

ANO
2022



**CADERNO DE RESPOSTAS DA
ATIVIDADE PRÁTICA DE:**

NLP

**ALUNO: (ILMAR JORGE TELES ANDRADE
JUNIOR – RU: 3330053)**

Caderno de Resposta Elaborado por:
Prof. MSc. Guilherme Ditzel Patriota

Prática 01 – Criação de modelo de classificação supervisionado para análise de fake News.

Questão 01 – Apresente aqui o código referente ao modelo gerado e a nuvem de palavras que foram usadas para identificar textos VERDADEIROS

ENUNCIADO: Veja o Roteiro da Atividade Prática para mais detalhes.

- | | |
|--|---|
| I. Apresentação do Código (não esquecer do identificador pessoal): | II. Apresentação das Imagens/Print do resultado (não esquecer |
|--|---|

RU: 3330053 - ILMAR

Palavras Positivas

```
In [13]: wordcloud = WordCloud(stopwords=stopwords,
                               background_color='white', width=1600,
                               height=800, max_words=2000, mask=m,
                               min_font_size=1).generate(Palavras)

fig, ax = plt.subplots(figsize=(16,8))
plt.title("RU: 3330053 - ILMAR JORGE", fontsize=30, color='red')
ax.imshow(wordcloud, interpolation='bilinear')
ax.set_axis_off()
plt.imshow(wordcloud);

wordcloud.to_file("positivo.png")
```

Figura 1: (Máximo de 2000 palavras positivas para a análise do exercício)



do identificador):

III. Responda à pergunta: Quantas palavras, bigramas e trigramas foram usados dos textos rotulados como REAL para a criação de seu modelo e qual a acurácia?

Resposta: (A palavra que mais apareceu foi LAVA JATO e a quantidade total de palavras usadas foram 17421313 - acurácia de 93,61%)

Prática 01 – Criação de modelo de classificação supervisionado para análise de fake News.

Questão 02 – Apresente aqui o código referente ao modelo gerado e a nuvem de palavras que foram usadas para identificar textos FALSOS.

ENUNCIADO: Veja o Roteiro da Atividade Prática para mais detalhes.

- ## **I. Apresentação do Código (não esquecer do identificador pessoal):**

Palavras Negativas

```
[14]: wordcloud = WordCloud(stopwords=stopwords,
                           background_color='white', width=1800,
                           height=800, max_words=2000, mask=mask2, max_
                           min_font_size=1).generate(PalavrasNegativo)

fig, ax = plt.subplots(figsize=(16,8))
plt.title("RU: 3330053 - ILMAR JORGE", fontsize=30, color="red")
ax.imshow(wordcloud, interpolation='bilinear')
ax.set_axis_off()
plt.imshow(wordcloud);

wordcloud.to_file("negativo.png")
```

Figura 3: (INSERIR LEGENDA)

**resultado
esquecer
identificador):**

RU: 3330053 - ILMAR JOF



Figura 4: (INSERIR LEGENDA)

- III. Responda à pergunta: Quantas palavras, bigramas e trigramas foram usados dos textos rotulados como FAKE para a criação de seu modelo, quais foram as técnicas usadas em seu pré-processamento e qual tipo de modelo foi escolhido para este classificador?

Resposta: (A palavra que mais apareceu foi LULA e a quantidade total de palavras usadas foram 2942662 - acurácia de 93,61%)

Link do trabalho:

https://github.com/lilmar-ir/analise_sentimentos_ILMAR.git