

Recomendação de Conteúdo

Everton Gago



Agenda

- Por que recomendar?
- Tipos de recomendação
- Como medir
- Experimento prático
- Extrapolação do experimento
- Desafios
- Pergunta chave

Por que recomendar?

Tipos de recomendação

- Por similaridade de usuários
- Por similaridade de itens
- Por associação

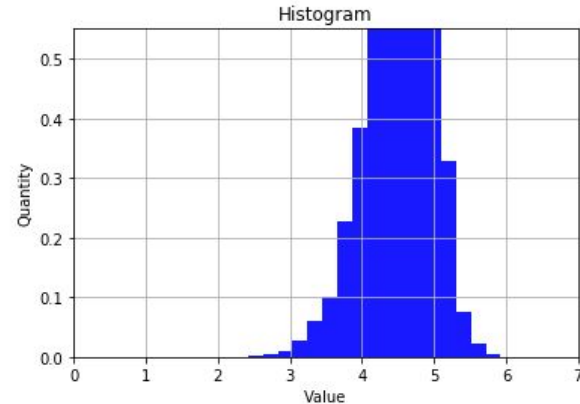
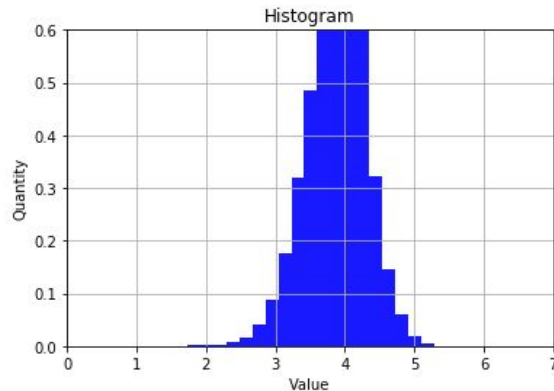
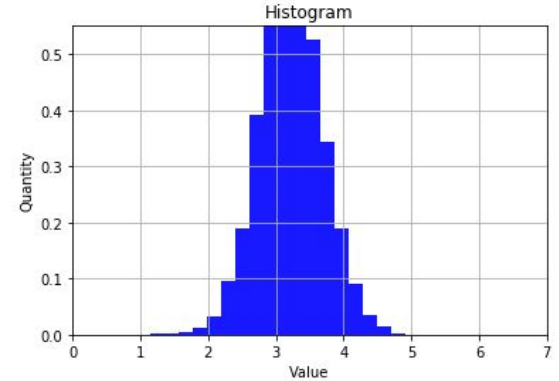
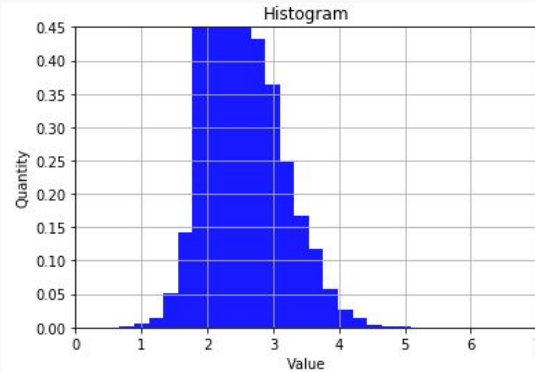
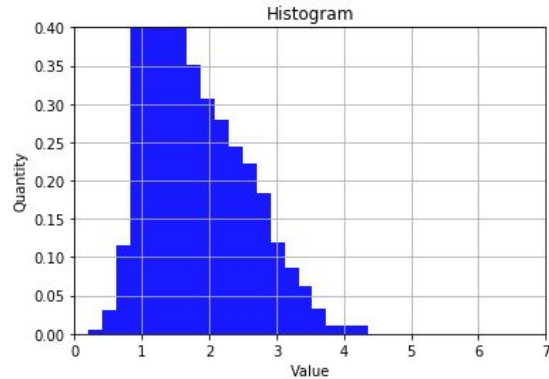
Como medir

- Medição off-line
- Medição on-line

Medição off-line

- Separar os dados em conjunto de treino e conjunto de teste (validação)
- Medir:
 - Precision x Recall
 - Medida de regressão (ex: Erro Quadrático Médio)
 - Distribuição de frequência das predições

Medições off-line - Distribuição de Frequências



Medição on-line

- Definir grupos de teste (Teste A/B)
- Definir hipóteses que serão validadas
- Definir e calcular métricas que possam validar / invalidar as hipóteses

Medição on-line

Exemplo

Hipótese: Quanto mais cedo o usuário encontrar o filme que procura, mais filmes ele vai assistir.

Métricas:

- Tempo entre a autenticação e a escolha do filme
- Quantidade de filmes que os usuários assistem na mesma sessão autenticada.
- Engajamento com a plataforma (frequência de uso)
- *etc ...*

Experimento Prático

<https://github.com/evertongago/devcircle>

Extrapolação do Experimento

<https://github.com/evertongago/experimento-recomendacao>

Desafios

- Grande volume de dados
- Definir critério de nota, quando não é explícita
- Definir hipóteses
- Definir e acompanhar métricas
 - Garantir relevância estatística
 - Coletar dados que garantam a realização das métricas
 - Interpretação das métricas
- Escopo do modelo (recomendador global ou contextual?)

Pergunta chave

Quando utilizar uma API e quando desenvolver o algoritmo?