



Estácio

Campus: Rua Manoel João Gonçalves, 410/412 – Alcântara CEP: 24711-080

Curso: Desenvolvimento Full-Stack

Disciplina: RPG0034 - DANDO INTELIGÊNCIA AO SOFTWARE

Turma: 9001

Semestre letivo: 2024.4 FLEX

Integrante:

Nome: Samir Campos Lima

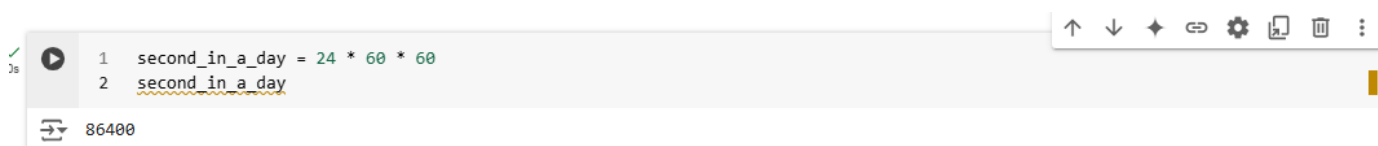
Matrícula: 2022.11.47141-1

Link do repositório no GIT: [samircamposlima/Miss-o-Pr-tica-N-vel-4-Mundo-5](https://github.com/samircamposlima/Miss-o-Pr-tica-N-vel-4-Mundo-5)

Microatividade 1: Configurar o ambiente Google Colab

- Procedimentos

1. Acesse a página do Google Colab: <https://colab.google>
2. No menu superior direito, selecione a opção “Open Colab”;
3. Na nova aba aberta no navegador, faça login (botão no canto superior direito) com uma conta Google;
4. Após o login, feche a janela modal exibida, para visualizar o notebook previamente existente “Olá, este é o Colaboratory”;
5. Navegue pela página e repare que existem blocos de texto seguidos de blocos de código. Esses últimos são caracterizados por possuírem uma cor de fundo diferente, em cinza, e por terem uma setinha que pode ser clicada, o que fará com o código contido no bloco seja executado e seu resultado apresentado a seguir. Veja o print abaixo (onde a setinha que permite a execução foi destacada com um quadrado vermelho):
6. Leia os blocos de texto de cada seção e, a seguir, clique para executar os exemplos de código. Isso lhe permitirá um bom entendimento do funcionamento do Google Colab.



The screenshot shows a code cell in Google Colab. On the left, there is a green checkmark and a play button icon. The code cell contains two lines of Python code: `1 second_in_a_day = 24 * 60 * 60` and `2 second_in_a_day`. The second line is highlighted with a red dashed box. Below the code, the output is displayed as `86400`. On the right side of the code cell, there is a toolbar with icons for undo, redo, insert, link, settings, copy, delete, and a menu icon.

Microatividade 2: Descrever tarefas diversas relacionadas ao Processamento de Linguagem Natural

- Procedimentos:

1. Estando logado no Google Colab, clique no menu “Arquivo” e selecione a opção “Novo notebook”;
2. Na nova aba aberta no navegador, dê um nome ao seu notebook, clicando e alterando o nome automaticamente gerado – Untitled0.ipynb – para nlp.ipynb;
3. Na janela de código, clique na opção “+Texto” (destacada no print abaixo) para inserir um bloco de texto;
4. No bloco de texto, insira um texto que explique o que será executado, a seguir, no bloco de código a ser inserido. Segue uma sugestão, que pode ser complementada posteriormente por você:

```
[ ] 1 Comece a programar ou gere código com IA.
```

Clique duas vezes (ou pressione "Enter") para editar

“Tarefas diversas de Processamento de Linguagem Natural

Passo 1: Instalando a biblioteca e recarregando o ambiente.”

nlp.ipynb

Arquivo Editar Ver Inserir Ambiente de execução Ferramentas Ajuda Todas as alterações foram salvas

+ Código + Texto

1 !pip install spacy

2

3 import pkg_resources,imp

4

5 imp.reload(pkg_resources)

Requirement already satisfied: spacy in /usr/local/lib/python3.10/dist-packages (3.7.5)

Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.0.12)

Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.0.5)

Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.0.11)

Requirement already satisfied: cyemem<2.1.0,>=2.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.0.10)

Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.0.9)

Requirement already satisfied: thinc<8.3.0,>=8.2.2 in /usr/local/lib/python3.10/dist-packages (from spacy) (8.2.5)

Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.1.3)

Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.4.8)

Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.0.10)

Requirement already satisfied: weasel<0.5.0,>=0.1.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (0.4.1)

Requirement already satisfied: typer<1.0.0,>=0.3.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (0.15.0)

Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (4.66.6)

Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.32.3)

Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.10.3)

Requirement already satisfied: Jinja2 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.1.4)

Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from spacy) (75.1.0)

Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (24.2)

Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.5.0)

Requirement already satisfied: numpy>=1.19.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.26.4)

Requirement already satisfied: language-data>=1.2 in /usr/local/lib/python3.10/dist-packages (from langcodes<4.0.0,>=3.2.0->spacy) (1.3.0)

Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.10/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (0.7.0)

Requirement already satisfied: pydantic-core>=2.27.1 in /usr/local/lib/python3.10/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (2.27.1)

Requirement already satisfied: typing-extensions>=4.12.2 in /usr/local/lib/python3.10/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (4.12.2)

Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (3.4.0)

Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2.2.3)

Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2024.8.30)

Requirement already satisfied: blis<0.8.0,>=0.7.8 in /usr/local/lib/python3.10/dist-packages (from thinc<8.3.0,>=8.2.2->spacy) (0.7.11)

Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.10/dist-packages (from thinc<8.3.0,>=8.2.2->spacy) (0.1.5)

Requirement already satisfied: click>=8.0.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy) (8.1.7)

Requirement already satisfied: shellingham>=1.3.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy) (1.5.4)

Requirement already satisfied: rich>=10.11.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy) (13.9.4)

Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from weasel<0.5.0,>=0.1.0->spacy) (0.20.0)

Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/python3.10/dist-packages (from weasel<0.5.0,>=0.1.0->spacy) (7.0.5)

Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from Jinja2->spacy) (3.0.2)

Requirement already satisfied: marisa-trie>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from language-data>=1.2->langcodes<4.0.0,>=3.2.0->spacy) (1.2.1)

Requirement already satisfied: markdown-it-py>=2.2.0 in /usr/local/lib/python3.10/dist-packages (from rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy) (3.0.0)

Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy) (2.18.0)

Requirement already satisfied: wrapt in /usr/local/lib/python3.10/dist-packages (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.1.0->spacy) (1.17.0)

Requirement already satisfied: mdurl==0.1 in /usr/local/lib/python3.10/dist-packages (from markdown-it-py>=2.2.0->rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy) (0.1.2)

<ipython-input-1-51f74c8fa278>:3: DeprecationWarning: pkg_resources is deprecated as an API. See https://setuptools.pypa.io/en/latest/pkg_resources.html

import pkg_resources,imp

<module 'pkg_resources' from '/usr/local/lib/python3.10/dist-packages/pkg_resources/__init__.py'>

Ativar o Windows

Acesse Configurações para ativar o Windows

4s conclusão: 20:10

7. Crie um novo bloco de texto e insira o seguinte conteúdo: “Passo 2: Baixando o modelo a ser utilizado e recarregando o ambiente.”;

8. Crie um novo bloco de código e insira as linhas abaixo:

Passo 2: Baixando o modelo a ser utilizado e recarregando o ambiente.

```
[2] 1 import spacy.cli
    2
    3 spacy.cli.download("en_core_web_trf")
    4
    5 import pkg_resources,imp
    6
    7 imp.reload(pkg_resources)
```

✓ Download and installation successful
You can now load the package via `spacy.load('en_core_web_trf')`
⚠ Restart to reload dependencies
If you are in a Jupyter or Colab notebook, you may need to restart Python in order to load all the package's dependencies. You can do this by selecting the 'Restart kernel' or 'Restart runtime' option.
<module 'pkg_resources' from '/usr/local/lib/python3.10/dist-packages/pkg_resources/__init__.py'>

9. Crie um bloco de texto com o conteúdo: “Passo 3: Importando a biblioteca e definindo o modelo a ser utilizado.”;

10. Crie um bloco de código com as linhas abaixo e o execute

“Passo 3: Importando a biblioteca e definindo o modelo a ser utilizado.”

```
[3] 1 import spacy
    2
    3 nlp = spacy.load('en_core_web_sm')
```

11. Crie um bloco de texto com o conteúdo: “Passo 4: Definindo o texto a ser utilizado nas tarefas de NLP”;

12. Crie um bloco de código, insira as linhas abaixo e o execute:

“Passo 4: Definindo o texto a ser utilizado nas tarefas de NLP”

```
[4] 1 text = """National Park Week starts on Saturday, and it also starts off with a
    2 bang-for-your-buck.
    3
    4 That's because every US National Park Service site will have free entry on
    5 Saturday. NPS manages almost 430 sites, and the majority of them already offer
    6 free entry every day.
    7
    8 But this is your chance to get into the coveted, big-name national parks and
    9 other sites without paying a fee.
    10
    11 That includes legendary parks such as Yosemite, which normally has an entry fee
    12 of $20 per person or $35 per vehicle.
    13
    14 (Note: You'll still need a reservation to drive into Yosemite on weekends and
    15 holidays from April 13 to June 30.)
    16
    17 """
    18
    19 doc = nlp(text)
```

13. Adicione um bloco de texto com o conteúdo “Passo 5: Realizando a tokenização do texto”;
14. Adicione um bloco de código, com as linhas abaixo, e o execute a seguir:

Passo 5: Realizando a tokenização do texto

```
1 for token in doc:
2     |
3     | print(token)
```

legendary
parks
such
as
Yosemite
,
which
normally
has
an
entry
fee

of
\$
20
per
person
or
\$
35
per

Ativar o Windows
Acesse Configurações para ativar o Windows.

Saída do passo 5 na próxima página:

National	National		includes	You
Park	Park		legendary	'll
Week	Service	But	parks	still
starts	site	this	such	need
on	will	is	as	a
Saturday	have	your	Yosemite	reservation
,	free	chance	,	to
and	entry	to	which	drive
it	on	get	normally	into
also	Saturday	into	has	Yosemite
starts	.	the	an	on
off	NPS	coveted	entry	weekends
with	manages	,	fee	and
a	almost	big		
	430	-		
	sites	name	of	holidays
bang	,	national	\$	from
-	and	parks	20	April
for	the	and	per	13
-	majority		person	to
your	of		or	June
-	them	other	\$	30
buck	already	sites	35	.
.	offer	without	per)
		paying	vehicle	
		a	.	
	free	fee		
That	entry	.		
's	every			
because	day		(
every	.		Note	
US		That	:	

15. Adicione um bloco de texto com o conteúdo: “Passo 6: Obtendo as tags de classe gramatical para os tokens individuais.”;
16. Adicione o bloco de código abaixo e o execute:

Passo 6: Obtendo as tags de classe gramatical para os tokens individuais.

```
1 for token in doc:
2
3     # Print the token and the POS tags
4
5     print(token, token.pos_, token.tag_)
```

```
⇒ National PROPN NNP
Park PROPN NNP
Week PROPN NNP
starts VERB VBZ
on ADP IN
Saturday PROPN NNP
, PUNCT ,
and CCONJ CC
it PRON PRP
also ADV RB
starts VERB VBZ
off ADP RP
with ADP IN
a DET DT

SPACE _SP
bang NOUN NN
- PUNCT HYPH
for ADP IN
- PUNCT HYPH
your PRON PRP$
- PUNCT HYPH
```

Saída do passo 6 nas próximas páginas:

National PROPN NNP
Park PROPN NNP
Week PROPN NNP
starts VERB VBZ
on ADP IN
Saturday PROPN NNP
, PUNCT ,
and CCONJ CC
it PRON PRP
also ADV RB
starts VERB VBZ
off ADP RP
with ADP IN
a DET DT

SPACE _SP
bang NOUN NN
- PUNCT HYPH
for ADP IN
- PUNCT HYPH
your PRON PRP\$
- PUNCT HYPH
buck NOUN NN
. PUNCT .

SPACE _SP
That PRON DT
's VERB VBZ
because SCONJ IN

every DET DT
US PROPN NNP
National PROPN NNP
Park PROPN NNP
Service PROPN NNP
site NOUN NN
will AUX MD
have VERB VB
free ADJ JJ
entry NOUN NN
on ADP IN

SPACE _SP
Saturday PROPN NNP
. PUNCT .

NPS PROPN NNP
manages VERB VBZ
almost ADV RB
430 NUM CD
sites NOUN NNS
, PUNCT ,
and CCONJ CC
the DET DT
majority NOUN NN
of ADP IN

them PRON PRP
already ADV RB
offer VERB VBP

SPACE _SP

free ADJ JJ
entry NOUN NN
every DET DT
day NOUN NN
. PUNCT .

SPACE _SP
But CCONJ CC
this PRON DT
is AUX VBZ
your PRON PRP\$
chance NOUN NN
to PART TO
get VERB VB

into ADP IN
the DET DT
coveted ADJ JJ
, PUNCT ,
big ADJ JJ
- PUNCT HYPH
name NOUN NN
national ADJ JJ
parks NOUN NNS
and CCONJ CC

SPACE _SP
other ADJ JJ
sites NOUN NNS
without ADP IN

paying VERB VBG
a DET DT
fee NOUN NN
. PUNCT .

SPACE _SP
That PRON DT
includes VERB VBZ
legendary ADJ JJ
parks NOUN NNS
such ADJ JJ
as ADP IN
Yosemite PROPN NNP
, PUNCT ,
which PRON WDT
normally ADV RB
has VERB VBZ
an DET DT
entry NOUN NN
fee NOUN NN

SPACE _SP
of ADP IN
\$ SYM \$
20 NUM CD
per ADP IN
person NOUN NN
or CCONJ CC
\$ SYM \$
35 NUM CD
per ADP IN
vehicle NOUN NN
. PUNCT .

SPACE _SP
(PUNCT -LRB-
Note NOUN NN
: PUNCT :
You PRON PRP
'll AUX MD
still ADV RB
need VERB VB
a DET DT

reservation NOUN NN
to PART TO
drive VERB VB
into ADP IN
Yosemite PROPN NNP
on ADP IN
weekends NOUN NNS
and CCONJ CC

SPACE _SP
holidays NOUN NNS
from ADP IN
April PROPN NNP
13 NUM CD
to ADP IN
June PROPN NNP
30 NUM CD
. PUNCT .
) PUNCT -RRB-
SPACE _SP

17. Crie o bloco de texto: “Passo 7: Imprimindo os tokens e sua análise morfológica”;
18. Crie o bloco de código abaixo e o execute:

Passo 7: Imprimindo os tokens e sua análise morfológica

```
[7] 1 for token in doc:
    2
    3     print(token, token.morph)
```

legendary Degree=Pos
parks Number=Plur
such Degree=Pos
as
Yosemite Number=Sing
, PunctType=Comm
which
normally
has Mood=Ind|Number=Sing|Person=3|Tense=Pres|VerbForm=Fin
an Definite=Ind|PronType=Art
entry Number=Sing
fee Number=Sing

of
\$
20 NumType=Card
per
person Number=Sing
or ConjType=Cmp
\$
35 NumType=Card
per
vehicle Number=Sing
. PunctType=Peri

Saída do passo 7 nas próximas páginas:

National Number=Sing

Park Number=Sing

Week Number=Sing

starts

Number=Sing | Person=3 | Tense=Pres | VerbForm=Fin

on

Saturday Number=Sing

, PunctType=Comm

and ConjType=Cmp

it

Case=Nom | Gender=Neut | Number=Sing | Person=3 | PronType=Prs

also

starts

Number=Sing | Person=3 | Tense=Pres | VerbForm=Fin

off

with

a Definite=Ind | PronType=Art

bang Number=Sing

- PunctType=Dash

for

- PunctType=Dash

your Person=2 | Poss=Yes | PronType=Prs

- PunctType=Dash

buck Number=Sing

. PunctType=Peri

That Number=Sing | PronType=Dem

's Number=Sing | Person=3 | Tense=Pres | VerbForm=Fin

because

every

US Number=Sing

National Number=Sing

Park Number=Sing

Service Number=Sing

site Number=Sing

will VerbForm=Fin

have VerbForm=Inf

free Degree=Pos

entry Number=Sing

on

Saturday Number=Sing

. PunctType=Peri

NPS Number=Sing

manages

Number=Sing | Person=3 | Tense=Pres | VerbForm=Fin

almost

430 NumType=Card

sites Number=Plur

, PunctType=Comm

and ConjType=Cmp

the Definite=Def | PronType=Art

majority Number=Sing

of

them

Case=Acc | Number=Plur | Person=3 | PronType=Prs

already

offer Tense=Pres | VerbForm=Fin

free Degree=Pos

entry Number=Sing

every

day Number=Sing

. PunctType=Peri

But ConjType=Cmp
this Number=Sing | PronType=Dem
is
Mood=Ind | Number=Sing | Person=3 | Tense=Pres | Verb
Form=Fin
your Person=2 | Poss=Yes | PronType=Prs
chance Number=Sing
to
get VerbForm=Inf
into
the Definite=Def | PronType=Art
coveted Degree=Pos
, PunctType=Comm
big Degree=Pos
- PunctType=Dash
name Number=Sing
national Degree=Pos
parks Number=Plur
and ConjType=Cmp

other Degree=Pos
sites Number=Plur
without
paying Aspect=Prog | Tense=Pres | VerbForm=Part
a Definite=Ind | PronType=Art
fee Number=Sing
. PunctType=Peri

That Number=Sing | PronType=Dem

includes
Number=Sing | Person=3 | Tense=Pres | VerbForm=Fin
legendary Degree=Pos
parks Number=Plur
such Degree=Pos
as
Yosemite Number=Sing
, PunctType=Comm
which
normally
has
Mood=Ind | Number=Sing | Person=3 | Tense=Pres | Verb
Form=Fin
an Definite=Ind | PronType=Art
entry Number=Sing
fee Number=Sing

of
\$
20 NumType=Card
per
person Number=Sing
or ConjType=Cmp
\$
35 NumType=Card
per
vehicle Number=Sing
. PunctType=Peri

(PunctSide=Ini | PunctType=Brck
Note Number=Sing
:
You Case=Nom | Person=2 | PronType=Prs

'll VerbType=Mod

still

need VerbForm=Inf

a Definite=Ind|PronType=Art

reservation Number=Sing

to

drive VerbForm=Inf

into

Yosemite Number=Sing

on

weekends Number=Plur

and ConjType=Cmp

holidays Number=Plur

from

April Number=Sing

13 NumType=Card

to

June Number=Sing

30 NumType=Card

. PunctType=Peri

) PunctSide=Fin|PunctType=Brck

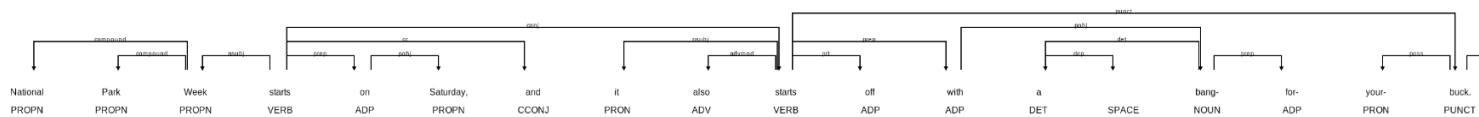
19. Crie o bloco de texto: “Passo 8: Visualizando a árvore de análise sintática por frase e a relação entre as palavras.”;
20. Crie o bloco de código abaixo e o execute:

Passo 8: Visualizando a árvore de análise sintática por frase e a relação entre as palavras.

```
[8] 1 from spacy import displacy
    2
    3 displacy.render(doc, style='dep', options={'compact': True})
```



Saída do passo 8 :



1Clique na imagem para ser encaminhado para o google drive

Microatividade 3:

Descrever o processo de identificação de entidades a partir de textos

- Procedimentos

1. Estando logado no Google Colab, clique no menu “Arquivo” e selecione a opção “Novo notebook”;
2. Na nova aba aberta no navegador, dê um nome ao seu notebook, clicando e alterando o nome automaticamente gerado – Untitled0.ipynb – para ner.ipynb;
3. Na janela de código, clique na opção “+Texto” (destacada no print abaixo) para inserir um bloco de texto;
4. No bloco de texto, insira um texto que explique o que será executado, a seguir, no bloco de código a ser inserido. Segue uma sugestão, que pode ser complementada posteriormente por você:
5. Clique, agora, na opção “+Código” para inserir um bloco de código logo abaixo do bloco de texto criado anteriormente. Nesse bloco de código, insira as linhas abaixo:
6. Execute o código acima;

1-Importando as bibliotecas python e configurando o modelo a ser utilizado pelo

```
1 import pandas as pd
2
3 import spacy
4
5 import requests
6
7 from bs4 import BeautifulSoup
8
9 nlp = spacy.load("en_core_web_sm")
10
11 pd.set_option("display.max_rows", 200)
```

7. Crie um novo bloco de texto e insira o seguinte conteúdo: “Definição do texto a ser analisado e exibição das entidades encontradas”;
8. Crie um novo bloco de código e insira as linhas abaixo:
9. Clique na setinha para executar o código. Como resultado deverá ser exibido, logo abaixo do bloco de código, o resultado da análise de identificação de entidades no texto fornecido, conforme abaixo:

2-Definição do texto a ser analisado e exibição das entidades encontradas

```
[2] 1 doc = nlp("Apple is looking at buying U.K. startup for $1 billion")
    2
    3 for ent in doc.ents:
    4
    5     print(ent.text, ent.start_char, ent.end_char, ent.label_)
```

```
Apple 0 5 ORG
U.K. 27 31 GPE
$1 billion 44 54 MONEY
```

10. Insira um novo bloco de texto com o conteúdo: “Importando a lib displacy e exibindo o resultado da análise NER de forma visual”;
11. Insira um novo bloco de código com as linhas abaixo:
12. Clique na setinha para executar o código. Você deverá ver, abaixo do bloco, o resultado da análise NER de forma visual, com destaque no próprio texto para o tipo de cada entidade identificada;

3-Importando a lib displacy e exibindo o resultado da análise NER de forma visual

```
[3] 1 from spacy import displacy
    2
    3 displacy.render(doc, style="ent")
```



The image shows the output of the NER analysis. The text "Apple is looking at buying U.K. startup for \$1 billion" is displayed with colored boxes highlighting the identified entities: "Apple" (ORG), "U.K." (GPE), and "\$1 billion" (MONEY).

13. Insira um novo bloco de texto com o conteúdo “Analisando um bloco de texto”;
14. Insira um novo bloco de código com as linhas abaixo:
15. Execute o Código. Como resultado você deverá ver o resultado da análise NER sobre o bloco de texto definido e atribuído à variável “sample_txt”;

4-Analisando um bloco de texto

```
1 sample_txt = ""
2
3 Hello Zhang Wei. Your AnyCompany Financial Services, LLC credit card account
4
5 1111-0000-1111-0000 has a minimum payment of $24.53 that is due by July 31st.
6
7 Based on your autopay settings, we will withdraw your payment on the due date
8 from
9
10 your bank account XXXXXX1111 with the routing number XXXXX0000.
11
12
13
14 Your latest statement was mailed to 100 Main Street, Anytown, WA 98121.
15
16 After your payment is received, you will receive a confirmation text message
17
18 at 206-555-0100.
19
20 If you have questions about your bill, AnyCompany Customer Service is available
21 by
22
23 phone at 206-555-0199 or email at support@anycompany.com.
24
25 ""
26
27 newdoc = nlp(sample_txt)
28
29 displacy.render(newdoc, style="ent")
```

Hello **Zhang Wei PERSON** . Your **AnyCompany Financial Services ORG** , **LLC ORG** credit card account

1111-0000-1111-0000 DATE has a minimum payment of \$ **24.53 MONEY** that is due by **July 31st DATE** .

Based on your autopay settings, we will withdraw your payment on the due date
from

your bank account **XXXXXX1111 ORG** with the routing number XXXXX0000.

Your latest statement was mailed to **100 CARDINAL** **Main Street FAC** , **Anytown GPE** , WA 98121.

After your payment is received, you will receive a confirmation text message

at **206-555-0100 CARDINAL** .

If you have questions about your bill, **AnyCompany Customer Service ORG** is available
by

phone at **206-555-0199 CARDINAL** or email at support@anycompany.com.

16. Adicione um novo bloco de texto com o conteúdo “Visualizando o resultado da análise de forma tabular”;
17. Adicione um novo bloco de código com as linhas abaixo:
18. Execute o código. Como resultado, você deverá visualizar o resultado da análise NER de forma tabulada – graças à utilização da lib Pandas, através da qual o resultado da análise foi atribuído a um dataframe.

5-Visualizando o resultado da análise de forma tabular

```
1 entities = [(ent.text, ent.label_, ent.lemma_) for ent in newdoc.ents]
2
3 df = pd.DataFrame(entities, columns=['text', 'type', 'lemma'])
4
5 df.head()
```



	text	type	lemma
0	Zhang Wei	PERSON	Zhang Wei
1	AnyCompany Financial Services	ORG	AnyCompany Financial Services
2	LLC	ORG	LLC
3	1111-0000-1111-0000	DATE	1111-0000-1111-0000
4	24.53	MONEY	24.53

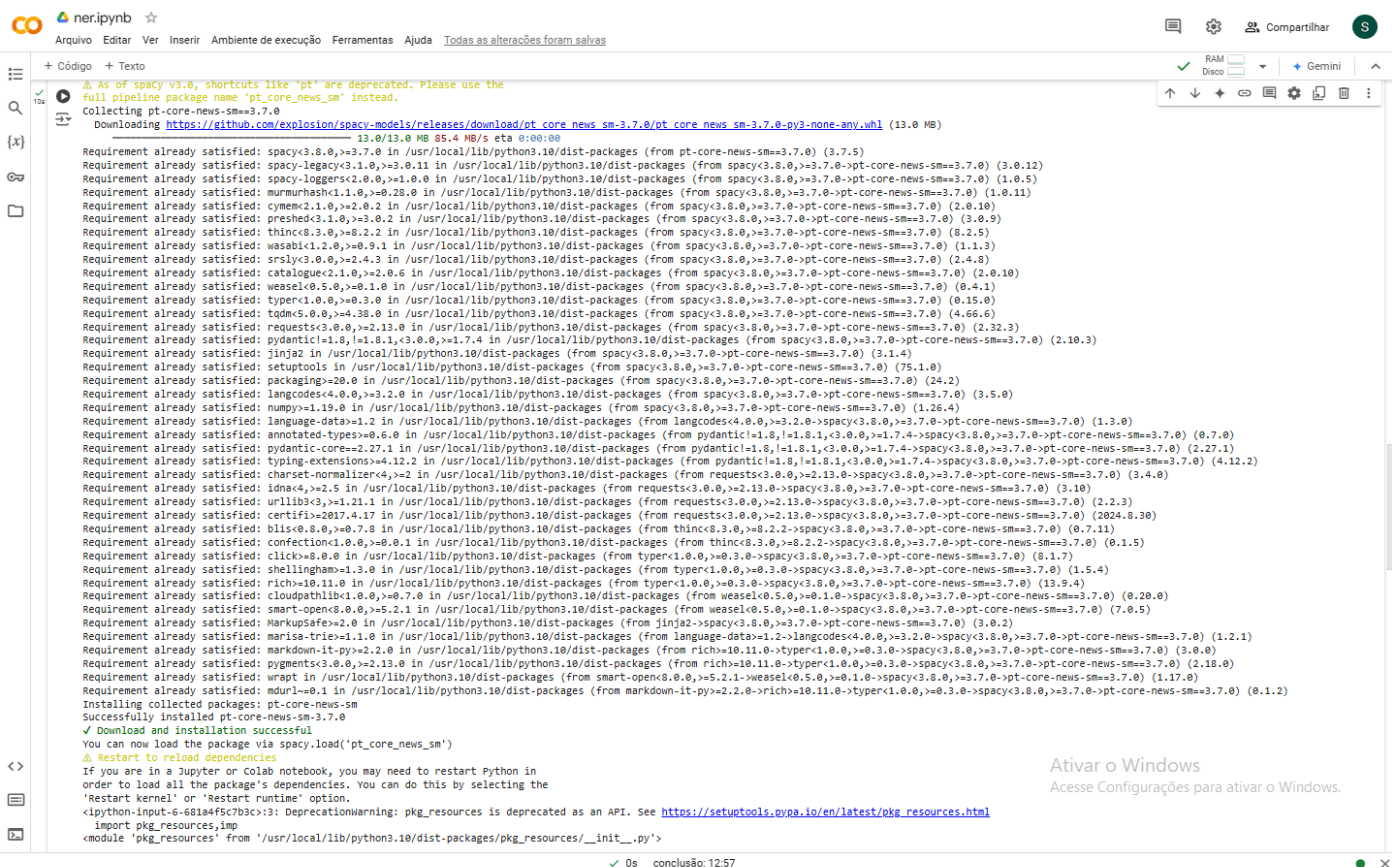
Próximas etapas:

[Gerar código com df](#)[Ver gráficos recomendados](#)[New interactive sheet](#)

19. Crie um novo bloco de texto com o conteúdo: “Utilizando modelos no idioma português”;
20. Adicione um novo bloco de código com as linhas abaixo e o execute na sequência:

6-Utilizando modelos no idioma português

```
1 !python -m spacy download pt
2
3 import pkg_resources,imp
4
5 imp.reload(pkg_resources)
```



```
ner.ipynb
Arquivo  Editar  Ver  Inserir  Ambiente de execução  Ferramentas  Ajuda  Todas as alterações foram salvas

+ Código  + Texto

A As of spacy v3.0, shortcuts like "pt" are deprecated. Please use the
full pipeline package name "pt_core_news_sm" instead.
Collecting pt-core-news-sm==3.7.0
  Downloading https://github.com/explosion/spacy-models/releases/download/pt_core_news_sm-3.7.0/pt_core_news_sm-3.7.0-py3-none-any.whl (13.0 MB)
    13.0/13.0 MB 85.4 MB/s eta 0:00:00
Requirement already satisfied: spacy<3.8.0,>=3.7.0 in /usr/local/lib/python3.10/dist-packages (from pt-core-news-sm==3.7.0) (3.7.5)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.20.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (1.0.11)
Requirement already satisfied: cytoolz<2.1.0,>=2.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (2.0.10)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (3.0.9)
Requirement already satisfied: thinc<8.3.0,>=8.2.2 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (8.2.5)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (2.4.0)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.1.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (0.4.1)
Requirement already satisfied: typer<1.0.0,>=0.3.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (0.15.0)
Requirement already satisfied: tqdm<5.0.0,>=4.30.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (4.66.6)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (2.32.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (2.4.0)
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (75.1.0)
Requirement already satisfied: packaging<20.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (24.2)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (3.5.0)
Requirement already satisfied: numpy<=1.19.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (1.26.4)
Requirement already satisfied: language-data<=1.2 in /usr/local/lib/python3.10/dist-packages (from langcodes<4.0.0,>=3.2.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (1.3.0)
Requirement already satisfied: pydantic<1.8.1,>=1.8.1 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (0.7.0)
Requirement already satisfied: pydantic-core<=2.27.1 in /usr/local/lib/python3.10/dist-packages (from pydantic<1.8.1,>=1.8.1,>=1.7.4->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (2.27.1)
Requirement already satisfied: typing-extensions<=4.12.2 in /usr/local/lib/python3.10/dist-packages (from pydantic<1.8.1,>=1.8.1,>=1.7.4->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (4.12.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (3.4.0)
Requirement already satisfied: idna<=2.5 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (2.2.3)
Requirement already satisfied: certifi<=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (2024.8.30)
Requirement already satisfied: blis<0.8.0,>=0.7.8 in /usr/local/lib/python3.10/dist-packages (from thinc<8.3.0,>=8.2.2->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (0.7.11)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.10/dist-packages (from thinc<8.3.0,>=8.2.2->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (0.1.5)
Requirement already satisfied: click<=8.0.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (8.1.7)
Requirement already satisfied: shellingham<1.3.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (1.5.4)
Requirement already satisfied: rich<=10.11.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (13.9.4)
Requirement already satisfied: cloudpathlib<1.0.0,>=7.0 in /usr/local/lib/python3.10/dist-packages (from weasel<0.5.0,>=0.1.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (0.20.0)
Requirement already satisfied: smart-open<6.0.0,>=5.2.1 in /usr/local/lib/python3.10/dist-packages (from weasel<0.5.0,>=0.1.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (7.0.5)
Requirement already satisfied: MarkupSafe<=2.0 in /usr/local/lib/python3.10/dist-packages (from Jinja2->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (3.0.2)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from rich<=10.11.0->typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (1.2.1)
Requirement already satisfied: markdown-it-py<=2.2.0 in /usr/local/lib/python3.10/dist-packages (from rich<=10.11.0->typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (3.0.0)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from rich<=10.11.0->typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (2.18.0)
Requirement already satisfied: wrapt in /usr/local/lib/python3.10/dist-packages (from smart-open<6.0.0,>=5.2.1->weasel<0.5.0,>=0.1.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (1.17.0)
Requirement already satisfied: mdurl<=0.1 in /usr/local/lib/python3.10/dist-packages (from markdown-it-py<=2.2.0->rich<=10.11.0->typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->pt-core-news-sm==3.7.0) (0.1.2)
Installing collected packages: pt-core-news-sm
Successfully installed pt-core-news-sm-3.7.0
✓ Download and installation successful
You can now load the package via spacy.load('pt_core_news_sm')

Restart to reload dependencies
If you are in a Jupyter or Colab notebook, you may need to restart Python in
order to load all the package's dependencies. You can do this by selecting the
'Restart kernel' or 'Restart runtime' option.
<ipython-input-6-681a4f5c7b3c>:3: DeprecationWarning: pkg_resources is deprecated as an API. See https://setuptools.pypa.io/en/latest/pkg_resources.html
import pkg_resources,imp
<module 'pkg_resources' from '/usr/local/lib/python3.10/dist-packages/pkg_resources/__init__.py'>

0s  conclusão: 12:57
```

21. Adicione um novo bloco de código com as seguintes linhas:

22. Execute o bloco de código. Como resultado você deverá ver a análise de entidades de forma visual.

23. Ao final da atividade, caso queira, você poderá salvar uma cópia do código no Google Drive ou no Github. Tais opções encontram-se disponíveis a partir do menu Arquivo.

ner.ipynb

Arquivo Editar Ver Inserir Ambiente de execução Ferramentas Ajuda Todas as alterações foram salvas

+ Código + Texto

1 import spacy

2

3 import requests

4

5 from bs4 import BeautifulSoup

6

7 from spacy import displacy

8

9

10

11 nlp = spacy.load("en_core_web_sm")

12

13

14

15 txt_br=""

16

17 ONU aprova missão internacional para restabelecer segurança no Haiti

18

19 O Conselho de Segurança das Nações Unidas (ONU) aprovou na noite de

20 segunda-feira (2) a criação e o envio de uma força internacional para a

21 manutenção de paz no Haiti, devido aos conflitos entre as gangues que dominam

22 o país.

23

24 O Haiti é o país mais pobre do hemisfério ocidental. A história do país é marcada

25 por golpes, deposições e massacres que geraram grande instabilidade política,

26 turbulência econômica e crise social. Em 30 de abril de 2004 o Conselho de

27 Segurança da Organização das Nações Unidas aprovou, por unanimidade, a

28 criação da Missão de Estabilização do Haiti, a MINUSTAH.

29

30 A missão, planejada para ter uma duração inicial de seis meses, vem sendo

31 prorrogada. O objetivo é combater a insegurança no país após a crise que forçou

32 a saída do ex-presidente Jean Bertrand Aristide, em fevereiro de 2004.

33

34 Coube ao Brasil a chefia da missão, instalada no mês de junho de 2004, bem

35 como constituir o maior contingente nacional de "boinas azuis" : 1.470 militares,

36 de um total de 6.700, militares de 21 países. Apesar das dificuldades estruturais

37 do Haiti, as atividades da MINUSTAH trouxeram avanços, como a realização das

38 eleições ocorridas em fevereiro deste ano.

39

40 ""

41

42 docbr = nlp(txt_br)

43

44

45

46 displacy.render(docbr, style="ent")

Ativar o Windows
Acesse Configurações para ativar o Windows.

0s conclusão: 12:57

O Conselho de Segurança PERSON das Nações Unidas (ONU) aprovou na PERSON noite de segunda-feira (2 CARDINAL) a criação e o envio de uma força internacional para a manutenção de paz no Haiti GPE , devido aos conflitos PERSON entre as gangues que dominam o país.

O Haiti GPE é o país mais pobre do hemisfério ocidental. A história GPE do país é marcada por golpes, deposições e massacres que geraram grande instabilidade política, turbulência econômica e crise social. Em 30 de abril de QUANTITY 2004 DATE o Conselho de GPE Segurança da Organização PERSON das Nações Unidas ORG aprovou, por unanimidade, a criação da Missão de Estabilização do Haiti GPE , a MINUSTAH ORG .

A missão, planejada para PERSON ter uma duração inicial de seis meses, vem sendo prorrogada. O objetivo é combater a insegurança no país após a crise que forçou a saída do ex-presidente Jean Bertrand Aristide PRODUCT , em fevereiro de 2004.

Coube ORG ao Brasil a chefia da missão, instalada GPE no mês de ORG junho de 2004 DATE , bem como constituir o maior contingente nacional de "boinas azuis": 1.470 militares QUANTITY , de um total de 6.700 CARDINAL , militares CARDINAL de 21 CARDINAL países. Apesar das dificuldades estruturais PERSON do Haiti GPE , as atividades da MINUSTAH ORG trouxeram avanços ORG , como a realização das eleições ocorridas em fevereiro GPE deste ano.

Microatividade 4:

Descrever o processo de extração de frases-chave a partir de textos

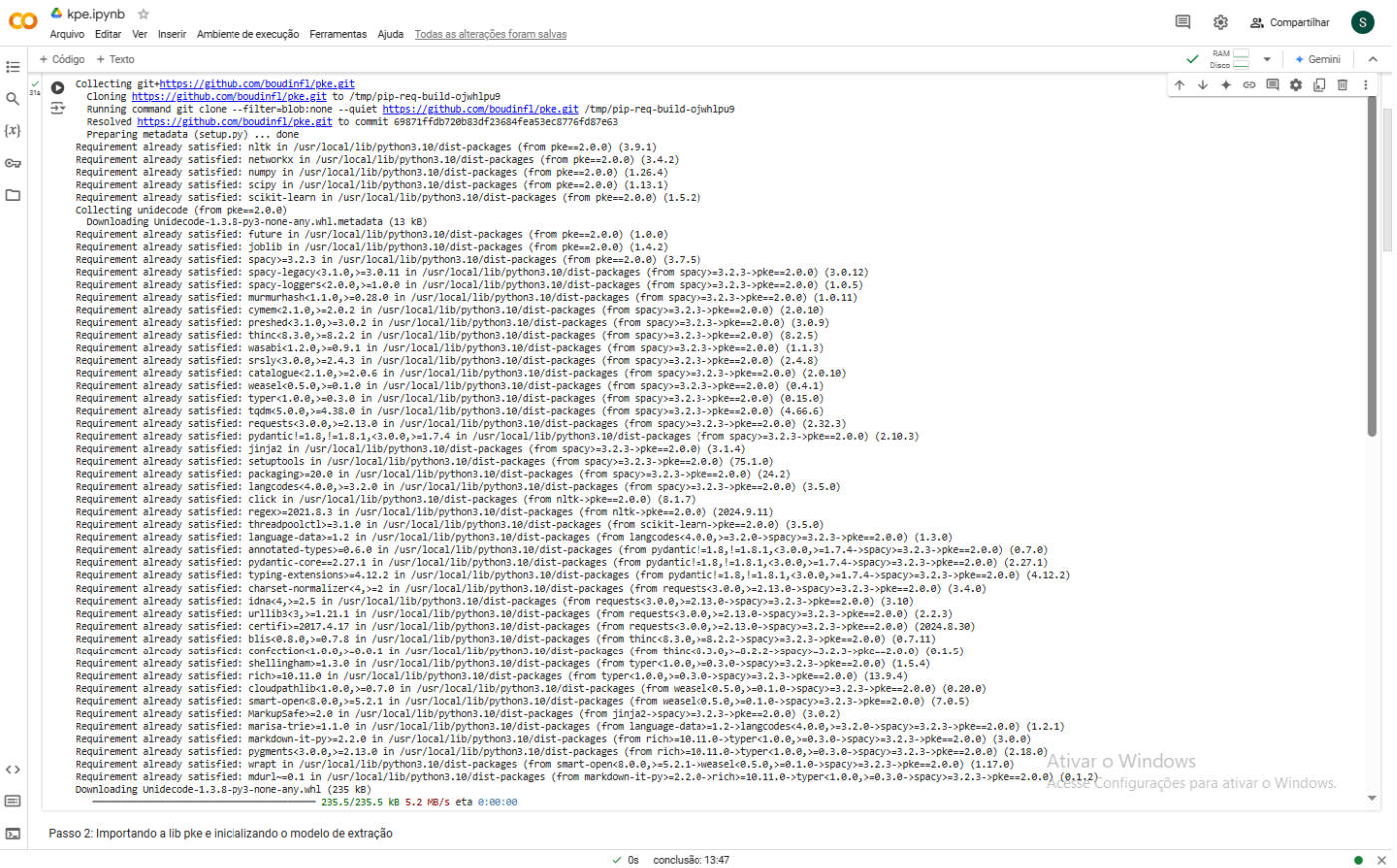
- Procedimentos

1. Estando logado no Google Colab, clique no menu “Arquivo” e selecione a opção “Novo notebook”;
2. Na nova aba aberta no navegador, dê um nome ao seu notebook, clicando e alterando o nome automaticamente gerado – Untitled0.ipynb – para kpe.ipynb;
3. Na janela de código, clique na opção “+Texto” (destacada no print abaixo) para inserir um bloco de texto;
4. No bloco de texto, insira um texto que explique o que será executado, a seguir, no bloco de código a ser inserido. Segue uma sugestão, que pode ser complementada posteriormente por você:
5. Clique, agora, na opção “+Código” para inserir um bloco de código logo abaixo do bloco de texto criado anteriormente. Nesse bloco de código, insira as linhas abaixo:
6. Execute o código acima;

Extração de frases-chave a partir de textos.

Passo 1: Instalação das bibliotecas..

```
[1] 1 !pip install git+https://github.com/boudinfl/pke.git
    2
    3 !python -m spacy download en_core_web_sm
    4
    5 import pkg_resources,imp
    6
    7 imp.reload(pkg_resources)
```

7. Crie um bloco de texto com o texto: “Passo 2: Importando a lib pke e inicializando o modelo de extração”;
8. Crie um bloco de código, insira as linhas abaixo e o execute a seguir

Passo 2: Importando a lib pke e inicializando o modelo de extração

```
1 import pke
2
3 # initialize a TopicRank keyphrase extraction model
4
5 extractor = pke.unsupervised.TopicRank()
```

9. Crie um bloco de texto com o conteúdo: “Passo 3: Definindo e carregando o texto a ser analisado.”
10. Crie um bloco de código, insira as linhas a seguir e o execute na sequência

Passo 3: Definindo e carregando o texto a ser analisado.

```
[3] 1 sample = """Tesla has been ordered to recall nearly 4,000 of its Cybertrucks due
2     to an accelerator pedal that
3
4     can stick in place when pressed down.
5
6     The cause, according to the regulator: soap.
7
8     “An unapproved change introduced lubricant (soap) to aid in the component
9     assembly of the pad onto the accelerator pedal.
10
11    Residual lubricant reduced the retention of the pad to the pedal,” the NHTSA
12    wrote in the recall document.
13
14    Tesla has yet to detail how many of the futuristic looking Cybertrucks it has
15    produced. But it has said that it would
16
17    be slow ramping up production of the vehicle, which had its first deliveries in late
18    November.
19
20    The NHTSA said the recall affects “all Model Year (‘MY’) 2024 Cybertruck
21    vehicles manufactured from November 13, 2023, to April 4, 2024.”
22
23    """.replace("\n", " ")
24
25 extractor.load_document(input=sample, language='en')
```

11. Crie um bloco de texto com o conteúdo: “Passo 4: Imprimindo as informações das sentenças”;
12. Crie um bloco de código com as linhas abaixo e o execute

Passo 4: Imprimindo as informações das sentenças

```
[5] 1 for i, sentence in enumerate(extractor.sentences):  
2  
3  
4  
5     # print out the sentence id, its tokens, its stems and the corresponding Part-of-Speech tags  
6  
7     print("sentence {}: ".format(i))  
8  
9     print(" - words: {} ...".format(' '.join(sentence.words[:5])))  
10  
11    print(" - stems: {} ...".format(' '.join(sentence.stems[:5])))  
12  
13    print(" - PoS: {} ...".format(' '.join(sentence.pos[:5])))
```

```
→ sentence 0:  
- words: Tesla has been ordered to ...  
- stems: tesla ha been order to ...  
- PoS: PROPN AUX AUX VERB PART ...  
sentence 1:  
- words: The cause , according ...  
- stems: the caus , accord ...  
- PoS: SPACE DET NOUN PUNCT VERB ...  
sentence 2:  
- words: " An unapproved change ...  
- stems: " an unapprov chang ...  
- PoS: SPACE PUNCT DET ADJ NOUN ...  
sentence 3:  
- words: Residual lubricant reduced the ...  
- stems: residu lubric reduc the ...  
- PoS: SPACE ADJ NOUN VERB DET ...  
sentence 4:  
- words: Tesla has yet to ...  
- stems: tesla ha yet to ...  
- PoS: SPACE PROPN AUX ADV PART ...  
sentence 5:  
- words: But it has said that ...  
- stems: but it ha said that ...  
- PoS: CCONJ PRON AUX VERB SCONJ ...  
sentence 6:  
- words: The NHTSA said the ...  
- stems: the nhtsa said the ...  
- PoS: SPACE DET PROPN VERB DET ...  
sentence 7:  
- words: ...  
- stems: ...  
- PoS: SPACE ...
```

13. Adicione um bloco de texto com o conteúdo: “Passo 5: Identificando as frases-chave candidatas”;

14. Insira um bloco de código com as linhas abaixo e execute a seguir

Passo 5: Identificando as frases- chave candidatas

```
1  extractor.candidate_selection()
2
3  for i, candidate in enumerate(extractor.candidates):
4
5
6
7      # print out the candidate id, its stemmed form
8
9      print("candidate {}: {} (stemmed form)".format(i, candidate))
10
11
12
13     # print out the surface forms of the candidate
14
15     print(" - surface forms:", [ " ".join(u) for u in
16 extractor.candidates[candidate].surface_forms])
17
18
19
20     # print out the corresponding offsets
21
22     print(" - offsets:", extractor.candidates[candidate].offsets)
23
24
25
26     # print out the corresponding sentence ids
27
28     print(" - sentence_ids:", extractor.candidates[candidate].sentence_ids)
29
30
31
32     # print out the corresponding PoS patterns
33
34     print(" - pos_patterns:", extractor.candidates[candidate].pos_patterns)
```

Saída do passo 5 nas próximas páginas:

candidate 0: tesla (stemmed form)

- surface forms: ['Tesla', 'Tesla']
- offsets: [0, 84]
- sentence_ids: [0, 4]
- pos_patterns: [['PROPN'], ['PROPN']]

candidate 1: cybertruck (stemmed form)

- surface forms: ['Cybertrucks', 'Cybertrucks']
- offsets: [10, 95]
- sentence_ids: [0, 4]
- pos_patterns: [['PROPN'], ['NOUN']]

candidate 2: acceler pedal (stemmed form)

- surface forms: ['accelerator pedal', 'accelerator pedal']
- offsets: [14, 58]
- sentence_ids: [0, 2]
- pos_patterns: [['NOUN', 'NOUN'], ['NOUN', 'NOUN']]

candidate 3: place (stemmed form)

- surface forms: ['place']
- offsets: [21]
- sentence_ids: [0]
- pos_patterns: [['NOUN']]

candidate 4: caus (stemmed form)

- surface forms: ['cause']
- offsets: [28]
- sentence_ids: [1]
- pos_patterns: [['NOUN']]

candidate 5: regul (stemmed form)

- surface forms: ['regulator']
- offsets: [33]

- sentence_ids: [1]

- pos_patterns: [['NOUN']]

candidate 6: soap (stemmed form)

- surface forms: ['soap', 'soap']
- offsets: [35, 45]
- sentence_ids: [1, 2]
- pos_patterns: [['NOUN'], ['NOUN']]

candidate 7: unapprov chang (stemmed form)

- surface forms: ['unapproved change']
- offsets: [40]
- sentence_ids: [2]
- pos_patterns: [['ADJ', 'NOUN']]

candidate 8: lubric (stemmed form)

- surface forms: ['lubricant']
- offsets: [43]
- sentence_ids: [2]
- pos_patterns: [['NOUN']]

candidate 9: compon assembl (stemmed form)

- surface forms: ['component assembly']
- offsets: [51]
- sentence_ids: [2]
- pos_patterns: [['NOUN', 'NOUN']]

candidate 10: pad (stemmed form)

- surface forms: ['pad', 'pad']
- offsets: [55, 69]
- sentence_ids: [2, 3]
- pos_patterns: [['NOUN'], ['NOUN']]

candidate 11: residu lubric (stemmed form)

- surface forms: ['Residual lubricant']

- offsets: [62]
- sentence_ids: [3]
- pos_patterns: [['ADJ', 'NOUN']]

candidate 12: retent (stemmed form)

- surface forms: ['retention']
- offsets: [66]
- sentence_ids: [3]
- pos_patterns: [['NOUN']]

candidate 13: pedal (stemmed form)

- surface forms: ['pedal']
- offsets: [72]
- sentence_ids: [3]
- pos_patterns: [['NOUN']]

candidate 14: nhtsa (stemmed form)

- surface forms: ['NHTSA', 'NHTSA']
- offsets: [76, 128]
- sentence_ids: [3, 6]
- pos_patterns: [['PROPN'], ['PROPN']]

candidate 15: recal document (stemmed form)

- surface forms: ['recall document']
- offsets: [80]
- sentence_ids: [3]
- pos_patterns: [['NOUN', 'NOUN']]

candidate 16: futurist (stemmed form)

- surface forms: ['futuristic']
- offsets: [93]
- sentence_ids: [4]
- pos_patterns: [['ADJ']]

candidate 17: slow (stemmed form)

- surface forms: ['slow']

- offsets: [109]

- sentence_ids: [5]

- pos_patterns: [['ADJ']]

candidate 18: product (stemmed form)

- surface forms: ['production']
- offsets: [112]
- sentence_ids: [5]
- pos_patterns: [['NOUN']]

candidate 19: vehicl (stemmed form)

- surface forms: ['vehicle']
- offsets: [115]
- sentence_ids: [5]
- pos_patterns: [['NOUN']]

candidate 20: late novemb (stemmed form)

- surface forms: ['late November']
- offsets: [123]
- sentence_ids: [5]
- pos_patterns: [['ADJ', 'PROPN']]

candidate 21: recal (stemmed form)

- surface forms: ['recall']
- offsets: [131]
- sentence_ids: [6]
- pos_patterns: [['NOUN']]

candidate 22: model year (stemmed form)

- surface forms: ['Model Year']
- offsets: [135]
- sentence_ids: [6]
- pos_patterns: [['PROPN', 'PROPN']]

candidate 23: cybertruck vehicl (stemmed form)

- surface forms: ['Cybertruck vehicles']
- offsets: [143]
- sentence_ids: [6]
- pos_patterns: [['PROPN', 'NOUN']]

candidate 24: novemb (stemmed form)

- surface forms: ['November']
- offsets: [147]
- sentence_ids: [6]
- pos_patterns: [['PROPN']]

candidate 25: april (stemmed form)

- surface forms: ['April']
- offsets: [153]
- sentence_ids: [6]
- pos_patterns: [['PROPN']]

15. Crie um novo bloco de texto com o conteúdo: “Passo 6: Ranqueando as palavras-chave candidatas”;
16. Crie um bloco de código, adicione as linhas abaixo e o execute a seguir:
17. Tendo finalizado os passos acima, você poderá, entre outras atividades extras, refazer os passos alterando o texto analisado e visualizando o resultado da análise de identificação de frases-chave sobre o mesmo;
18. Ao final da atividade, caso queira, você poderá salvar uma cópia do código no Google Drive ou no Github. Tais opções encontram-se disponíveis a partir do menu Arquivo.

Passo 6: Ranqueando as palavras-chave candidatas

```
1 extractor.candidate_weighting()
2
3 for i, topic in enumerate(extractor.topics):
4
5
6
7     # print out the topic id and the candidates it groups together
8
9     print("topic {}: {}".format(i, ';'.join(topic)))
```

```
topic 0: cybertruck;cybertruck vehicl
topic 1: vehicl
topic 2: lubric;residu lubric
topic 3: recal;recal document
topic 4: acceler pedal;pedal
topic 5: late novemb;novemb
topic 6: april
topic 7: caus
topic 8: compon assembl
topic 9: futurist
topic 10: model year
topic 11: nhtsa
topic 12: pad
topic 13: place
topic 14: product
topic 15: regul
topic 16: retent
topic 17: slow
topic 18: soap
topic 19: tesla
topic 20: unapprov chang
```

Microatividade 5:

Descrever o processo de identificação de linguagem predominante a partir de textos

- Procedimentos

1. Estando logado no Google Colab, clique no menu “Arquivo” e selecione a opção “Novo notebook”;
2. Na nova aba aberta no navegador, dê um nome ao seu notebook, clicando e alterando o nome automaticamente gerado – Untitled0.ipynb – para lang.ipynb;
3. Na janela de código, clique na opção “+Texto” (destacada no print abaixo) para inserir um bloco de texto;
4. No bloco de texto, insira um texto que explique o que será executado, a seguir, no bloco de código a ser inserido. Segue uma sugestão, que pode ser complementada posteriormente por você:
5. Clique, agora, na opção “+Código” para inserir um bloco de código logo abaixo do bloco de texto criado anteriormente. Nesse bloco de código, insira as linhas abaixo:
6. Execute o código acima;

Identificação de idioma predominante em textos.

Passo 1: Instalação das bibliotecas.

```
✓ [1] 1 |pip install langdetect
8s

Collecting langdetect
  Downloading langdetect-1.0.9.tar.gz (981 kB)
    981.5/981.5 kB 13.2 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Requirement already satisfied: six in /usr/local/lib/python3.10/dist-packages (from langdetect) (1.16.0)
Building wheels for collected packages: langdetect
  Building wheel for langdetect (setup.py) ... done
  Created wheel for langdetect: filename=langdetect-1.0.9-py3-none-any.whl size=993222 sha256=9a592903d5b61703c644d08f7f68b8f6ea8dad06aac922950214831e79d7c0fe
  Stored in directory: /root/.cache/pip/wheels/95/03/7d/59ea870c70ce4e5a370638b5462a7711ab78fba2f655d05106
Successfully built langdetect
Installing collected packages: langdetect
Successfully installed langdetect-1.0.9
```

7. Crie um bloco de texto com o texto: “Passo 2: Definindo as frases, em diferentes idiomas, que serão analisadas.”;

8. Crie um bloco de código, insira as linhas abaixo e o execute a seguir

Passo 2: Definindo as frases, em diferentes idiomas, que serão analisadas.

```
1 text = ["Wear masks, keep distance, wash hands, be safe in these difficult days." ,
2
3         "Viseljen maszkot, tartson távolságot, mosson kezét, legyen biztonságban ezekben a nehéz napokban",
4
5         "Deaths are increasing, be vigilant.",
6
7         "Носите маски, соблюдайте дистанцию, мойте руки, будьте осторожны в эти тяжелые дни.",
8
9         "Covid-19: Indians flock to vaccination centers as vaccines are now available for 60+ in India since the 1st of March",
10
11        "Indossa maschere, mantieni le distanze, lavati le mani, sii al sicuro in questi giorni difficili.",
12
13        "Portez des masques, gardez vos distances, lavez-vous les mains, soyez en sécurité en ces jours difficiles.",
14
15        "Brug masker, hold afstand, vask hænder, vær sikker i disse vanskelige dage.",
16
17        "We are facing a global education crisis. No effort should be spared to safely bring every child back into the classroom.",
18
19        "Bruk masker, hold avstand, vask hendene, vær trygg i disse vanskelige dagene.",
20
21        "Portu maskojn, tenu distancon, lavu manojn, estu sekuraj en ĉi tiuj malfacilaj tagoj.",
22
23        "Tragen Sie Masken, halten Sie Abstand, waschen Sie Ihre Hände, seien Sie in diesen schwierigen Tagen sicher.",
24
25        "Носіть маски, тримайтеся на відстані, мийте руки, будьте в безпеці в ці важкі дні.",
26
27        "Lock down , working from home are the keys words for these days.",
28
29        "Lives have changed drastically across the planet and this period will forever be remembered as the beginning of something we have yet to witness.",
30
31        "Este é um exemplo de texto escrito em português."]
```

9. Crie um bloco de texto com o conteúdo: “Passo 3: Recarragendo o ambiente python após a instalação da lib.”

10. Crie um bloco de código, insira as linhas a seguir e o execute na sequência

Passo 3: Recarregando o ambiente python após a instalação da lib.

```
1 import pkg_resources, imp
2
3 imp.reload(pkg_resources)
```

<ipython-input-5-65ce4e9120e8>:1: DeprecationWarning: pkg_resources is deprecated as an API. See https://setuptools.pypa.io/en/latest/pkg_resources.html
<module 'pkg_resources' from '/usr/local/lib/python3.10/dist-packages/pkg_resources/__init__.py'>

11. Crie um bloco de texto com o conteúdo: “Passo 4: Detectando a linguagem predominante na lista de frases.”;
12. Crie um bloco de código com as linhas abaixo e o execute
13. Ao final dessa primeira roda de execução, altere a lista de frases, inserindo textos em outros idiomas para verificar se os mesmos são identificados corretamente;
14. Por fim, caso queira, você poderá salvar uma cópia do código no Google Drive ou no Github. Tais opções encontram-se disponíveis a partir do menu Arquivo.

Passo 4: Detectando a linguagem predominante na lista de frases.

```
1 from langdetect import detect
2
3 for x in text:
4
5     print ('Frase: ', x)
6
7     print ('Idioma: ', detect(x), '\n\n')
```

Saída do passo 4 :

Frase: Wear masks, keep distance, wash hands, be safe in these difficult days.

Idioma: en

Frase: Viseljen maszkot, tartson távolságot, mosson kezet, legyen biztonságban ezekben a nehéz napokban

Idioma: hu

Frase: Deaths are increasing, be vigilant.

Idioma: en

Frase: Носите маски, соблюдайте дистанцию, мойте руки, будьте осторожны в эти тяжелые дни.

Idioma: ru

Frase: Covid-19: Indians flock to vaccination centers as vaccines are now available for 60+ in India since the 1st of March

Idioma: en

Frase: Indossa maschere, mantieni le distanze, lavati le mani, sii al sicuro in questi giorni difficili.

Idioma: it

Frase: Portez des masques, gardez vos distances, lavez-vous les mains, soyez en sécurité en ces jours difficiles.

Idioma: fr

Frase: Brug masker, hold afstand, vask hænder, vær sikker i disse vanskelige dage.

Idioma: da

Frase: We are facing a global education crisis. No effort should be spared to safely bring every child back into the classroom.

Idioma: en

Frase: Bruk masker, hold avstand, vask hendene, vær trygg i disse vanskelige dagene.

Idioma: da

Frase: Portu maskojn, tenu distancon, lavu manojn, estu sekuraj en ĉi tiuj malfacilaj tagoj.

Idioma: hr

Frase: Tragen Sie Masken, halten Sie Abstand, waschen Sie Ihre Hände, seien Sie in diesen schwierigen Tagen sicher.

Idioma: de

Frase: Носіть маски, тримайтеся на відстані, мийте руки, будьте в безпеці в ці важкі дні.

Idioma: uk

Frase: Lock down , working from home are the keys words for these days.

Idioma: en

Frase: Lives have changed drastically across the planet and this period will forever be remembered as the beginning of something we have yet to witness.

Idioma: en

Frase: Este é um exemplo de texto escrito em português.

Idioma: pt

Dando Inteligência ao Software

- Procedimentos

1. Estando logado no Google Colab, clique no menu “Arquivo” e selecione a opção “Novo notebook”;
2. Na nova aba aberta no navegador, dê um nome ao seu notebook, clicando e alterando o nome automaticamente gerado – Untitled0.ipynb – para sentiment.ipynb;
3. Na janela de código, clique na opção “+Texto” (destacada no print abaixo) para inserir um bloco de texto;
4. No bloco de texto, insira um texto que explique o que será executado, a seguir, no bloco de código a ser inserido. Segue uma sugestão, que pode ser complementada posteriormente por você:
5. Insira um novo bloco de texto com o conteúdo: “Passo 1: Instalando as bibliotecas e recarregando o ambiente”;
6. Insira um bloco de código com o conteúdo abaixo:
7. Execute o código acima. Durante o processo, caso receba, na tela, a mensagem dizendo que a sessão precisa ser reiniciada, clique no respectivo botão;

Análise de Sentimentos. Passo 1: Instalando as bibliotecas e recarregando o ambiente

```
1  !pip install -U pip setuptools wheel
2
3  !pip install -U spacy
4
5  !python -m spacy download en_core_web_sm
6
7  !pip install spacytextblob
8
9  import pkg_resources,imp
10
11  imp.reload(pkg_resources)
```

CO sentiment.ipynb

Arquivo Editar Ver Inserir Ambiente de execução Ferramentas Ajuda Todas as alterações foram salvas

+ Código + Texto

215

```
Requirement already satisfied: pip in /usr/local/lib/python3.10/dist-packages (24.3.1)
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (75.6.0)
Requirement already satisfied: wheel in /usr/local/lib/python3.10/dist-packages (0.45.1)
Requirement already satisfied: spacy in /usr/local/lib/python3.10/dist-packages (3.8.3)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.0.11)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.0.10)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.0.9)
Requirement already satisfied: thinc<8.4.0,>=8.3.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (8.3.2)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.4.8)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.1.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (0.4.1)
Requirement already satisfied: typer<1.0.0,>=0.3.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (0.15.0)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (4.66.6)
Requirement already satisfied: numpy<1.19.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.0.2)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.32.3)
Requirement already satisfied: pydantic<1.8.1,>=1.8.1,<3.0.0,>=1.7.4 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.10.3)
Requirement already satisfied: Jinja2 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.1.4)
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from spacy) (75.6.0)
Requirement already satisfied: packaging>20.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (24.2)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.5.0)
Requirement already satisfied: language-data==1.2 in /usr/local/lib/python3.10/dist-packages (from langcodes<4.0.0,>=3.2.0->spacy) (1.3.0)
Requirement already satisfied: annotated-types==0.6.0 in /usr/local/lib/python3.10/dist-packages (from pydantic!=1.8.1,>=1.8.1,<3.0.0,>=1.7.4->spacy) (0.7.0)
Requirement already satisfied: pydantic-core==2.27.1 in /usr/local/lib/python3.10/dist-packages (from pydantic!=1.8.1,>=1.8.1,<3.0.0,>=1.7.4->spacy) (2.27.1)
Requirement already satisfied: typing-extensions==4.12.2 in /usr/local/lib/python3.10/dist-packages (from pydantic!=1.8.1,>=1.8.1,<3.0.0,>=1.7.4->spacy) (4.12.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (3.4.0)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2.2.3)
Requirement already satisfied: certifi>2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2024.8.30)
Requirement already satisfied: blis<1.1.0,>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from thinc<8.4.0,>=8.3.0->spacy) (1.0.2)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.10/dist-packages (from thinc<8.4.0,>=8.3.0->spacy) (0.1.5)
Requirement already satisfied: click>8.0.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy) (8.1.7)
Requirement already satisfied: shellingham>1.3.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy) (1.5.4)
Requirement already satisfied: rich>10.11.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy) (13.9.4)
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from weasel<0.5.0,>=0.1.0->spacy) (0.20.0)
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/python3.10/dist-packages (from weasel<0.5.0,>=0.1.0->spacy) (7.0.5)
Requirement already satisfied: MarkupSafe>2.0 in /usr/local/lib/python3.10/dist-packages (from Jinja2->spacy) (3.0.2)
Requirement already satisfied: marisa-trie>1.1.0 in /usr/local/lib/python3.10/dist-packages (from language-data==1.2->langcodes<4.0.0,>=3.2.0->spacy) (1.2.1)
Requirement already satisfied: markdown-it-py>2.2.0 in /usr/local/lib/python3.10/dist-packages (from rich>10.11.0->typer<1.0.0,>=0.3.0->spacy) (3.0.0)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from rich>10.11.0->typer<1.0.0,>=0.3.0->spacy) (2.18.0)
Requirement already satisfied: wrapt in /usr/local/lib/python3.10/dist-packages (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.1.0->spacy) (1.17.0)
Requirement already satisfied: mdurl==0.1 in /usr/local/lib/python3.10/dist-packages (from markdown-it-py>2.2.0->rich>10.11.0->typer<1.0.0,>=0.3.0->spacy) (0.1.2)
Collecting en-core-web-sm==3.8.0
  Downloading https://github.com/explosion/spacy-models/releases/download/en_core_web_sm-3.8.0/en_core_web_sm-3.8.0-py3-none-any.whl (12.8 MB)
    12.8/12.8 MB 81.4 MB/s eta 0:00:00
✓ Download and installation successful
You can now load the package via spacy.load('en_core_web_sm')
```

sentiment.ipynb

Arquivo Editar Ver Inserir Ambiente de execução Ferramentas Ajuda Todas as alterações foram salvas

+ Código + Texto

215

```
Restart to reload dependencies
If you are in a Jupyter or Colab notebook, you may need to restart Python in order to load all the package's dependencies. You can do this by selecting the 'Restart kernel' or 'Restart runtime' option.
Requirement already satisfied: spacytextblob in /usr/local/lib/python3.10/dist-packages (5.0.0)
Requirement already satisfied: spacy<3.0.0 in /usr/local/lib/python3.10/dist-packages (from spacytextblob) (3.8.3)
Requirement already satisfied: textblob>=0.18.0.post0 in /usr/local/lib/python3.10/dist-packages (from spacytextblob) (0.18.0.post0)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (1.0.11)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (2.0.10)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (3.0.9)
Requirement already satisfied: thinc<8.4.0,>=8.3.0 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (8.3.2)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (2.4.8)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.1.0 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (0.4.1)
Requirement already satisfied: typer<1.0.0,>=0.3.0 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (0.15.0)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (4.66.6)
Requirement already satisfied: numpy<1.19.0 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (2.0.2)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (2.32.3)
Requirement already satisfied: pydantic<1.8.1,>=1.8.1,<3.0.0,>=1.7.4 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (2.10.3)
Requirement already satisfied: Jinja2 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (3.1.4)
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (75.6.0)
Requirement already satisfied: packaging>20.0 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (24.2)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (3.5.0)
Requirement already satisfied: nltk>3.8 in /usr/local/lib/python3.10/dist-packages (from textblob>=0.18.0.post0->spacytextblob) (3.9.1)
Requirement already satisfied: language-data==1.2 in /usr/local/lib/python3.10/dist-packages (from langcodes<4.0.0,>=3.2.0->spacy>=3.0.0->spacytextblob) (1.3.0)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk>=3.8->textblob>=0.18.0.post0->spacytextblob) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk>=3.8->textblob>=0.18.0.post0->spacytextblob) (1.4.2)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk>=3.8->textblob>=0.18.0.post0->spacytextblob) (2024.9.11)
Requirement already satisfied: annotated-types==0.6.0 in /usr/local/lib/python3.10/dist-packages (from pydantic!=1.8.1,>=1.8.1,<3.0.0,>=1.7.4->spacy>=3.0.0->spacytextblob) (0.7.0)
Requirement already satisfied: pydantic-core==2.27.1 in /usr/local/lib/python3.10/dist-packages (from pydantic!=1.8.1,>=1.8.1,<3.0.0,>=1.7.4->spacy>=3.0.0->spacytextblob) (2.27.1)
Requirement already satisfied: typing-extensions==4.12.2 in /usr/local/lib/python3.10/dist-packages (from pydantic!=1.8.1,>=1.8.1,<3.0.0,>=1.7.4->spacy>=3.0.0->spacytextblob) (4.12.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy>=3.0.0->spacytextblob) (3.4.0)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy>=3.0.0->spacytextblob) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy>=3.0.0->spacytextblob) (2.2.3)
Requirement already satisfied: certifi>2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy>=3.0.0->spacytextblob) (2024.8.30)
Requirement already satisfied: blis<1.1.0,>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from thinc<8.4.0,>=8.3.0->spacy>=3.0.0->spacytextblob) (1.0.2)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.10/dist-packages (from thinc<8.4.0,>=8.3.0->spacy>=3.0.0->spacytextblob) (0.1.5)
Requirement already satisfied: shellingham>1.3.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy>=3.0.0->spacytextblob) (1.5.4)
Requirement already satisfied: rich>10.11.0 in /usr/local/lib/python3.10/dist-packages (from typer<1.0.0,>=0.3.0->spacy>=3.0.0->spacytextblob) (13.9.4)
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from weasel<0.5.0,>=0.1.0->spacy>=3.0.0->spacytextblob) (0.20.0)
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/python3.10/dist-packages (from weasel<0.5.0,>=0.1.0->spacy>=3.0.0->spacytextblob) (7.0.5)
Requirement already satisfied: MarkupSafe>2.0 in /usr/local/lib/python3.10/dist-packages (from Jinja2->spacy>=3.0.0->spacytextblob) (3.0.2)
Requirement already satisfied: marisa-trie>1.1.0 in /usr/local/lib/python3.10/dist-packages (from language-data==1.2->langcodes<4.0.0,>=3.2.0->spacy>=3.0.0->spacytextblob) (1.2.1)
Requirement already satisfied: markdown-it-py>2.2.0 in /usr/local/lib/python3.10/dist-packages (from rich>10.11.0->typer<1.0.0,>=0.3.0->spacy>=3.0.0->spacytextblob) (3.0.0)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from rich>10.11.0->typer<1.0.0,>=0.3.0->spacy>=3.0.0->spacytextblob) (2.18.0)
Requirement already satisfied: wrapt in /usr/local/lib/python3.10/dist-packages (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.1.0->spacy>=3.0.0->spacytextblob) (1.17.0)
Requirement already satisfied: mdurl==0.1 in /usr/local/lib/python3.10/dist-packages (from markdown-it-py>2.2.0->rich>10.11.0->typer<1.0.0,>=0.3.0->spacy>=3.0.0->spacytextblob) (0.1.2)
<module 'pkg_resources' from '/usr/local/lib/python3.10/dist-packages/pkg_resources/_init_.py'>
```

Passo 2: Importando as bibliotecas para análise de sentimento

[8] 1 import spacy

Conectado ao dispositivo: de back-end do Google Compute Engine em Python 3

8. Insira um novo bloco de texto com o conteúdo: “Passo 2: Importando as bibliotecas para análise de sentimento”;
9. Insira um bloco de código com as linhas abaixo e execute:

Passo 2: Importando as bibliotecas para análise de sentimento

```
✓ [8] 1 import spacy
      2
      3 from spacytextblob.spacytextblob import SpacyTextBlob
```

10. Insira um novo bloco de texto com o conteúdo: “Passo 3: Definindo o modelo e a pipeline a serem utilizadas na análise”;
11. Crie um bloco de código com as linhas abaixo e execute:

Passo 3: Definindo o modelo e a pipeline a serem utilizadas na análise

```
[9] 1 nlp = spacy.load('en_core_web_sm')
     2
     3 nlp.add_pipe('spacytextblob')
```

↗ <spacytextblob.spacytextblob.SpacyTextBlob at 0x7f89f038bb20>

12. Insira um novo bloco de texto com o conteúdo: “Passo 4: Definindo o texto inicial a ser analisado para verificação/validação da biblioteca”;
13. Crie um bloco de código com o conteúdo abaixo e o execute:

Passo 4: Definindo o texto inicial a ser analisado para verificação/validação da biblioteca

```
[10] 1 user_input = 'This is a wonderful campsite. I loved the serenity and the birds chirping in the morning.'
      2
      3 doc = nlp(user_input)
```

14. Insira um novo bloco de texto com o conteúdo: “Passo 5: Exibindo o resultado da primeira análise (um range entre -1 [avaliação negativa] e 1 [avaliação positiva]”;
15. Crie um bloco de código com as linhas abaixo e o execute:

Passo 5: Exibindo o resultado da primeira análise (um range entre -1 [avaliação negativa] e 1 [avaliação positiva])

```
[11] 1 input_polarity = doc._.blob.polarity
      2 sentiment = {
      3
      4     'score': input_polarity
      5 }
      6
      7 print(sentiment)
```

↗ {'score': 0.85}

16. Insira um novo bloco de texto com o conteúdo: “Passo 6: Definindo a lista de tweets a serem analisadas”;

17. Insira um bloco de código com as linhas abaixo e execute:

Passo 6: Definindo a lista de tweets a serem analisadas

```
[16] 1 tweets=["Bayer Leverkusen goalkeeper Bernd Leno will not be going to Napoli. His agent Uli Ferber to Bild: I can confirm that there were negotiations with Napoli, w
2
3 "Gary Speed v Blackburn at St James in 2001/02 anyone? #NUFC #BEL #JAP #WorldCup",
4
5 "@ChelseaFC Don't make him regret it and start him over Hoofiz",
6
7 "@LiverpoolFF @AnfieldEdition He's a liar, made up. I've unfollowed him as loads of others have. Pure blagger. #LFC",
8
9 "@theesk @Everton Didn't realise Kenwright is due to leave at the end of the month. In all seriousness could you see him being interested in us?",
10
11 "@hasanshahbaz19 @LFC My knowledge has decreased somewhat in the past few seasons",
12
13 "Report: Linked with #Everton and #Wolves, Italians set to sign £4.5m-rated winger",
14
15 "Am seeing tweets that there's been a fall out @Everton between the money men... I'm presuming it's just a quiet news day or some kopite with nothing better to do!
16
17 "@LFC @officialAL20 @IntChampionsCup @ManUtd Expect loads of excuses after tonight's game",
18
19 "@MartinDiamond17 @azryahmad @Baren_D @MathewLewis1997 @iamheinthu @DiMarzio @Alissonbecker @LFC @SkySportsNews @SkySport @OfficialASRoma I'm just fine I have your
20
21 "What a weekend of football results! @ManUtd @Glentoran @RangersFC & Hearts ????",
22
23 "@ChelseaFC For the first time in a long while, my heart was relaxed while watching Chelsea. Really enjoyed it today. Come on, CHELSEA!!!",
24
25 "@ChelseaFC @CesarAzpi What a fantastic signing worth every single penny ??",
26
```

Passo 6: Definindo a lista de tweets a serem analisadas

```
tweets=["Bayer Leverkusen goalkeeper Bernd Leno will not be going to Napoli.
His agent Uli Ferber to Bild: I can confirm that there were negotiations with
Napoli, which we have broken off. Napoli is not an option. Atletico Madrid and
Arsenal are the other strong rumours. #B04 #AFC",
```

```
"Gary Speed v Blackburn at St James in 2001/02 anyone? #NUFC #BEL #JAP
#WorldCup",
```

```
"@ChelseaFC Don't make him regret it and start him over Hoofiz",
```

```
"@LiverpoolFF @AnfieldEdition He's a liar, made up. I've unfollowed him as
loads of others have. Pure blagger. #LFC",
```

```
"@theesk @Everton Didn't realise Kenwright is due to leave at the end of the
month. In all seriousness could you see him being interested in us?",
```

```
"@hasanshahbaz19 @LFC My knowledge has decreased somewhat in the past few
seasons",
```

```
"Report: Linked with #Everton and #Wolves, Italians set to sign £4.5m-rated
winger",
```

```
"Am seeing tweets that there's been a fall out @Everton between the money
men... I'm presuming it's just a quiet news day or some kopite with nothing
better to do! @ALANMYERSMEDIA",
```

```
"@LFC @officialAL20 @IntChampionsCup @ManUtd Expect loads of excuses after
tonight's game",
```



```
"@MartinDiamond17 @azryahmad @Baren_D @MathewLewis1997 @iamheinthu @DiMarzio
@Alissonbecker @LFC @SkySportsNews @SkySport @OfficialASRoma I'm just fine I
have your fanbase angry over stating facts should ask them hun. Xo",

"What a weekend of football results! @ManUtd @Glentoran @RangersFC & Hearts
????",

"@ChelseaFC For the first time in a long while, my heart was relaxed while
watching Chelsea. Really enjoyed it today. Come on, CHELSEA!!!",

"@ChelseaFC @CesarAzpi What a fantastic signing worth every single penny ??",

"Pogba scores, Pogba assists. But tomorrow papers won't be telling you this,
instead they will tell you how he'll end up at Juve because he's unhappy,
frustrated, have grudges with Mourinho and so on and so forth #mufc",

"@WestHamUtd we need to keep @CH14_ and get @HirvingLozano70 to compliment",

"@kevdev9 @Everton Shouldn't be happening! Needs to stay away with his venomous
attitude until he is sold!",

"@brfootball @aguerosergiokun @ManCity What a genius. Pep taking winning
mentality with him, conquering league after league. Baller",

"@HMZ0709 Can we get a RT for our #lfc Mo Salah Liverpool Enamel Pin Badge"]
```

18. Insira um novo bloco de texto com o conteúdo: “Passo 7: Analisando os tweets”;

19. Insira e execute o bloco de código abaixo:

20. Por fim, caso queira, você poderá salvar uma cópia do código no Google Drive ou no Github. Tais opções encontram-se disponíveis a partir do menu Arquivo.

Passo 7: Analisando os tweets

```
1  for item in tweets:
2
3      doc = nlp(item)
4
5      input_polarity = doc._.blob.polarity
6
7      sentiment = {
8
9          'tweet': item,
10
11         'score': input_polarity
12     }
13
14
15     print(sentiment)
```

Saída do passo 7 nas próximas páginas

{'tweet': 'Bayer Leverkusen goalkeeper Bernd Leno will not be going to Napoli. His agent Uli Ferber to Bild: I can confirm that there were negotiations with Napoli, which we have broken off. Napoli is not an option. Atletico Madrid and Arsenal are the other strong rumours. #B04 #AFC', 'score': -0.030555555555555575}

{'tweet': 'Gary Speed v Blackburn at St James in 2001/02 anyone? #NUFC #BEL #JAP #WorldCup', 'score': 0.0}

{'tweet': "@ChelseaFC Don't make him regret it and start him over Hoofiz", 'score': 0.0}

{'tweet': "@LiverpoolFF @AnfieldEdition He's a liar, made up. I've unfollowed him as loads of others have. Pure blagger. #LFC", 'score': 0.21428571428571427}

{'tweet': "@theesk @Everton Didn't realise Kenwright is due to leave at the end of the month. In all seriousness could you see him being interested in us?", 'score': 0.0625}

{'tweet': '@hasanshahbaz19 @LFC My knowledge has decreased somewhat in the past few seasons', 'score': -0.28333333333333334}

{'tweet': 'Report: Linked with #Everton and #Wolves, Italians set to sign £4.5m-rated winger', 'score': 0.0}

{'tweet': 'Am seeing tweets that there's been a fall out @Everton between the money men... I'm presuming it's just a quiet news day or some kopite with nothing better to do! @ALANMYERSMEDIA', 'score': 0.3125}

{'tweet': '@LFC @officialAL20 @IntChampionsCup @ManUtd Expect loads of excuses after tonight's game', 'score': -0.4}

{'tweet': '@MartinDiamond17 @azryahmad @Baren_D @Mathewlewis1997 @iamheinthu @DiMarzio @Alissonbecker @LFC @SkySportsNews @SkySport @OfficialASRoma I'm just fine I have your fanbase angry over stating facts should ask them hun. Xo', 'score': -0.041666666666666666}

{'tweet': 'What a weekend of football results! @ManUtd @Glentoran @RangersFC & Hearts ?????', 'score': 0.0}

{'tweet': '@ChelseaFC For the first time in a long while, my heart was relaxed while watching Chelsea. Really enjoyed it today. Come on, CHELSEA!!!', 'score': 0.39218749999999997}

{'tweet': '@ChelseaFC @CesarAzpi What a fantastic signing worth every single penny ??', 'score': 0.20952380952380953}

{'tweet': "Pogba scores, Pogba assists. But tomorrow papers won't be telling you this, instead they will tell you how he'll end up at Juve because he's unhappy, frustrated, have grudges with Mourinho and so on and so forth #mufc", 'score': -0.6499999999999999}

{'tweet': '@WestHamUtd we need to keep @CH14_ and get @HirvingLozano70 to compliment', 'score': 0.0}

{'tweet': '@kevdev9 @Everton Shouldn't be happening! Needs to stay away with his venomous attitude until he is sold!', 'score': 0.0}

{'tweet': '@brfootball @aguerosergiokun @ManCity What a genius. Pep taking winning mentality with him, conquering league after league. Baller', 'score': 0.5}

{'tweet': '@HMZ0709 Can we get a RT for our #lfc Mo Salah Liverpool Enamel Pin Badge', 'score': 0.0}