



## Missão Prática – Mundo 5 – Nível 4

Eduardo Alencar Garcia Ferreira - Matrícula: **202208341081**

Polo Silva Lobo – BH

### **Nível 4: Dando Inteligência ao Software**

#### **Objetivos da prática:**

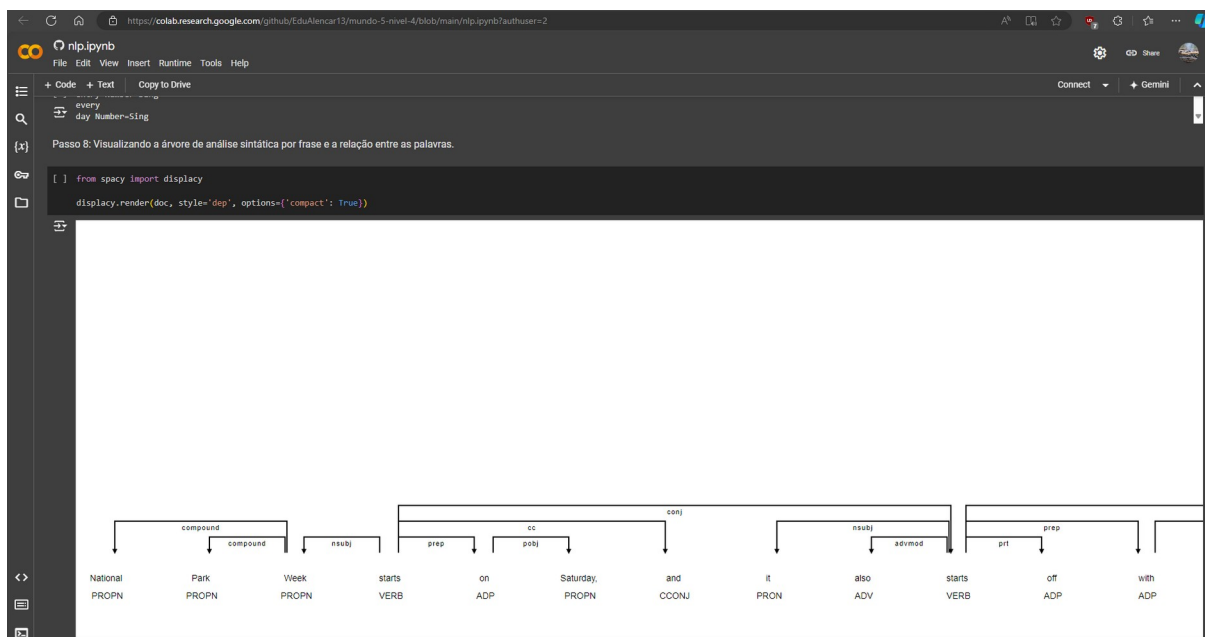
- Configurar o ambiente Google Colab;
- Descrever tarefas diversas relacionadas ao Processamento de Linguagem Natural;
- Descrever o processo de identificação de entidades a partir de textos;
- Descrever o processo de extração de frases-chave a partir de textos;
- Descrever o processo de identificação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

#### **Códigos do roteiro:**

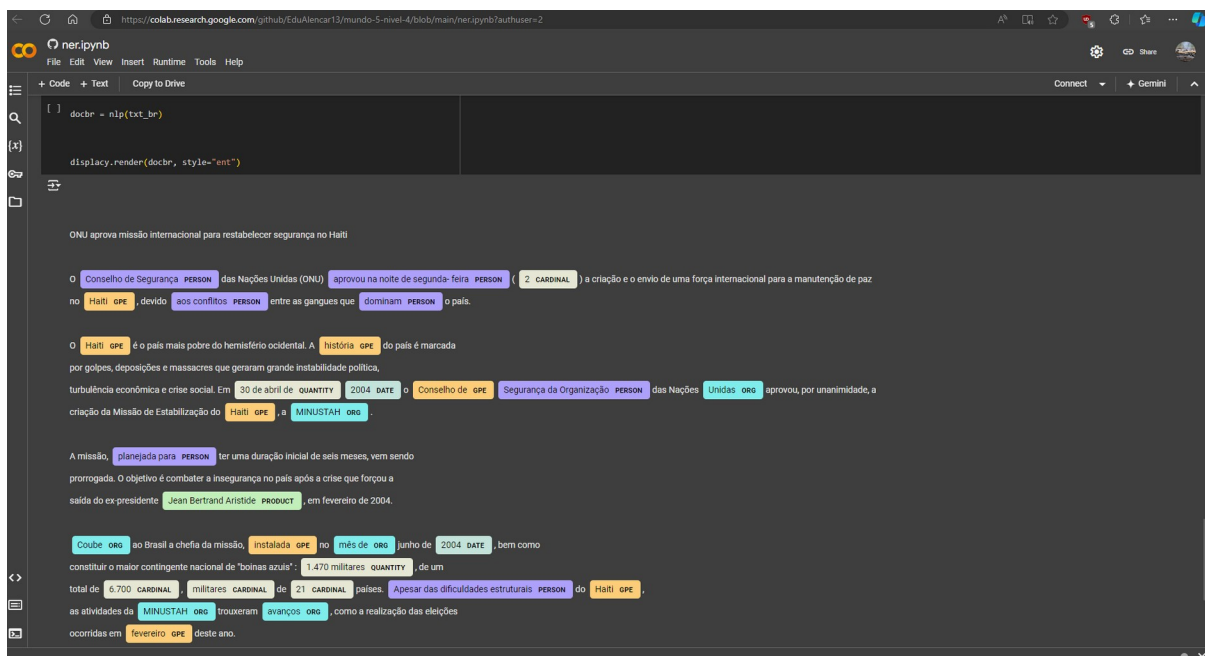
<https://github.com/EduAlencar13/mundo-5-nivel-4>

#### **Resultados da execução dos códigos:**

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



## Microatividade 3: Descrever o processo de identificação de entidades a partir de textos



## Microatividade 4: Descrever o processo de extração de frases-chave a partir de textos

```
kpeipynb
File Edit View Insert Runtime Tools Help

+ Code + Text Copy to Drive
[ ] - sentence_ids: [2]
    - pos_patterns: [['ADP', '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]

Passo 6: Ranqueando as palavras-chave candidatas

[ ] extractor.candidate_weighting()

for i, topic in enumerate(extractor.topics):
    # print out the topic id and the candidates it groups together
    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: accelar 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: recent
topic 17: slow
topic 18: soap
```

## Microatividade 5: Descrever o processo de identificação de linguagem predominante a partir de textos

```
lang.ipynb
File Edit View Insert Runtime Tools Help

+ Code + Text Copy to Drive
[ ] import pkg_resources,imp
    imp.reload(pkg_resources)

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

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

[ ] from langdetect import detect

for x in text:
    print ('Frase: ', x)
    print ('Idioma: ', detect(x), '\n\n')

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

Frase: Viseljén maszkot, tartsd 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.
```

## Missão prática: Dando Inteligência ao Software

```
https://colab.research.google.com/github/EduMencar13/mundo-5-nivel-4/blob/main/sentiment.ipynb?authuser=2

sentiment.ipynb
File Edit View Insert Runtime Tools Help

+ Code + Text Copy to Drive

[ ] input_polarity = doc._.polarity

sentiment = {
  'score': input_polarity
}

print(sentiment)

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
"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 Hoolitz",
"@liverpoolFF @Anfieldidiot 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?",
"@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 man... I'm presuming it's just a quiet news day on some kopite with nothing better to do! @ALANMYERSMEDIA",
"@LFC @officialAL20 @intchampionscup @ManUtd Expect loads of excuses after tonight's game",
"@MartirDiamond17 @azizahmad @Aren_D @MathewLewis1997 @iamheinth @DHarzlo @llissonbecker @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 @liverpool @RangersFC &mp, 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",
"@esttiamd we need to keep gchix and get @drvingiozamo to compliment",
"@kxdev9 @Everton Shouldn't be happening! Needs to stay away with his venomous attitude until he is sold!",
"@brfootball @agueroerikoun @ManCity What a genius. Pep taking winning mentality with him, conquering league after league. Baller",
"@KZ0709 Can we get a RT for our #lfc Mo Salah Liverpool Enamel Pin Badge"]

Passo 7: Analisando os tweets

[ ] for item in tweets:
    doc = nlp(item)
    input_polarity = doc._.polarity
    sentiment = {
        'tweet': item,
        'score': input_polarity
    }
    print(sentiment)
```