



Missão Prática

Nível - 4

Tema: Dando Inteligência ao Software

Aluno: Yvo Murilo Santos D'Albuquerque

Objetivos da prática

- Configurar o ambiente Google Colab;
- Descrever tarefas diversas relacionadas ao Processamento de Linguagem Natural;
- Descrever o processo de identificação de entidades a partir de textos;
- Descrever o processo de extração de frases-chave a partir de textos;
- Descrever o processo de identificação de linguagem predominante a partir de textos;

- Contextualização

As microatividades, a seguir, envolverão uma série de tarefas de Inteligência Artificial

voltadas para o Processamento de Linguagem Natural utilizando bibliotecas da

linguagem Python. Atualmente existem ferramentas em nuvem que suportam esse tipo

de processo, como a AWS Comprehend e a Google Cloud Natural Language API, por exemplo. Entretanto, utilizaremos, ao longo das atividades, um misto de ferramentas Cloud e bibliotecas Python, formando assim um ambiente que poderia ser replicado em sua própria estação de trabalho. Nesse contexto, serão utilizados o Google Colab e a biblioteca Spacy (disponível para a linguagem de programação Python). Caso ainda não conheça tais ferramentas, vale a pena pesquisar um pouco a respeito das mesmas antes de iniciar a resolução dos exercícios.

Procedimentos -1

1. Acesse a página do Google Colab:
<https://colab.google>
2. No menu superior direito, seleciona 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.

Procedimentos - 2

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:
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;
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:
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
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:
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:

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:

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:

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:

Resultados esperados

```
!pip install -U pip setuptools wheel
!pip install -U spacy
!python -m spacy download en_core_web_sm
!pip install -U spacytextblob
import pkg_resources
import imp

imp.reload(pkg_resources)
```

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) (1.16.0)
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 52.4 MB/s eta 0:00:00

✓ Download and installation successful
You can now load the package via spacy.load('en_core_web_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.
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) (3.8.2)
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) (1.0.10)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (2.0.8)
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)

```
[ ] !pip install --upgrade numpy scipy nltk textblob spacytextblob
```

Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (2.0.2)
Collecting numpy
 Using cached numpy-2.1.3-cp310-cp310-manylinux_2_17_x86_64_manylinux2014_x86_64.whl.metadata (62 kB)
Requirement already satisfied: scipy in /usr/local/lib/python3.10/dist-packages (1.14.1)
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (3.9.1)
Requirement already satisfied: textblob in /usr/local/lib/python3.10/dist-packages (0.18.0.post0)
Requirement already satisfied: spacytextblob in /usr/local/lib/python3.10/dist-packages (5.0.0)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk) (1.4.2)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk) (2024.9.11)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk) (4.66.6)
Requirement already satisfied: spacy>=3.0.0 in /usr/local/lib/python3.10/dist-packages (from spacytextblob) (3.8.2)
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) (1.0.10)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy>=3.0.0->spacytextblob) (2.0.8)
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)


```
[ ] import pkg_resources, imp
    imp.reload(pkg_resources)
```

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

```
[ ] import spacy
    from spacytextblob.spacytextblob import SpacyTextBlob
```

```
nlp = spacy.load('en_core_web_sm')
```

```
nlp.add_pipe('spacytextblob')
```

```
tweets = [
```

```
    "Bayer Leverkusen goalkeeper Bernd Leno will not be going to Napoli. His agent Uli Ferber to Bild: I can confirm that",
    "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",
    "@hasanshabbaz19 @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 new",
    "@LFC @officialAL20 @IntChampionsCup @ManUtd Expect loads of excuses after tonight's game",
    "@MartinDiamond17 @azryahmad @Baren_D @MathewLewis1997 @iamheinthu @DiMarzio @Alissonbecker @LFC @SkySportsNews @Sky",
    "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",
    "@ChelseaFC @CesarAzpi What a fantastic signing worth every single penny ??",
```

```

for tweet in tweets:
    doc = nlp(tweet)
    input_polarity = doc._.blob.polarity

    truncated_tweet = tweet if len(tweet) <= 46 else tweet[:46] + '...'

    sentiment = {
        'tweet': truncated_tweet,
        'score': input_polarity
    }

    print(sentiment)

```

```

{
  "tweet": "Bayer Leverkusen goalkeeper Bernd Leno will no...",
  "score": -0.0305555555555555575
},
{
  "tweet": "Gary Speed v Blackburn at St James in 2001/02 ...",
  "score": 0.0
},
{
  "tweet": "@ChelseaFC Don't make him regret it and start ...",
  "score": 0.0
},
{
  "tweet": "@LiverpoolFF @AnfieldEdition He's a liar, made...",
  "score": 0.21428571428571427
},
{
  "tweet": "@theesk @Everton Didn't realise Kenwright is d...",
  "score": 0.0625
},
{
  "tweet": "@hasanshahbaz19 @LFC My knowledge has decrease...",
  "score": -0.28333333333333334
},
{
  "tweet": "Report: Linked with #Everton and #Wolves, Ital...",
  "score": 0.0
},
{
  "tweet": "Am seeing tweets that there's been a fall out ...",
  "score": 0.3125
},
{
  "tweet": "@LFC @officialAL20 @IntChampionsCup @ManUtd Ex...",
  "score": -0.4
},
{
  "tweet": "@MartinDiamond17 @azryahmad @Baren_D @Mathewle...",
  "score": -0.041666666666666666
},
{
  "tweet": "What a weekend of football results! @ManUtd @G...",
  "score": 0.0
},
{
  "tweet": "@ChelseaFC For the first time in a long while,...",
  "score": 0.39218749999999997
},
{
  "tweet": "@ChelseaFC @CesarAzpi What a fantastic signing...",
  "score": 0.20952380952380953
},
{
  "tweet": "Pogba scores, Pogba assists. But tomorrow pape...",
  "score": -0.6499999999999999
},
{
  "tweet": "@WestHamUtd we need to keep @CH14_ and get @Hi...",
  "score": 0.0
},
{
  "tweet": "@kevdev9 @Everton Shouldn't be happening! Need...",
  "score": 0.0
}

```