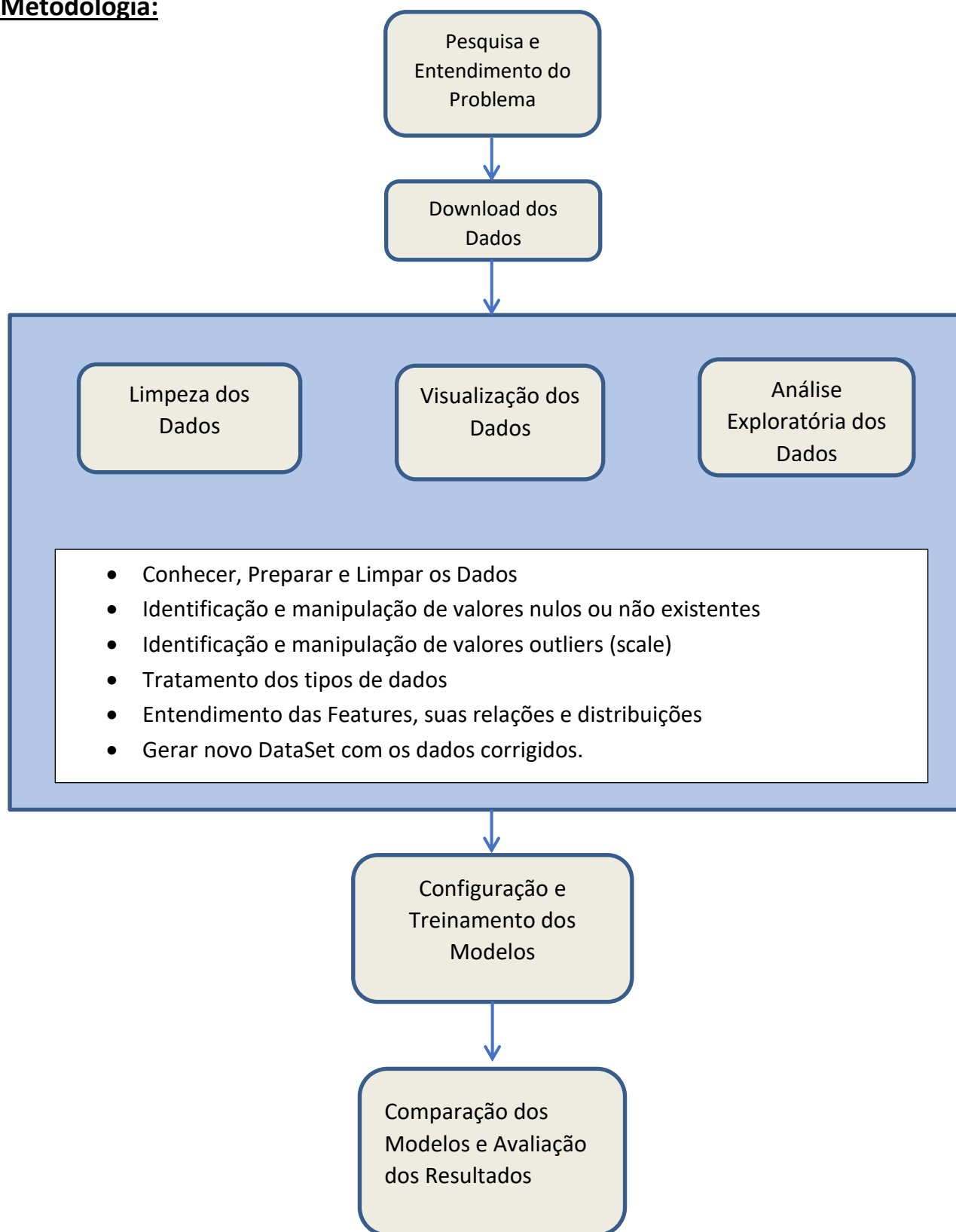
## Métricas de avaliação

Neste trabalho estamos optando por utilizar como métricas, o mesmo padrão encontrado nos desafios Kaggle:

- AUC - Area Under the ROC Curve - Receiver Operating Characteristic (ROC Curve)
- F1 Score, calculado a partir das medidas Precision e Recall

## Design do projeto

### Metodologia:





## Referências

- [1] BackOrder - Dicionário de Logística On-line  
Revista LOGÍSTICA & SUPPLY CHAIN  
<https://www.imam.com.br/logistica/dicionario-da-logistica/?pag=13&a=P>
- [2] LogWeb - Indicadores de atendimento  
<http://www.logweb.com.br/colunas/indicadores-de-atendimento/>
- [3] Impact of Backorder on Supply Chain Performance - an Experimental Study  
<https://www.sciencedirect.com/science/article/pii/S1474667016345785>  
<https://doi.org/10.3182/20130619-3-RU-3018.00477>
- [4] BNE - Banco Nacional de Empregos  
<https://www.bne.com.br/curriculo-de-assistente-de-pcp-em-fortaleza-ce/10684485>
- [5] Catho Empresas  
<https://www.catho.com.br/buscar/curriculos/curriculo/18295544/>
- [6] KPI - % of backorders: Percentage of unfulfilled orders.  
<http://kpilibrary.com/kpis/of-backorders>
- [7] Predicting backorders with 3 models  
<https://www.kaggle.com/dredlaw/predicting-backorders-with-3-models>
- [8] Predict Product Backorders with SMOTE and RF  
<https://www.kaggle.com/haimfeld87/predict-product-backorders-with-smote-and-rf>
- [9] Previsão de falta de materiais no contexto de gestão inteligente de inventário: uma aplicação de aprendizado desbalanceado  
<https://repositorio.ufjf.br/jspui/bitstream/ufjf/6861/1/rodrigobarbosadesantis.pdf>
- [10] Backorder Prediction Using Machine Learning For Danish Craft Beer Breweries  
[https://projekter.aau.dk/projekter/files/262657498/master\\_thesis.pdf](https://projekter.aau.dk/projekter/files/262657498/master_thesis.pdf)
- [11] Sales Analytics: How to Use Machine Learning to Predict and Optimize Product Backorders  
<https://www.r-bloggers.com/sales-analytics-how-to-use-machine-learning-to-predict-and-optimize-product-backorders/>
- [12] Stockout Prediction using machine learning  
<https://www.linkedin.com/pulse/stockout-prediction-using-machine-learning-vinay-arun/>
- [13] Backorder Prediction Modeling  
<https://rpubs.com/cschumacher16/backorders>
- [14] scikit-learn - Machine Learning in Python  
[http://scikit-learn.org/stable/user\\_guide.html](http://scikit-learn.org/stable/user_guide.html)
- [15] imbalanced-learn - SMOTE, under-sampling and over-sampling methods  
<http://contrib.scikit-learn.org/imbalanced-learn/stable/index.html>