

## INTRODUÇÃO

Muito dos artigos que vemos publicados na área de machine learning acabam focando muito no aspecto teórico, por isso decidimos realizar um summary de um paper aplicado em um caso prático e o escolhido foi o paper “Analytics for an Online Retailer: Demand Forecasting and Price Optimization” dos autores: Kris Johnson Ferreira, Bin Hong Alex Lee, David Simchi-Levi. O nosso intuito é explicar as principais ideias do artigo de forma simples para que pessoas fora da área consigam compreender como foi feito o trabalho em questão.

## CONTEXTUALIZAÇÃO

O trabalho apresentado no artigo foi realizado em uma loja de roupas online chamada Rue La La, os autores demonstraram como a loja poderia utilizar os dados coletados para otimizar a precificação dos seus produtos. A loja funciona publicando “eventos” no site, onde cada evento dura de 1 a 4 dias e é composto de um conjunto de produtos em promoção e cada produto é chamado de “estilo”.

Para compreender a abordagem dos autores é importante ressaltar que uma das maiores dificuldades da loja era precificar e entender a demanda de estilos que estão sendo vendidos pela primeira vez no site mesmo que uma grande parte da receita venha desse grupo de estilos. Devido a estes fatores eles decidiram desenvolver uma ferramenta de auxílio a precificação baseada em duas fases: 1-Previsão da demanda de estilos que estão sendo expostos pela primeira vez. 2- Utilização dos dados da previsão para criar um modelo de otimização de preço com o intuito de maximizar a receita.

Decidimos explicar alguns conceitos utilizados pelos autores que podem não ser comuns para alguns leitores, com o objetivo de facilitar a leitura completa do artigo.

- Algoritmos paramétricos: São algoritmos que fazem suposições sobre os dados utilizados. Ex: Assumem que os dados seguem uma distribuição normal.
- Otimizador de preço dinâmico: Atualização constante do preço de um produto a partir dos dados coletados.
- Markdown Pricing: É a redução do preço de um item para esvaziar o inventário.
- K-Fold Cross-Validation: Os dados de treino são divididos aleatoriamente em K grupos, depois o modelo é treinado utilizando K-1 grupos e então avaliado no grupo de dados que não foi utilizado, esse procedimento ocorre K vezes, de maneira que todos os grupos sejam utilizados para validar o modelo uma vez. Tal método é utilizado para avaliar a habilidade de generalização do modelo e possibilita ajustar os hiperparâmetros do modelo.

## METODOLOGIA

Para que o algoritmo funcione devidamente, uma peça chave é justamente poder prever com certa acurácia a demanda de novos produtos a serem exibidos em eventos.

Para tal, primeiro decide-se o nível de detalhamento das agregações de produtos a serem feitas, desconsiderando o tamanho desses produtos. Isso é feito basicamente porque o preço de certo estilo de produto é o mesmo para todos os seus tamanhos.

Em resumo, seguem-se os seguintes passos para fazer uma predição da demanda de novos produtos:

1. Registrar vendas de certos estilos no passado;
2. Estimar a demanda desses mesmos estilos no passado;
3. Prever a demanda e vendas de novos estilos que serão vendidos no futuro.

Para ser feito o registro de vendas no passado, os dados a serem utilizados nessa parte consistem das vendas feitas na loja desde o começo de 2011 até metade de 2013, em que cada registro representa a venda de certo item durante um evento. As características a serem consideradas de cada venda podem ser vistas no seguinte sumário:



Posteriormente, na parte de estimar a demanda desses mesmos estilos no passado, surge um empecilho de que a demanda prevista estará, muito provavelmente, subdimensionada devido aos casos em que houve falta de estoque. Para isso, é necessário achar uma maneira de estimar esses casos em que houve perda de vendas por causa da falta de estoque. Para tal, utiliza-se os dados referentes aos itens que não foram vendidos para fazer uma estimativa da perda em vendas dos itens que ficaram fora de estoque. Em outras palavras, as proporções entre itens que sobraram no estoque e itens que ficaram fora de estoque foram normalizadas para que se chegasse a uma estimativa de quantos itens a mais seriam vendidos caso esses se encontrassem em estoque.

Após essas duas etapas, depois os dados adquiridos foram computados num modelo de regressão linear para obter uma predição de futuras demandas e vendas para novos estilos.

## RESULTADOS

O trabalho resultou em uma ferramenta automática de suporte a precificação que roda diariamente produzindo recomendações de preços para eventos que irão começar no dia seguinte e através de um experimento realizado na loja durante os cinco primeiros meses de 2014, o uso da ferramenta acarretou em 9.7% de aumento na receita associado a um intervalo de confiança de 90% de [2.3%, 17.8%], tal resultado levou a loja a adotar o instrumento de precificação diariamente.

Referências: Analytics for an online retailer: Demand Forecasting and Price Optimization