

ShapleyGAN - Entendendo o Comportamento de Sistemas de Busca com Modelos Explicativos de Machine Learning

Dr. Junior A. Koch
Principal Data Scientist @ Elo7

linkedin.com/in/juniorkoch/
@ninjajrk

<https://www.meetup.com/elo7-dev-meetups/>





THE
DEVELOPER'S
CONFERENCE

Compradores
9 milhões

elo7

Vendedores
30 mil



entrepreneurship
innovation
artisanship
creativity
technology
humanization
personalization
handcraft



1 de cada 4 casamentos no Brasil tem o envolvimento do Elo7



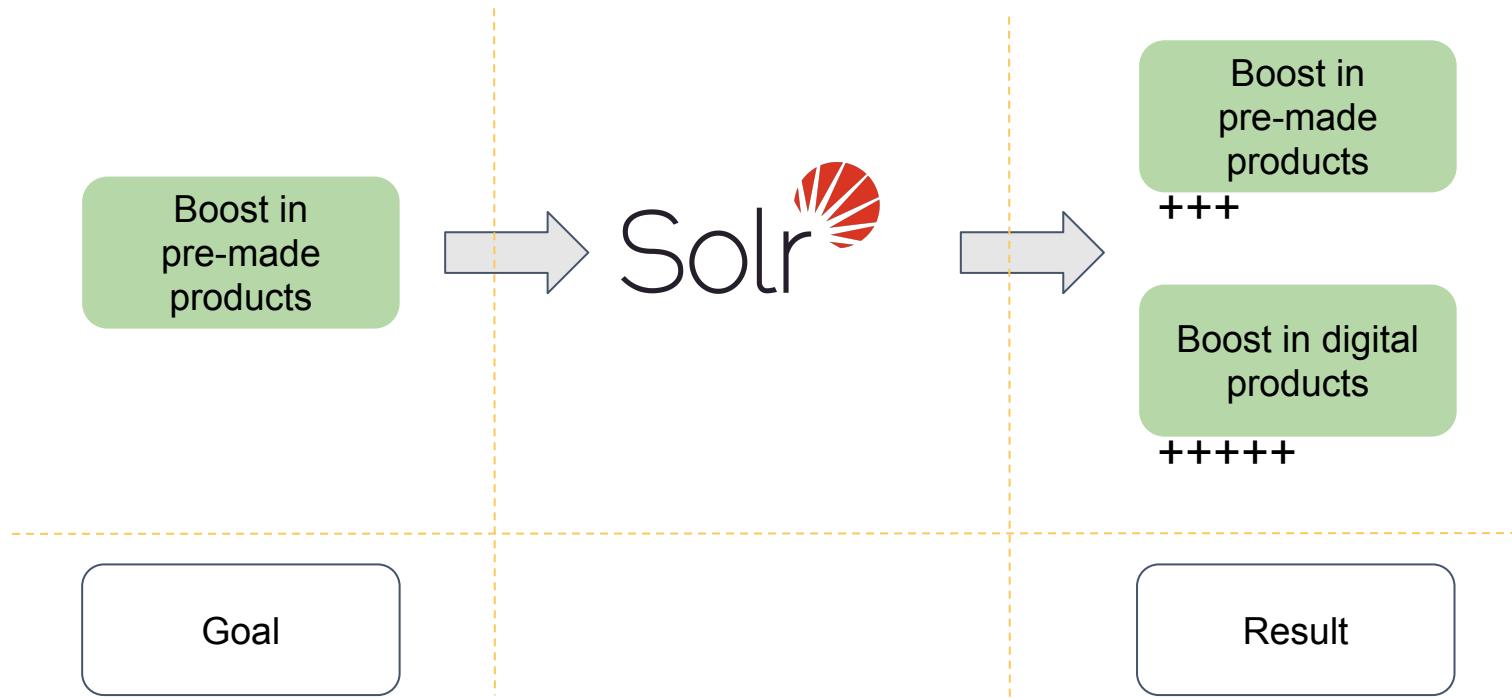
O time...



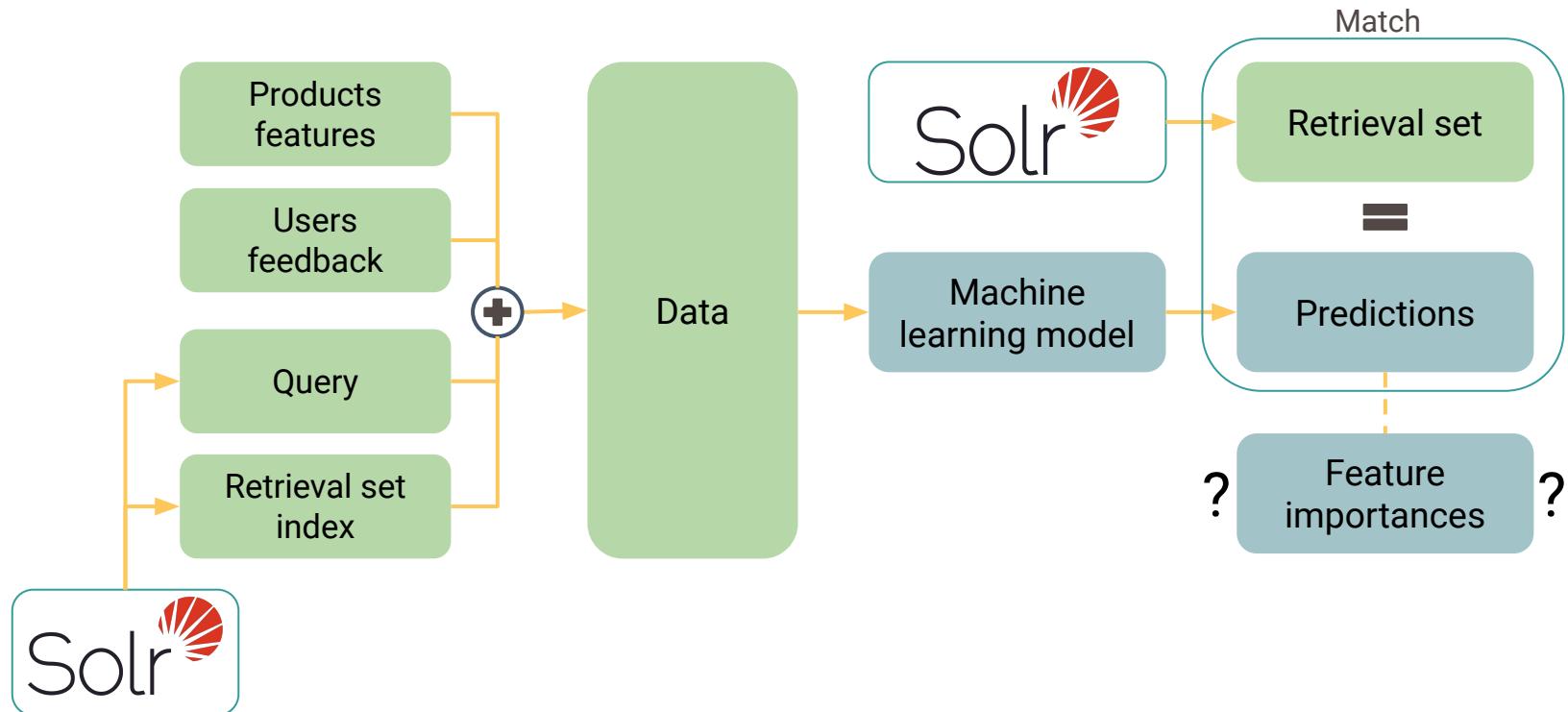
Engine de busca



Efeitos indesejados



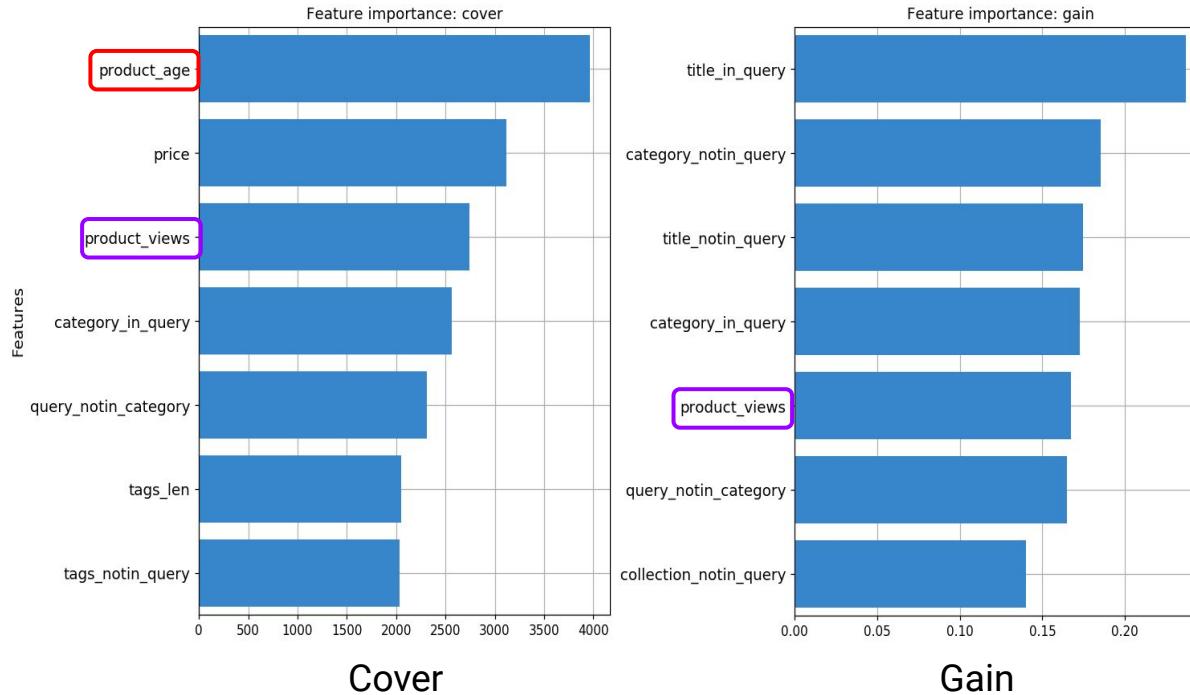
Explainable ML para Search Engine



Explainable ML para Search Engine



7 features mais importantes

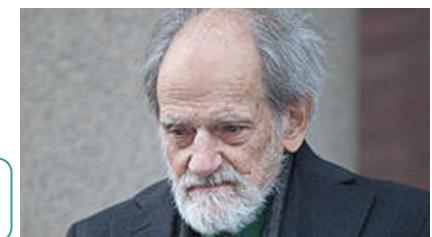
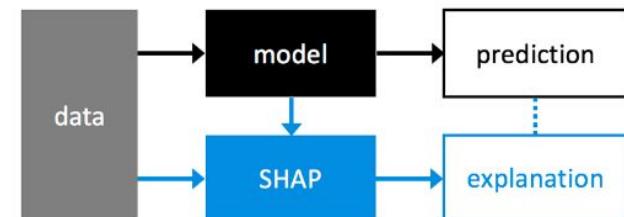


SHapley Additive exPlanations



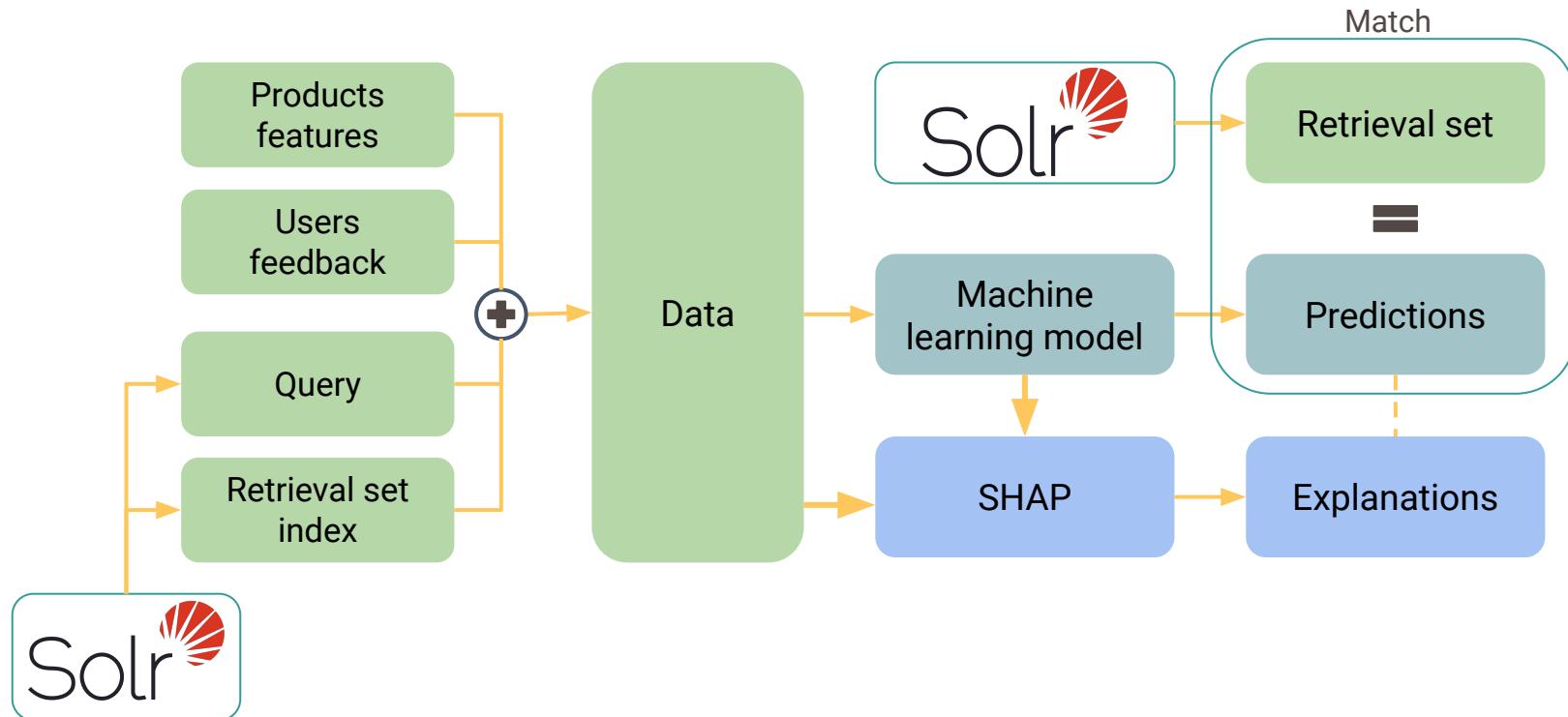
SHAP (SHapley Additive exPlanations) is a unified approach to explain the output of any machine learning model. SHAP connects game theory with local explanations, uniting several previous methods and representing the only possible consistent and locally accurate additive feature attribution method based on expectations.

A screenshot of the GitHub repository for "shap" by slundberg. The repository has 269 stars, 197 forks, and 865 issues. It includes sections for Code, Issues (359), Pull requests (17), Projects (1), Wiki, Security, and Insights. A description states: "A unified approach to explain the output of any machine learning model." Tags include interpretability, machine-learning, deep-learning, gradient-boosting, and shap. The repository has 976 commits, 3 branches, 34 releases, 65 contributors, and is licensed under MIT. It shows a "master" branch with a merge commit from slundberg. Buttons for Create new file, Upload files, Find File, and Clone or download are visible.

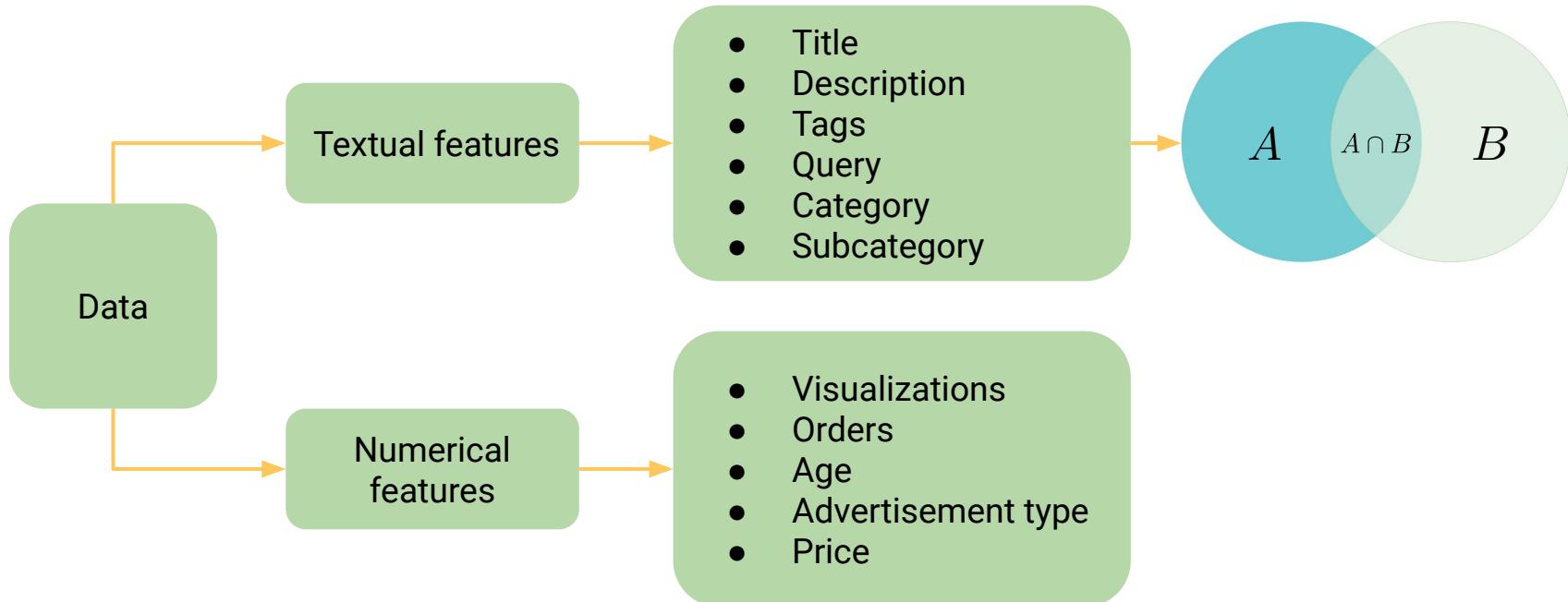


Nobel Laureate
Lloyd S. Shapley

Explainable ML para Search Engine

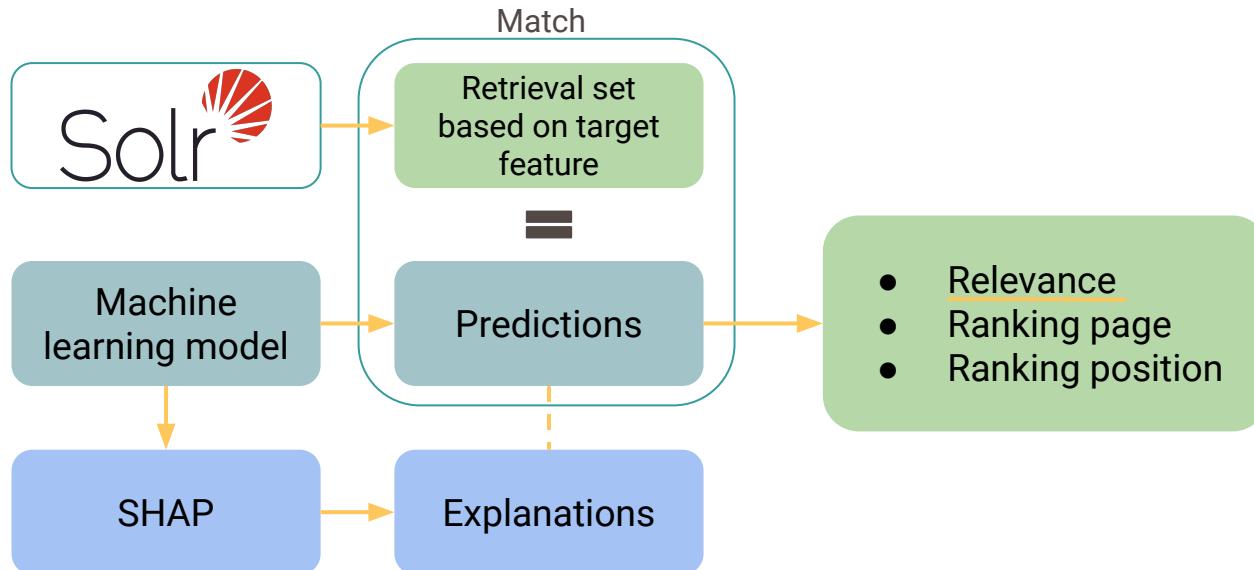


Feature Engineering



$$A = \# \text{ tokens in query}$$
$$B = \# \text{ tokens in features}$$

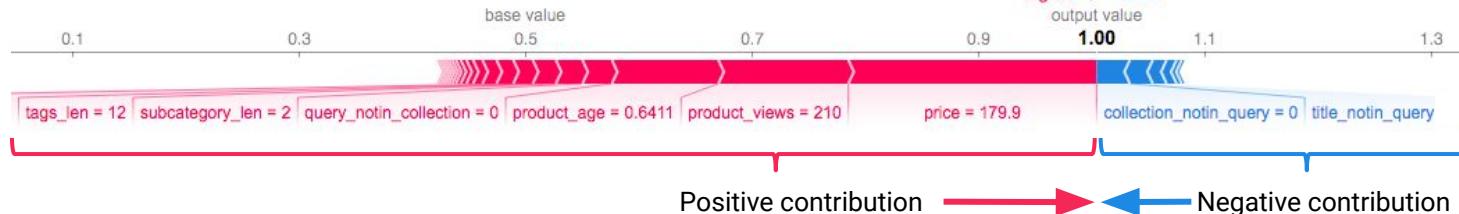
Target features



Resultados: relevância



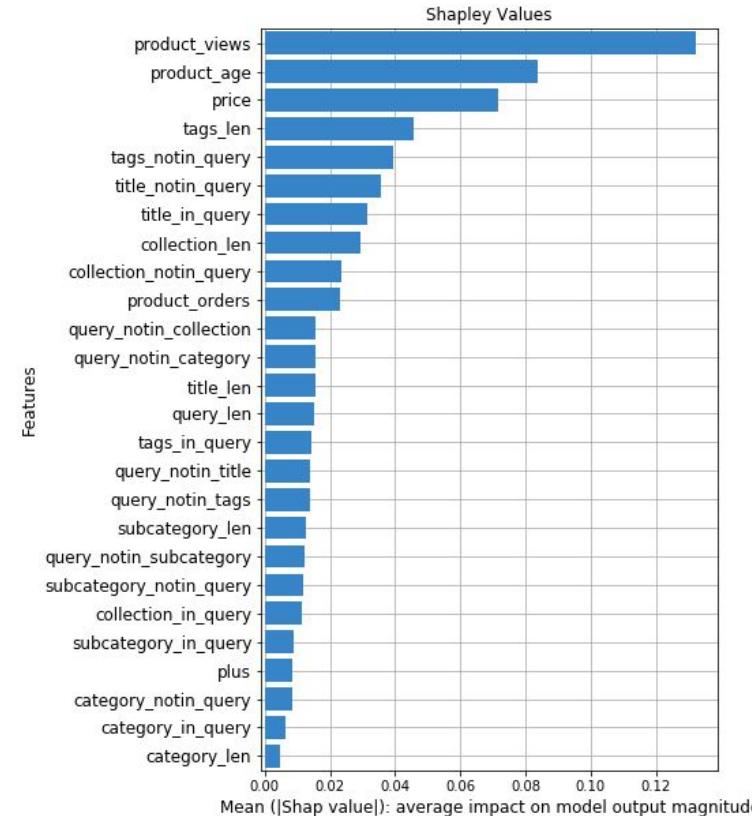
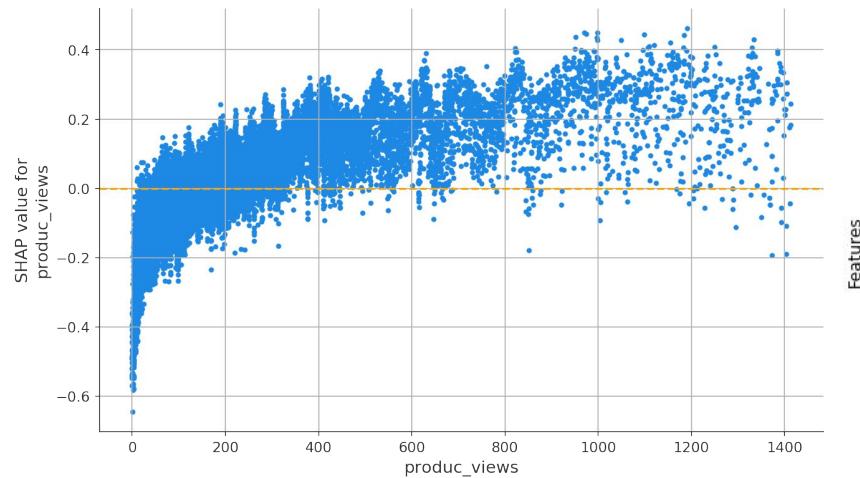
The screenshot shows a product page for a Pikachu Amigurumi. The product image is a yellow crocheted Pikachu sitting. To the left is a vertical sidebar with five smaller images of Pikachu. At the top, there's a navigation bar with 'Produtos', a search bar, and a 'CEP' input field. Below the product image, the title 'Pikachu Amigurumi' and price 'R\$ 70,00' are displayed, along with financing information. A large orange box highlights the text 'Relevant product'.





THE
DEVELOPER'S
CONFERENCE

Resultados: relevância



Conclusões dessa abordagem



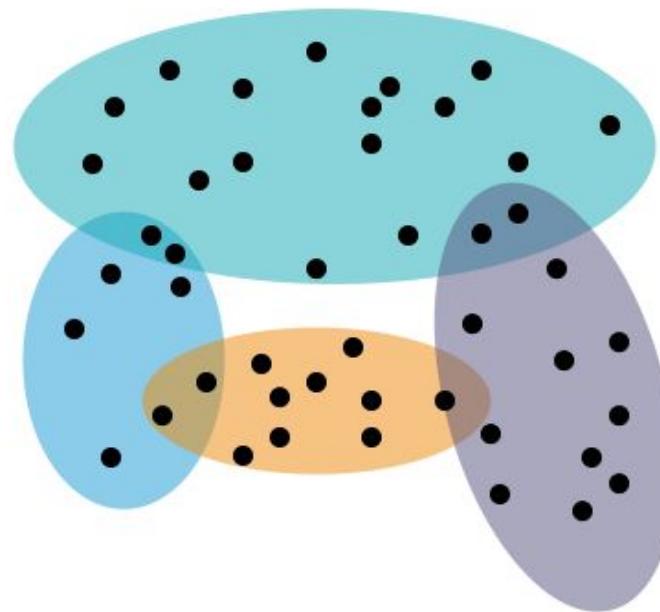
- Entender as características dos produtos que são relevantes para os usuários
- Agrupar produtos com explicações similares
- Verificar se mudanças na ferramenta de busca vão gerar o resultado esperado antes de ir para produção
- Validar características relevantes
- Construir engines de busca personalizadas para usos específicos

Indo além...

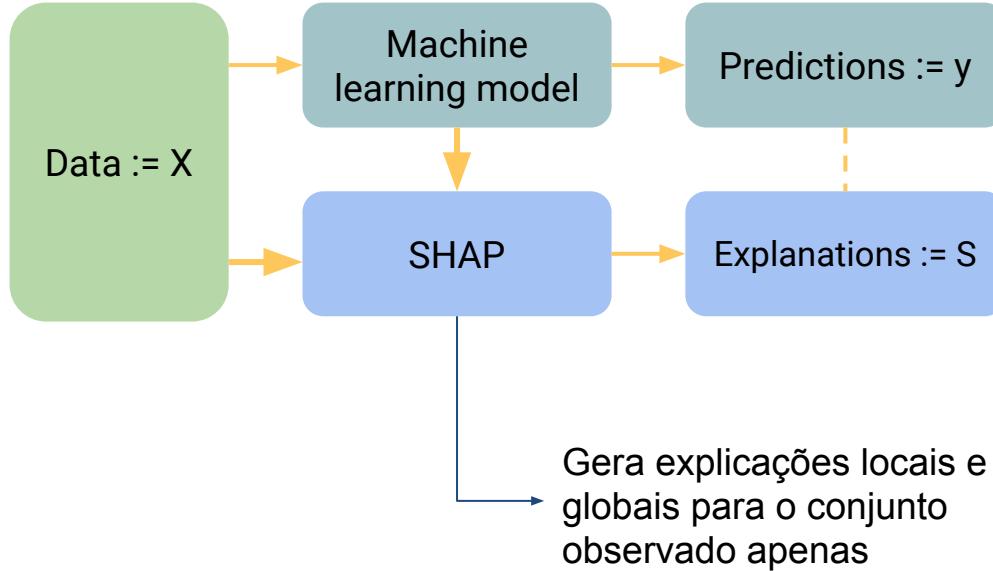


THE
DEVELOPER'S
CONFERENCE

- Views>300
- Price<=R\$30.00
- Free_shipment
- Age<1year



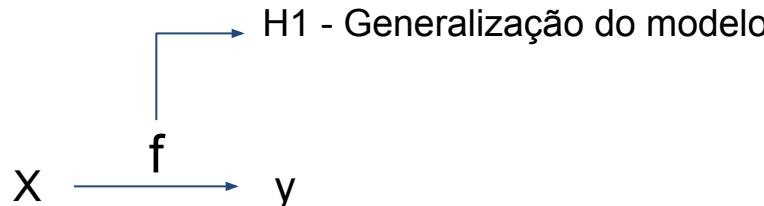
E se quisermos explicar uma nova observação?



X = Espaço de features
S = Espaço de valores de Shapley



Hipóteses



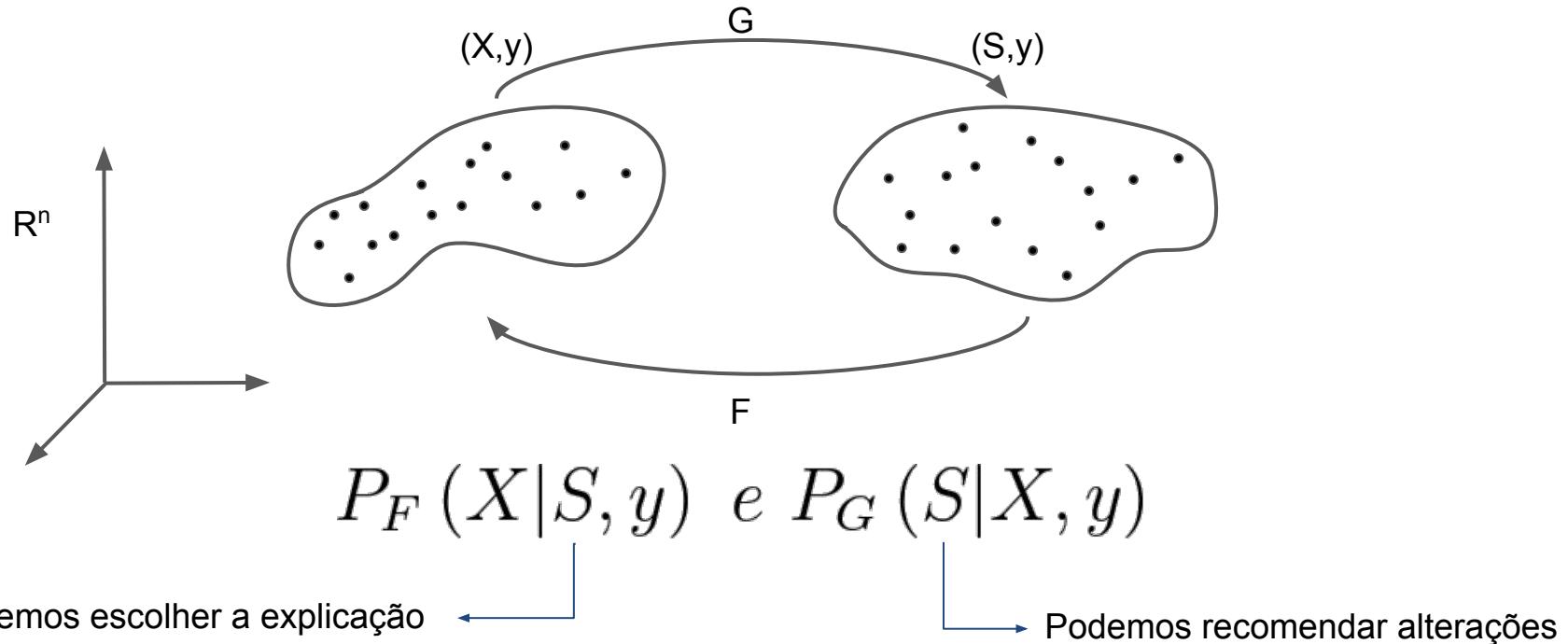
$x' \in x \sim \text{Dist}_X \Rightarrow \left\{ \begin{array}{l} ((X \oplus x'), y) \xrightarrow{\text{SHAP}} S' \\ (X, y) \xrightarrow{\text{SHAP}} S \end{array} \right\}$ tal que $S = S'$

Nova observação

```
graph LR; X[x] -- f --> Y[y]; H1[H1 - Generalização do modelo]; H2[H2 - Novas observações vêm da mesma distribuição e portanto seguem as mesmas explicações.]
```

SHAP: grande consumo de tempo para milhões de observações

Hipótese da Variedade*



*Manifold hypothesis



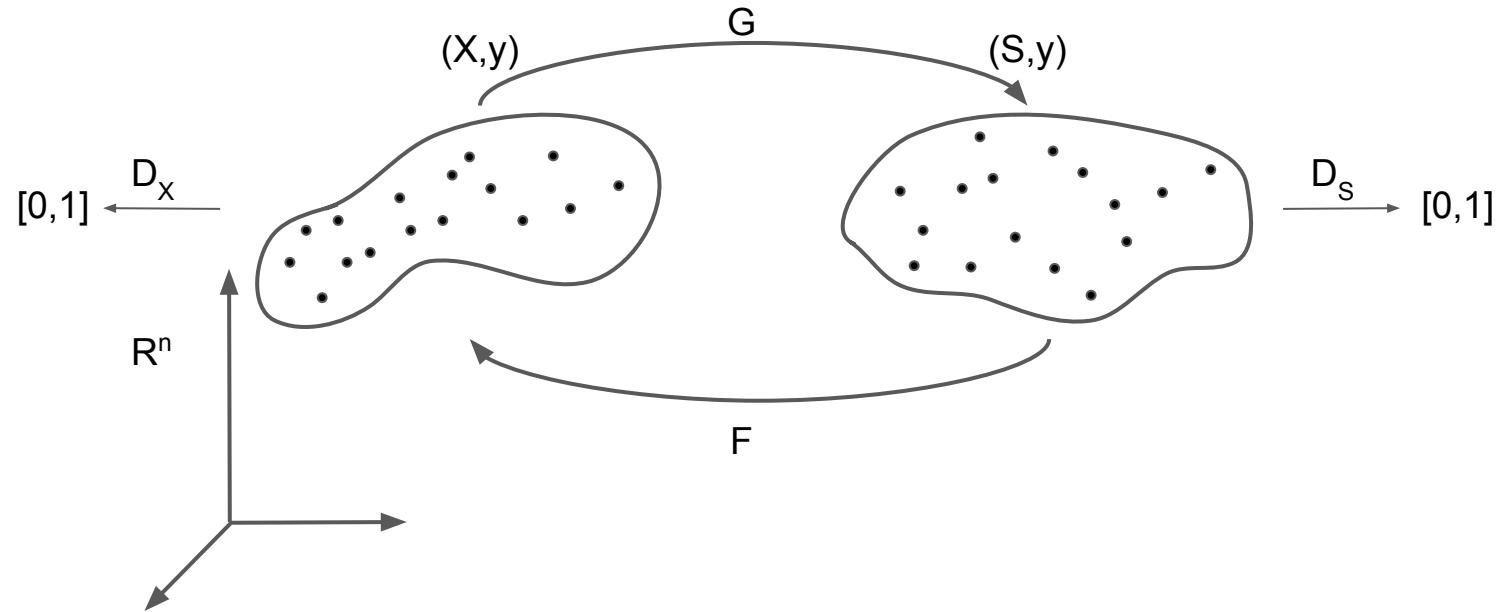
THE
DEVELOPER'S
CONFERENCE



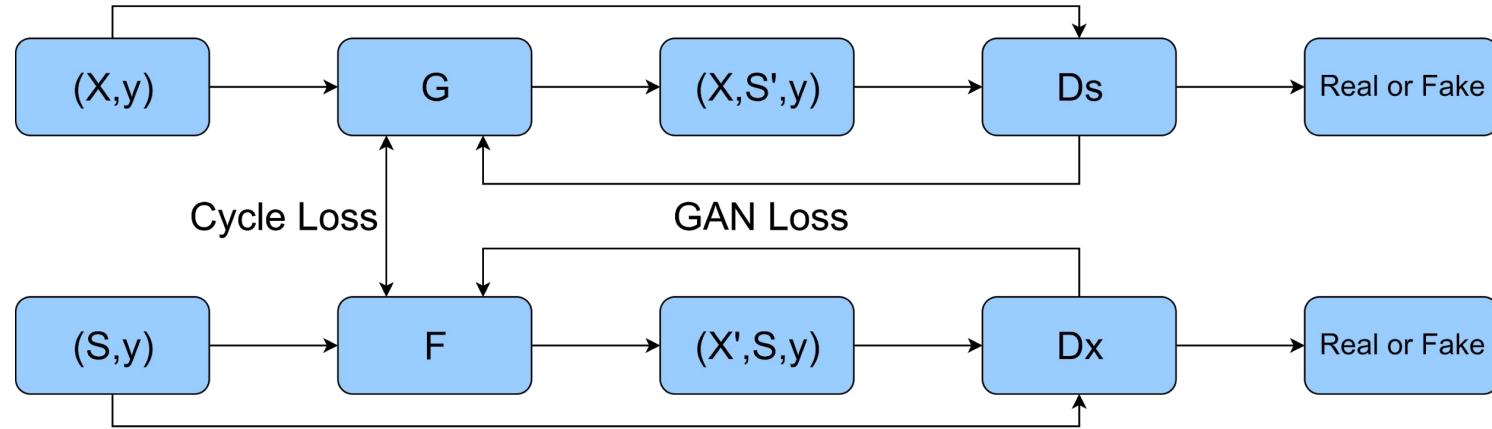
ShapleyGAN



THE
DEVELOPER'S
CONFERENCE



ShapleyGAN



ShapleyGAN



$$\mathcal{L}_{loss} = \mathcal{L}_{GAN} + \mathcal{L}_{Cycle}$$

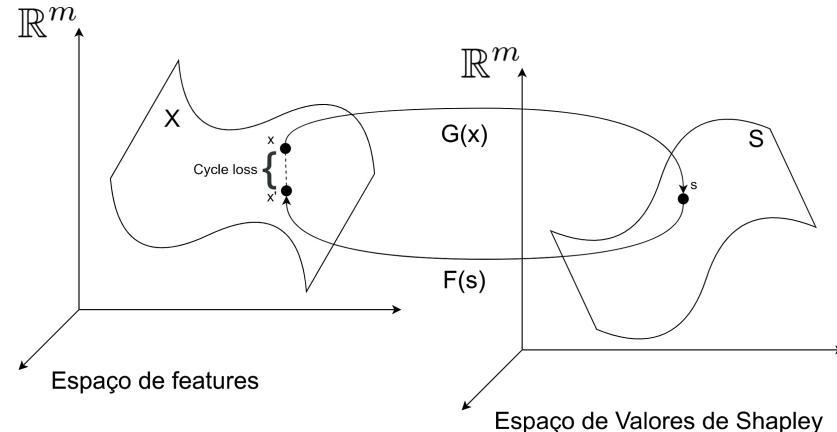
ou mais especificamente :

$$\mathcal{L}_{GAN}(G, D_S) = \mathbb{E}_{s \sim p_{data}(s), (x,y) \sim p_{data}(x,y)} [\log D_S(x, s, y)] + \mathbb{E}_{(x,y) \sim p_{data}(x,y)} [1 - \log D_S(x, G(x, y), y)],$$

$$\mathcal{L}_{GAN}(F, D_X) = \mathbb{E}_{x \sim p_{data}(x), (s,y) \sim p_{data}(s,y)} [\log D_X(x, s, y)] + \mathbb{E}_{(s,y) \sim p_{data}(s,y)} [1 - \log D_X(F(s, y), s, y)]$$

e

$$\mathcal{L}_{Cycle}(G, F) = \mathbb{E}_{(x,y) \sim p_{data}(x,y)} [| | | F(G(x, y), y) - x | |_1] + \mathbb{E}_{(s,y) \sim p_{data}(s,y)} [| | | G(F(s, y), y) - s | |_1].$$



Convergência e validação



Para acompanhar o treinamento das redes foram utilizadas as métricas usuais de GANs:

- D_x - taxa na qual o discriminador de X classificou dados falsos como reais;
- D_s - taxa na qual o discriminador de S classificou dados falsos como reais;
- Mode collapse - valor binário indicando se foi observado que o gerador aprendeu um conjunto de saída que engana facilmente o discriminador;
- G loss - perda do gerador de S;
- F loss - perda do gerador de X;
- D_x loss - perda do discriminador de X;
- D_s loss - perda do discriminador de S;
- $Cycle$ loss - perda específica da arquitetura do CycleGAN.

Convergência e validação



Métricas provenientes dos valores de Shapley:

- SHAP Efficiency - acurácia da soma dos valores de Shapley mais o predição média comparada com a classe real.

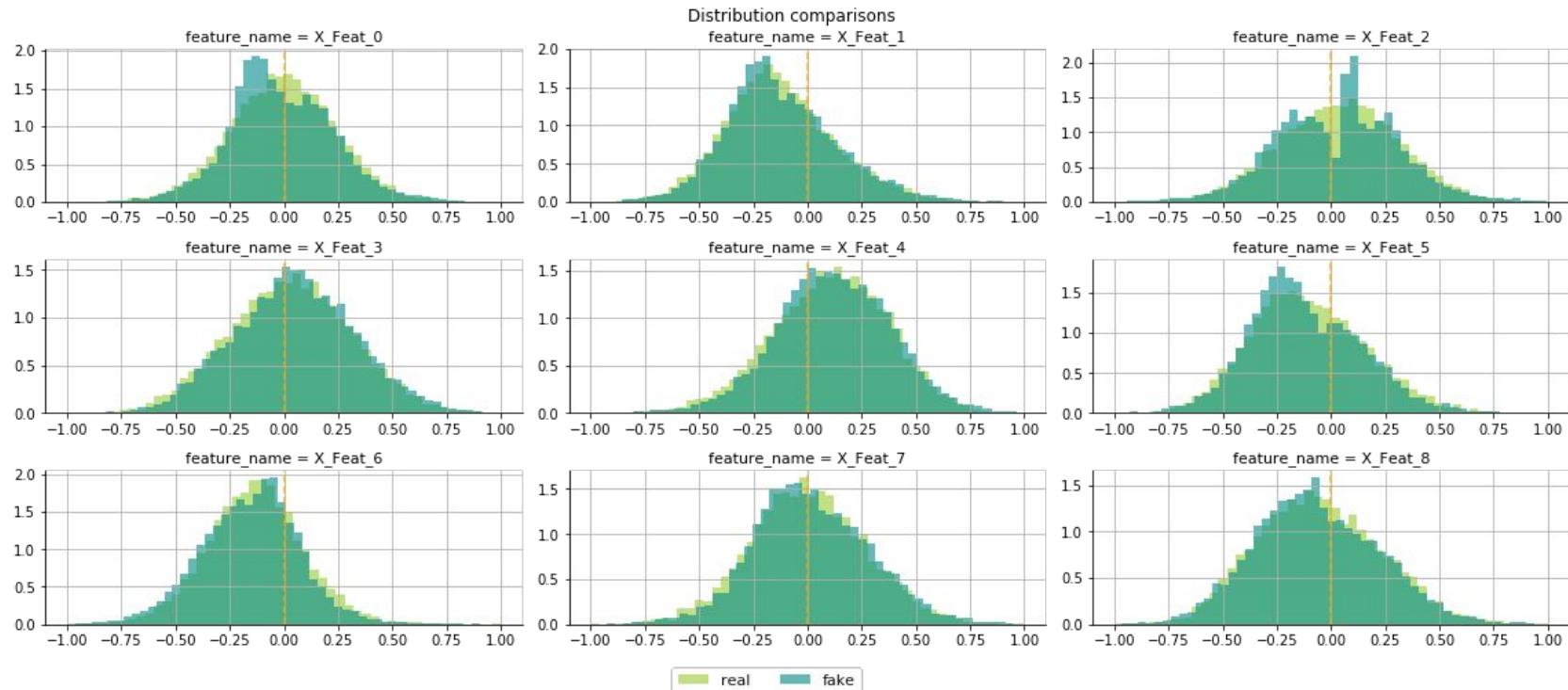
Métricas de análise topológica:

- GeometryScore - métrica baseada nas características topológicas dos dados;
- 1-d holes - número de buracos de dimensão 1 dos dados.

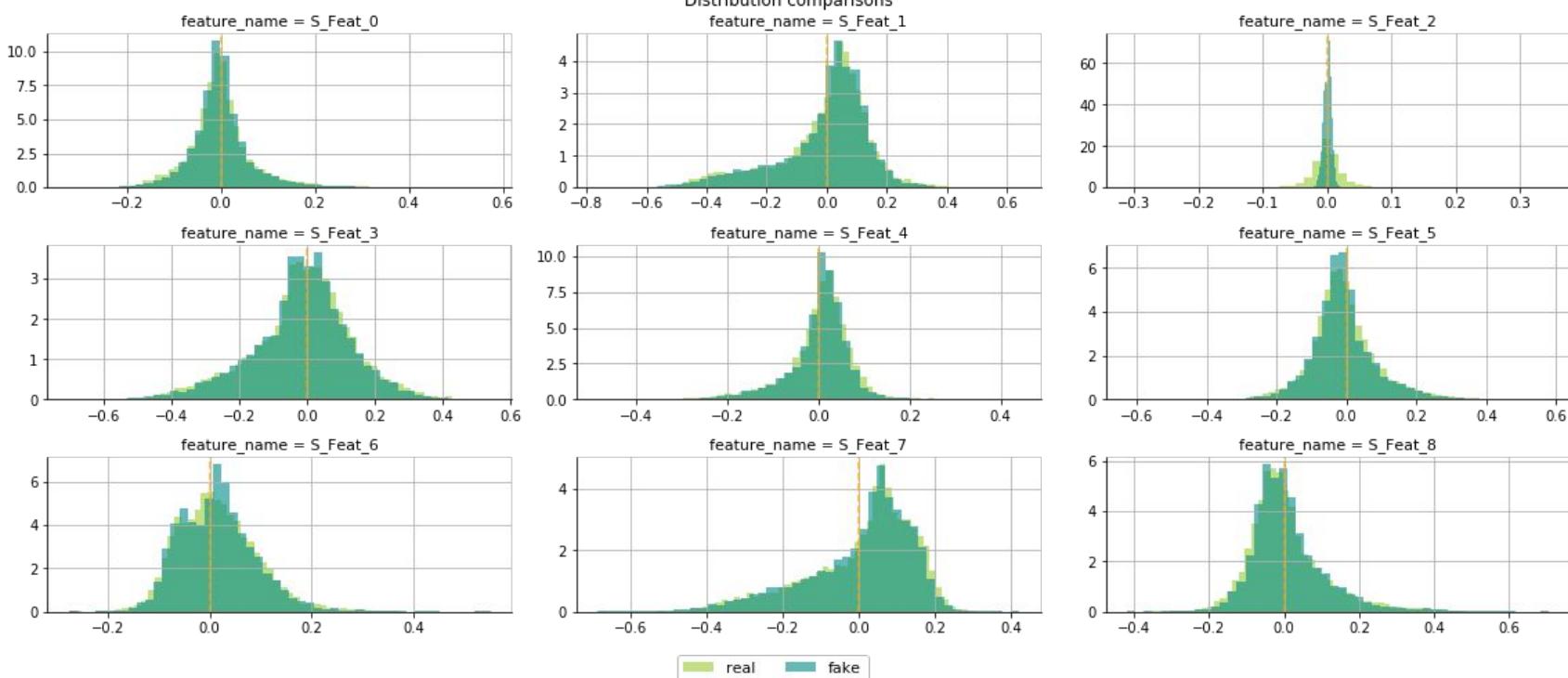
Métricas de distância:

- FID - Fréchet Inception Distance;
- Cos dist - distância de cosseno;
- L1 - distância L1;
- L2 - distância L2;
- KL - divergência de Kullback-Liebler;
- JSD - divergência de Jensen-Shannon.

Convergência e validação



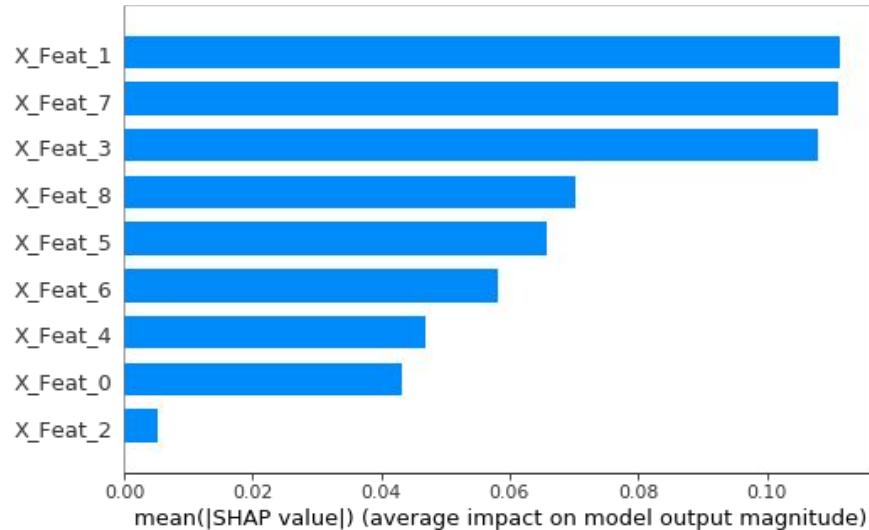
Convergência e validação



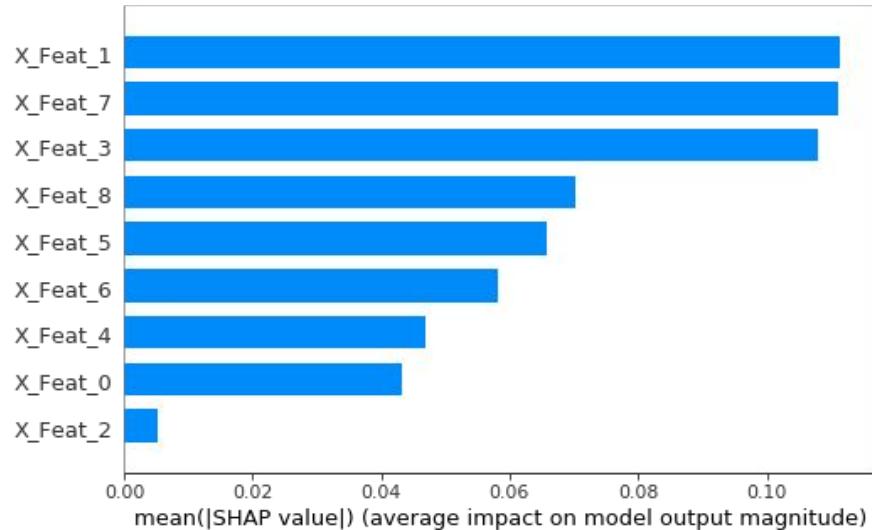
Explicações globais



ShapleyGAN explanation



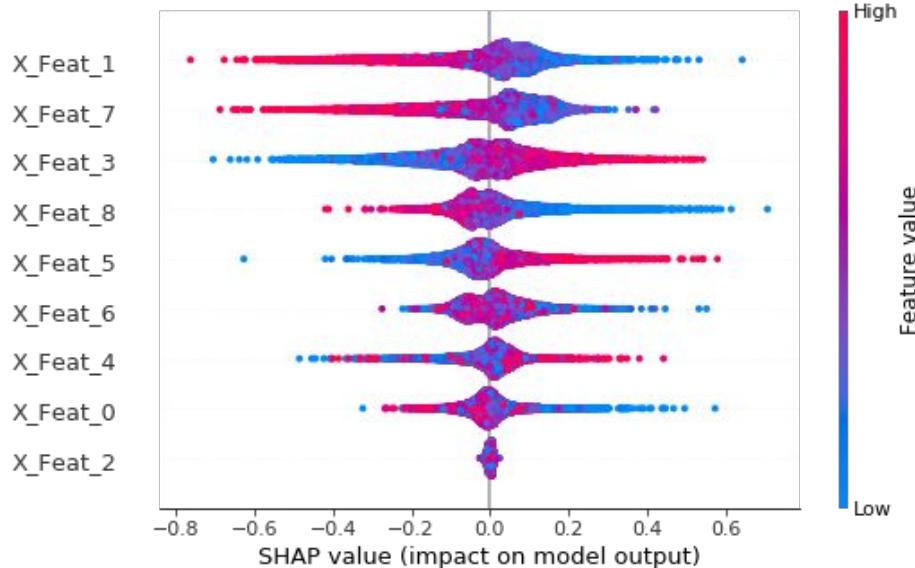
SHAP explanation



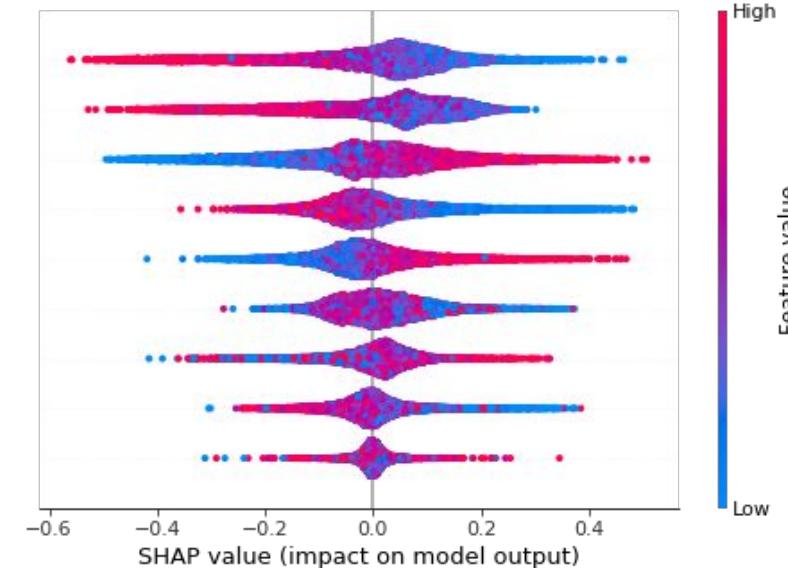
Explicações globais



ShapleyGAN explanation



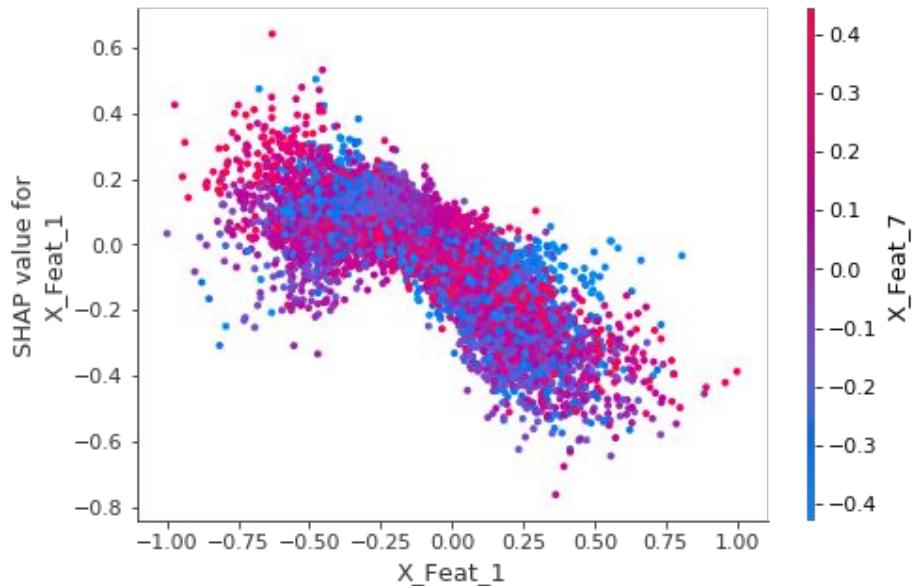
SHAP explanation



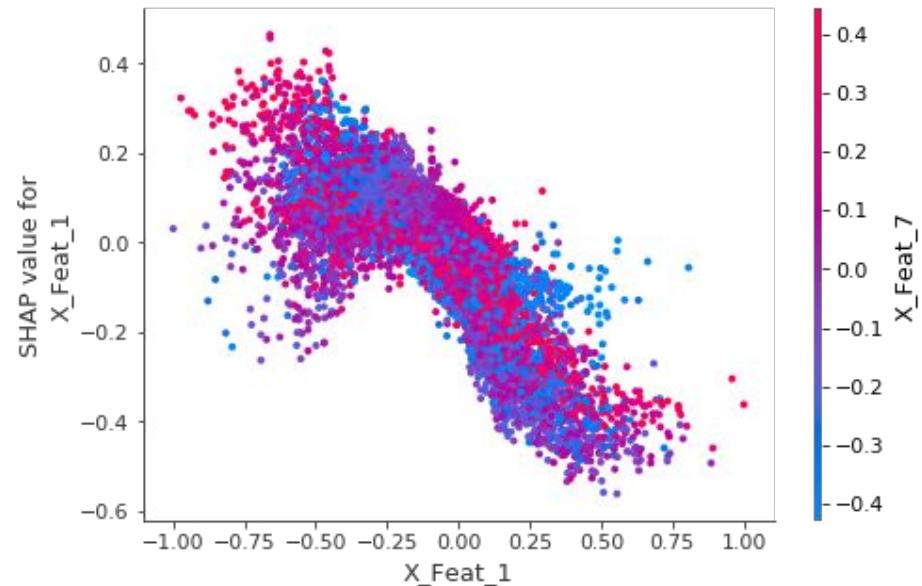
Explicações globais



ShapleyGAN explanation



SHAP explanation

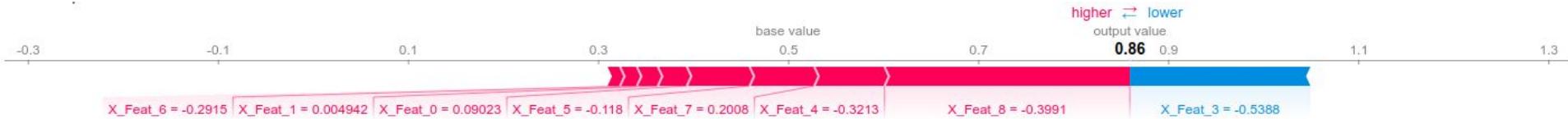




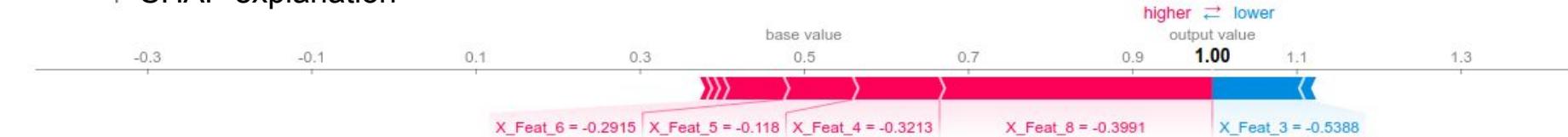
THE
DEVELOPER'S
CONFERENCE

Explicações locais

ShapleyGAN explanation



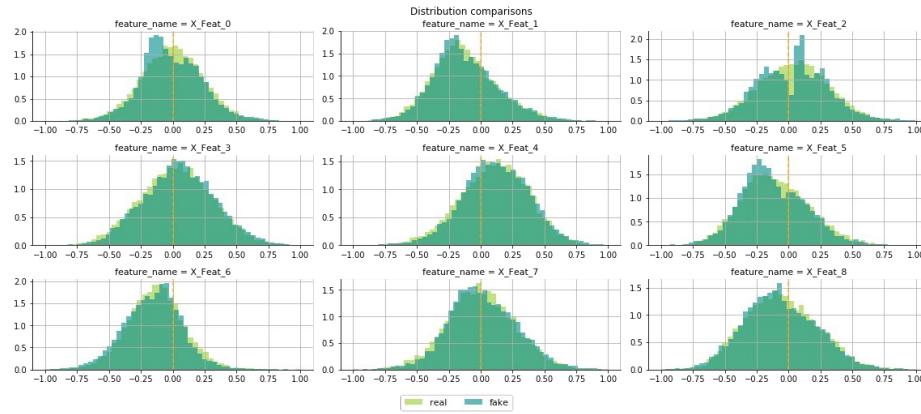
SHAP explanation



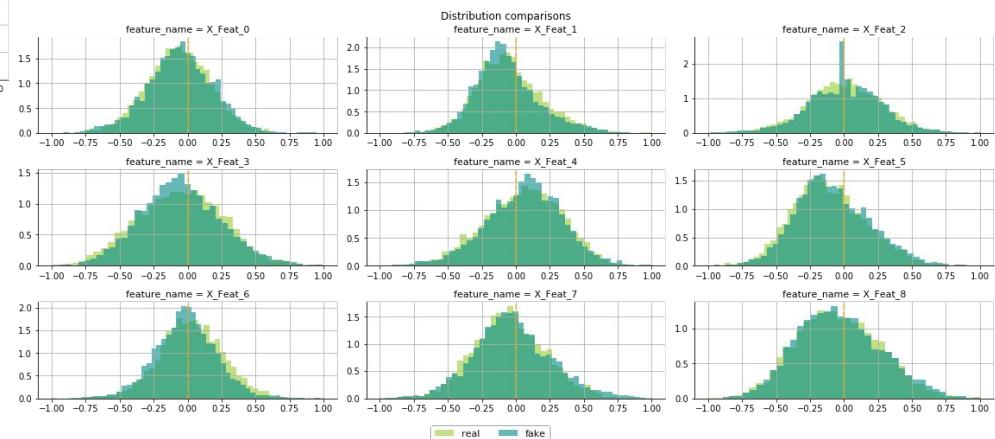
Generalização



THE
DEVELOPER'S
CONFERENCE



Dados de treino



Dados de validação

<https://www.meetup.com/elo7-dev-meetups/>

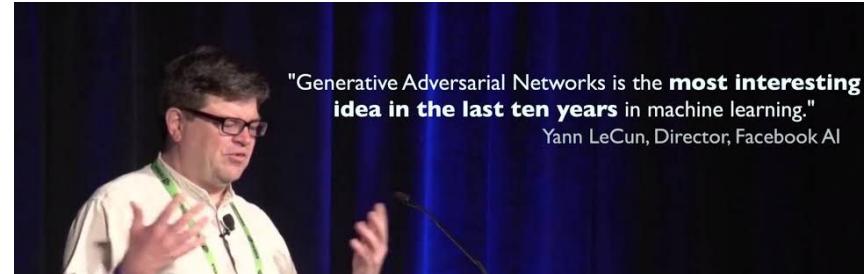
GAN School #01 - GAN tradicional e ConditionalGAN

29/09/2020 - 19-21h



GAN School #02 - InfoGAN e CycleGAN

06/10/2020 - 19-21h





: linkedin.com/in/juniorkoch/

elo7

www.elo7.com.br

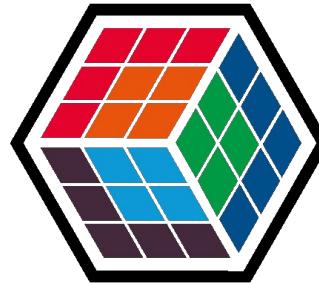
Elo7

Sobre Vagas Conheça Fazer login como candidato

EMPRESAS HUMANIZADAS
2019 - 2020

CAPITALISMO CONSCIENTE
2019 - 2020

Trabalhe conosco:
<https://elo7.gupy.io/>



THE DEVELOPER'S CONFERENCE