

CEIA-POSIA07 - Data Mining

Danilo Sipoli Sanches

Departamento Acadêmico de Computação
Universidade Tecnológica Federal do Paraná
Cornélio Procópio



- Apresentar e aplicar as principais técnicas de mineração de dados;
- Compreender o processo de descoberta de conhecimento (KDD);
- Detectar padrões importantes e não óbvios em grandes bancos de dados;
- Realizar a análise exploratória dos dados;
- Serão abordados bibliotecas em Python.

O que vamos aprender?

- 1 Processo de Descoberta de Conhecimento
- 2 Análise Exploratória de Dados
- 3 Técnicas de Pré-processamento
- 4 Seleção de características
- 5 Classificação e Agrupamento

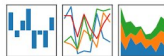
Data Science in Python



Pandas,

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



Scikit-learn, Numpy



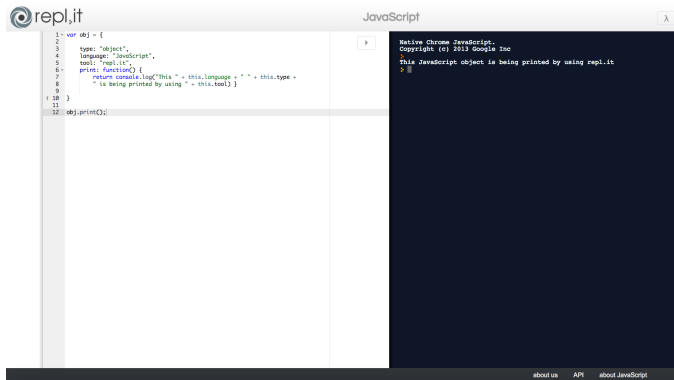
NumPy

Matplotlib



Interpretador de Python (Online)

<https://repl.it/>



The screenshot shows the repl.it online JavaScript interpreter interface. The left pane contains a JavaScript code editor with the following code:

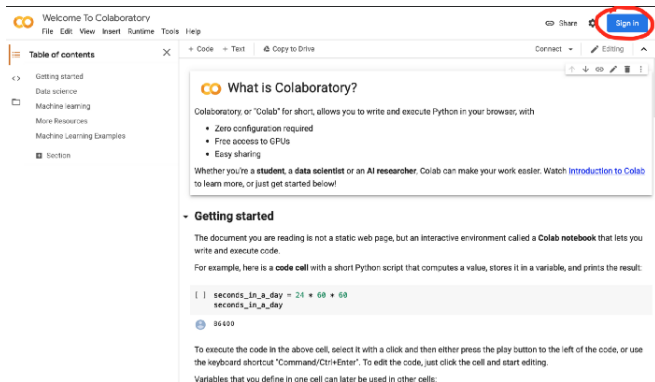
```
1 var obj = {  
2   type: "object",  
3   language: "JavaScript",  
4   tool: "repl.it",  
5   print: function() {  
6     return console.log("This " + this.language + " " + this.type +  
7       " is being printed by using " + this.tool }  
8   }  
9 }  
10  
11  
12 obj.print();
```

The right pane shows the output of the code execution, which is:

```
Native Chrome JavaScript.  
Copyright (c) 2013 Google Inc  
This JavaScript object is being printed by using repl.it  
>
```

At the bottom of the interface, there are links for "about us", "API", and "about JavaScript".

<https://colab.research.google.com/>



Welcome To Colaboratory

File Edit View Insert Runtime Tools Help

Table of contents

- Getting started
- Data science
- Machine learning
- More Resources
- Machine Learning Examples
- Section

What is Colaboratory?

Colaboratory, or "Colab" for short, allows you to write and execute Python in your browser, with

- Zero configuration required
- Free access to GPUs
- Easy sharing

Whether you're a **student**, a **data scientist** or an **AI researcher**, Colab can make your work easier. Watch [Introduction to Colab](#) to learn more, or just get started below!

Getting started

The document you are reading is not a static web page, but an interactive environment called a **Colab notebook** that lets you write and execute code.

For example, here is a **code cell** with a short Python script that computes a value, stores it in a variable, and prints the result:

```
[ ] seconds_in_a_day = 24 * 60 * 60
    seconds_in_a_day
```

36 000

To execute the code in the above cell, select it with a click and then either press the play button to the left of the code, or use the keyboard shortcut "Command/Ctrl+Enter". To edit the code, just click the cell and start editing.

Variables that you define in one cell can later be used in other cells:

kaggle™

