



Classifying text with **Machine Learning**

Ademir Ferreira
Full Stack Developer



Machine Learning: Conceito

Machine Learning é um conjunto de regras e procedimentos, que permite que os computadores possam agir e **tomar decisões** baseados em dados **ao invés de ser explicitamente programados** para realizar uma determinada tarefa.



Machine Learning: Conceito

✓ **Aprendizagem Supervisionada** (Supervised Learning)

Quando o algoritmo é treinado com um conjunto de dados contendo um resultado esperado.

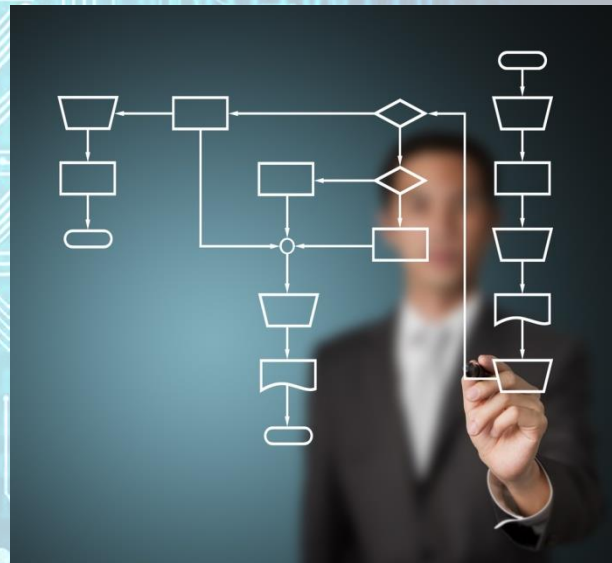
✓ **Aprendizagem Não Supervisionada** (Unsupervised Learning)

Quando o algoritmo pode encontrar um padrão e relações em um conjunto de dados, sem que exista um conhecimento prévio.



Machine Learning: Algoritmos

- ✓ Classificação (Classification)
- ✓ Regressão (Regression)
- ✓ Árvores de Decisão (Decision Trees)
- ✓ Modelo Gerador (Generative Model)
- ✓ Modelos Condicionais (Discriminative Model)
- ✓ Aprendizagem Profunda (Deep Learning)



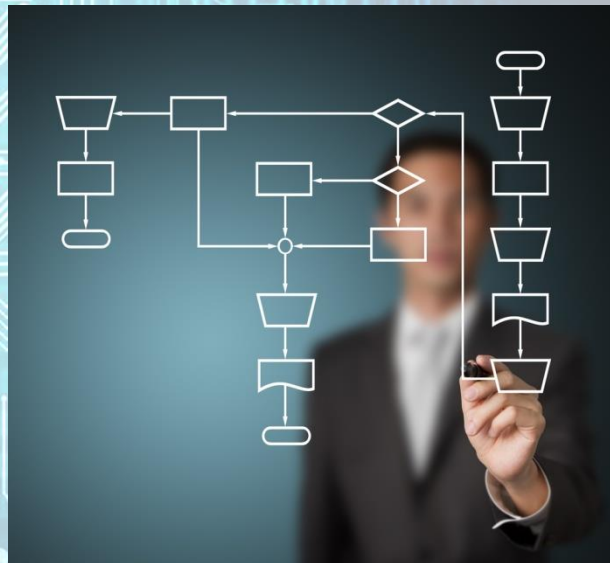
Machine Learning: Algoritmos

Classificação

É uma sub-categoria de aprendizagem **supervisionada**.

É o processo de tomar algum tipo de entrada e atribuir um **rótulo** (características) a ela.

Exemplo: Mapeamento de uma imagem de uma pessoa e classificação como masculino ou feminino.



Machine Learning: Tecnologias

A **scikit-learn** é uma biblioteca de aprendizado de máquina de código aberto para a linguagem de programação Python.



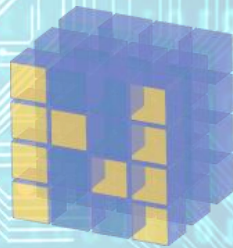
Machine Learning: Tecnologias

Ela inclui vários algoritmos de classificação, regressão e agrupamento incluindo máquinas de vetores de suporte, florestas aleatórias, gradient boosting, k-means e DBSCAN, e é projetada para interagir com as bibliotecas Python numéricas e científicas

NumPy e **SciPy**.



SciPy



NumPy

matplotlib

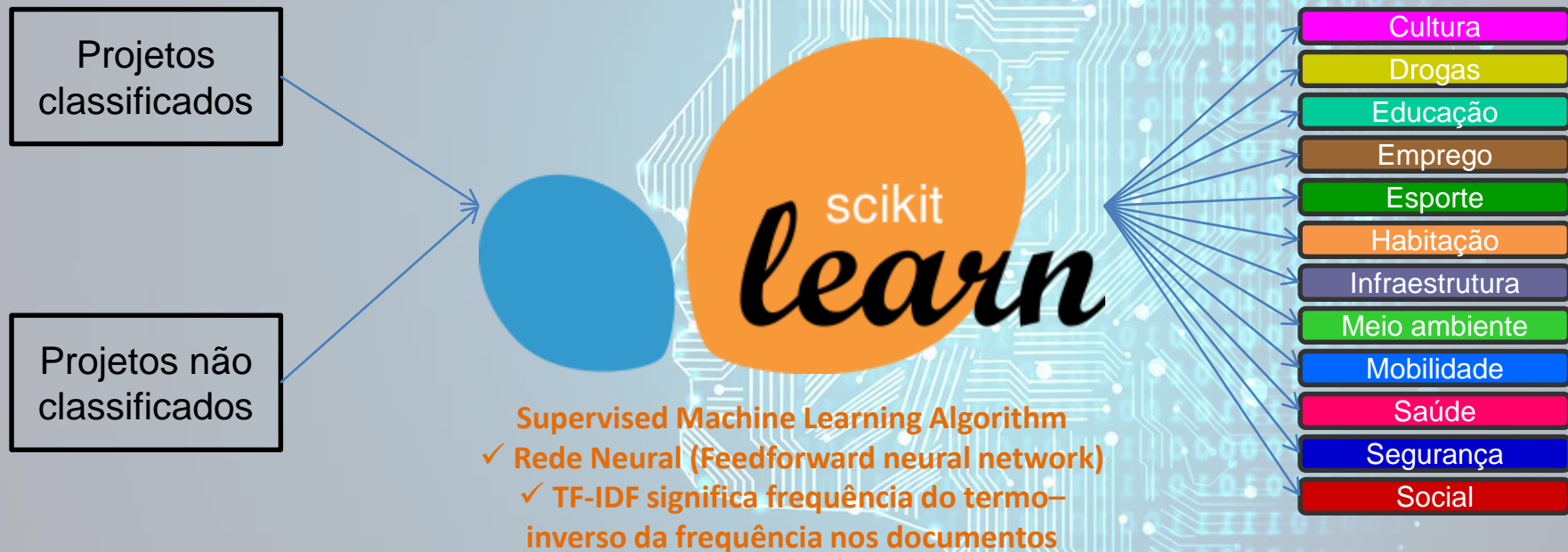


Machine Learning: Meu projeto

O objetivo é identificar textos referente a um projetos proposto pelo poder legislativo.



Machine Learning: Meu projeto



Machine Learning: Exemplo

Dados de treinamento (Corpus)

```
1 def __getTrainingCorpusTest(self):
2     try:
3         print("Learning.__getTrainingCorpusTest")
4         resp = {}
5         resp["corpus"] = {
6             "targetNames": [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26,
7             "trainingData": [
8                 "Estimular a produção de materiais pedagógicos para crianças cegas e com baixa visão, que serão disponibiliza
9                 "Grupos para discussão de prevenção à gravidez indesejada, doenças sexualmente transmissíveis, gestação e men
10            ],
11            "target": {
12                "culture": [0, 0],
13                "drugs": [0, 0],
14                "education": [50, 0],
15                "employment": [0, 0],
16                "sport": [0, 0],
17                "housing": [0, 0],
18                "infrastructure": [0, 0],
19                "environment": [0, 0],
20                "mobility": [0, 0],
21                "cheers": [0, 50],
22                "safety": [0, 0],
23                "social": [50, 50]
24            },
25        }
26
27     except Exception as e:
28         resp["error"] = e
29
30     finally:
31         return resp
32
```

Machine Learning: Exemplo

Gerando vetor a partir dos dados

```
154     def __teachMachine(self, trainingData):
155         try:
156             print("Learning.__teachMachine")
157             resp = {}
158             xTrainTfidfVectorizer = self.vectorizer.fit_transform(trainingData)
159             resp["vector"] = xTrainTfidfVectorizer
160
161         except Exception as e:
162             resp["error"] = e
163
164         finally:
165             return resp
166
```

Machine Learning: Exemplo

Classificando os dados

```
183     def __dataClassifier(self, vector, target):
184         try:
185             print("Learning.__dataClassifier")
186             resp = {}
187             self.clf.fit(vector, target)
188
189         except Exception as e:
190             resp["error"] = e
191
192         finally:
193             return resp
194
```


Machine Learning: Exemplo

Testando os dados

```
195     def __identifyContent(self, data, targetNames):
196         try:
197             print("Learning.__identifyContent")
198             resp = {}
199             xNewTfidfVectorizer = self.vectorizer.transform(data)
200             predicted = self.clf.predict(xNewTfidfVectorizer)
201             resp["result"] = list(zip(data, predicted))
202
203         except Exception as e:
204             resp["error"] = e
205
206         finally:
207             return resp
```

Machine Learning: Demonstração

<http://mydemonstration.com>



Machine Learning: Mercado

Brasil ▼

11/04/2013 – 04/12/2018 ▼

Todas as categorias ▼

Pesquisa na Web ▼

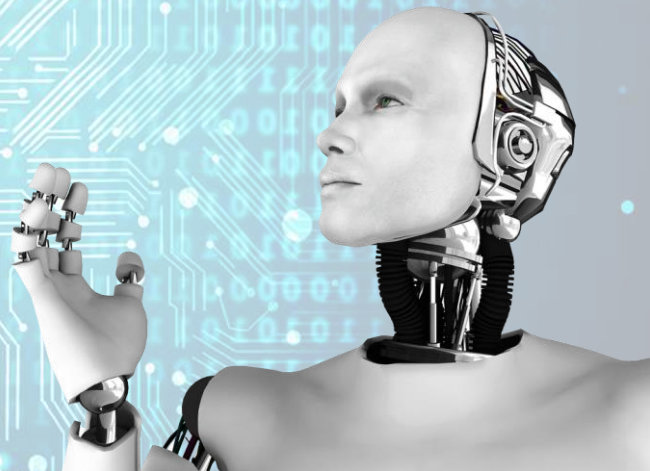
Interesse ao longo do tempo ⓘ



Machine Learning: Conclusão

Além do que foi apresentado, é possível aplicar machine learning em tudo.

É uma forma de potencializar a inteligência humana, e se usada de forma benéfica pode ajudar a civilização a ter uma evolução exponencial.



Machine Learning: Conclusão



Machine Learning: Para saber mais

Scikit-Learn

<https://scikit-learn.org/stable/>

Outras páginas para iniciar estudos

<http://www.cienciaedados.com/conceitos-fundamentais-de-machine-learning/>

<https://medium.com/luisfredgs/classificando-textos-com-machine-learning-e054ca7bf4e0>

<https://medium.com/@dehhmesquita/classificando-textos-com-redes-neurais-e-tensorflow-5063784a1b31>





Muito obrigado!

Ademir Ferreira

Full Stack Developer

ademir.fteo@gmail.com

