

# Grupo Boticário: Inteligência Artificial para recomendação do melhor presente

## Introdução

Um dos grandes dilemas que todos temos quando estamos procurando algo para presentear é justamente o que comprar! São sempre tantas opções e há uma grande pressão, pois estamos em busca do presente que mais vai surpreender e agradar.

Conhecemos a pessoa a qual queremos presentear, mas como fazer para identificar qual é o presente ideal para ela?

Baseado neste dilema, o Grupo Boticário, em busca da melhor experiência para seus clientes, quer dar uma ajudinha tecnológica para essa decisão. E se pudéssemos ter além das melhores opções de presente, um assistente virtual com um sistema de recomendação de produtos?

A ideia é que o cliente possa interagir com um assistente virtual e, após responder algumas perguntas sobre a pessoa a ser presenteada, o assistente forneça sugestões de produtos mais aderentes ao perfil.

## Desafio

O objetivo deste desafio é criar um sistema de recomendação de produtos para seus clientes, baseado nos perfis das linhas. Para realizar esta atividade, o participante deve:

1. Construir um modelo preditivo usando o *Watson Machine Learning* (WML) e os dados disponíveis no arquivo *dataset.xls*;
2. Será fornecido ao participante um *Watson Assistant* (WA) já configurado. Para que o Assistente Virtual, implementado pelo WA, possa recomendar produtos é necessário conectar o WA ao modelo preditivo disponibilizado pelo WML. Para que isso aconteça, o participante deverá:
  - a. Implementar uma página web que recebe os dados do WA e envia para o endpoint do WML, e retorna o resultado do WML para o WA.
  - b. Conectar o WA com a *página* implementada.

Cada participante deve criar seu próprio modelo de classificação no *Watson Machine Learning*, utilizando os algoritmos de classificação já presentes na ferramenta. Depois de devidamente criado, o modelo deve ser público e acessado pelo *Watson Assistant*.

Desta forma, através de uma interação com o *Watson Assistant* serão coletadas as informações do perfil desejado e assim o modelo gerado no *Watson Machine Learning* será consultado e a recomendação retornada para o cliente.

## Construção do Modelo Preditivo

Para a construção do modelo preditivo, função de recomendação, o participante deverá utilizar o dataset disponível no arquivo dataset.xlsx. Os dados disponíveis neste arquivo têm o formato apresentado na tabela 1.

	A	B	C	D	E
1	Sexo <input type="text"/>	Estilo <input type="text"/>	Lugar <input type="text"/>	Bebida <input type="text"/>	Produto <input type="text"/>
2	homem	netflix	floresta	vinho	coffee
3	homem	netflix	floresta	chope	coffee
4	homem	netflix	floresta	milk_shake	coffee
5	homem	netflix	floresta	suco_detox	coffee
6	homem	netflix	floresta	café	coffee
7	homem	balada	floresta	vinho	malbec
8	homem	balada	floresta	chope	malbec
9	homem	balada	floresta	milk_shake	malbec
10	homem	balada	floresta	suco_detox	malbec
11	homem	balada	floresta	café	malbec
12	homem	bar	floresta	vinho	malbec
13	homem	bar	floresta	chope	malbec
14	homem	bar	floresta	milk_shake	malbec
15	homem	bar	floresta	suco_detox	malbec
16	homem	bar	floresta	café	malbec
17	homem	museu	floresta	vinho	zaad
18	homem	museu	floresta	chope	zaad
19	homem	museu	floresta	milk_shake	zaad
20	homem	museu	floresta	suco_detox	zaad

Tabela 1: Exemplo de dados contidos no arquivo dataset.xls

Este dataset possui 5 atributos categóricos. O atributo que queremos aprender, ou seja, a recomendação a ser fornecida é o atributo **Produto**. Os demais atributos (**Gênero**, **Estilo**, **Lugar** e **Bebida**) são os atributos independentes utilizados para prever o valor de **Produto**.

Este dataset possui 1.000 exemplos. Estes exemplos foram gerados através de computador simulando dados oriundos de base histórica de venda. Os atributos possuem os seguintes valores:

- **Gênero**: homem, mulher e nem\_um\_nem\_outro;
- **Estilo**: balada, bar, museu, netflix e parque;
- **Lugar**: campo, cidade, floresta e praia;
- **Bebida**: café, chope, milk\_shake, suco\_detox e vinho;

- **Produto:** siage\_hydraprotein\_eudora; boticollection\_boticario; coffee\_boticario; eau\_de\_parfum\_eudora; batom\_liquido\_mate\_quemdisseberenice; batom\_volumao\_quemdisseberenice; malbec\_boticario; nativaspa\_boticario; impression\_eudora

Este dataset será a entrada para o *Watson Machine Learning*. O WML só manipula arquivos CSV. Então, em algum momento, o participante terá que converter este arquivo, cujo o formato é XLSX, em CSV. Uma opção é utilizar a função de *Refinery* do *Watson Studio*. Ao utilizar a função de *Refinery*, o participante poderá também identificar a frequência com que determinados valores são apresentados no dataset. Um exemplo é apresentado na figura 1.

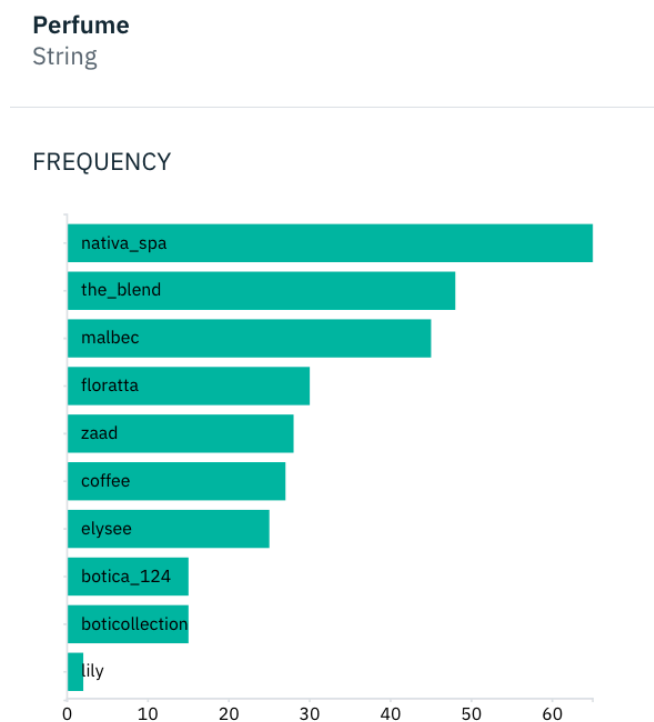


Figura 1: Frequência dos tipos de perfumes no dataset.

De qualquer forma, o importante é ter um arquivo no formato CSV com os dados originais para então submeter ao WML.

Ao utilizar o WML, o participante terá que informar o arquivo CSV, escolher qual é a coluna que queremos aprender (no caso, Produto), escolher as demais colunas como variáveis que serão utilizadas pelo modelo e selecionar todos os algoritmos de *machine learning* adequados para o problema, como apresentado na figura 2.

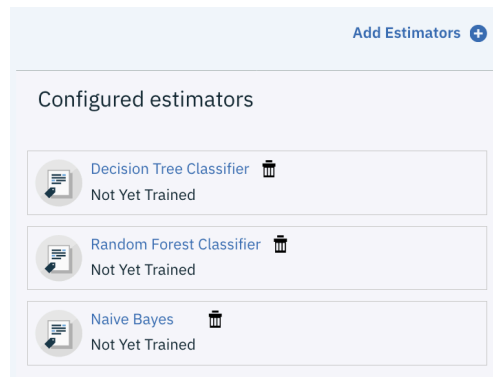


Figura 2: Algoritmos adequados para o problema

O WML irá criar e avaliar diferentes modelos para o problema. O participante terá que escolher o melhor modelo, salvar e depois publicar como uma API Rest (*deployment type = web service*).

Na figura 3 é apresentado a tela onde um teste é realizado com o endpoint disponibilizado pelo WML. Neste exemplo, é considerado uma mulher, com estilo igual a parque, lugar preferido igual a campo e bebida suco detox. Para este caso, o modelo construído recomenda o perfume Boticollection.

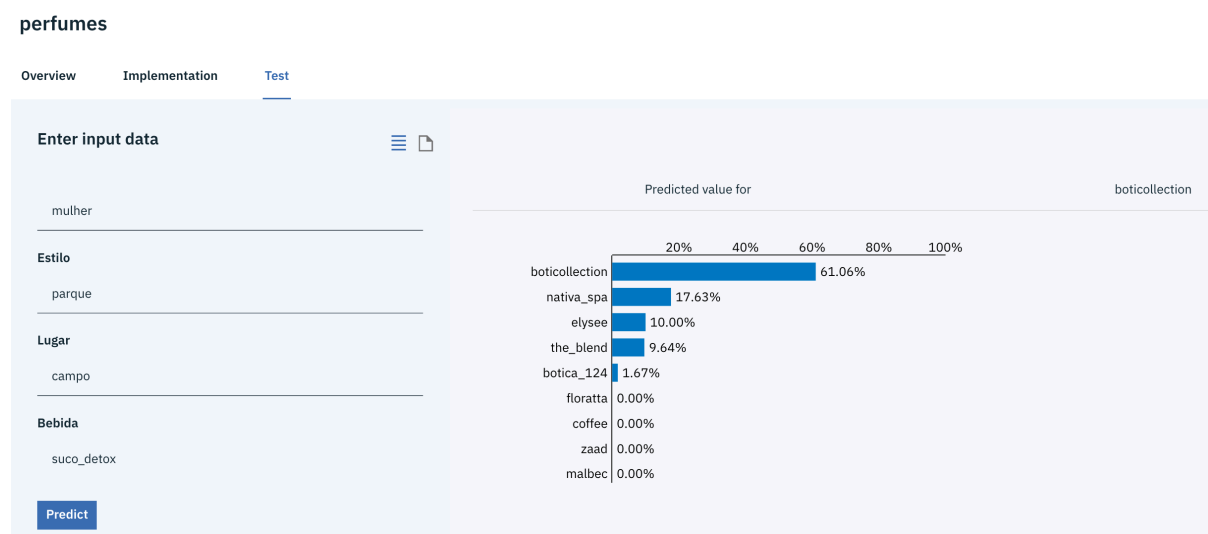


Figura 3: Exemplo de teste usando o endpoint do WML