

nexus

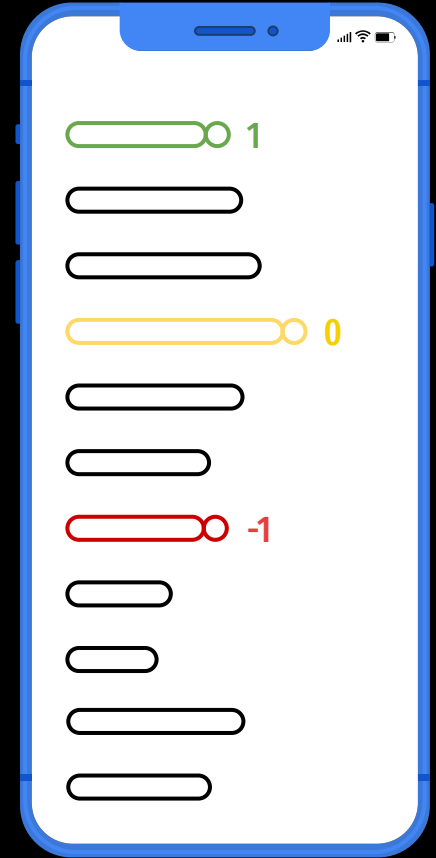
Conexão

Entre a Uber
e seus usuários



Desafio

Criar um modelo para
análise de sentimentos

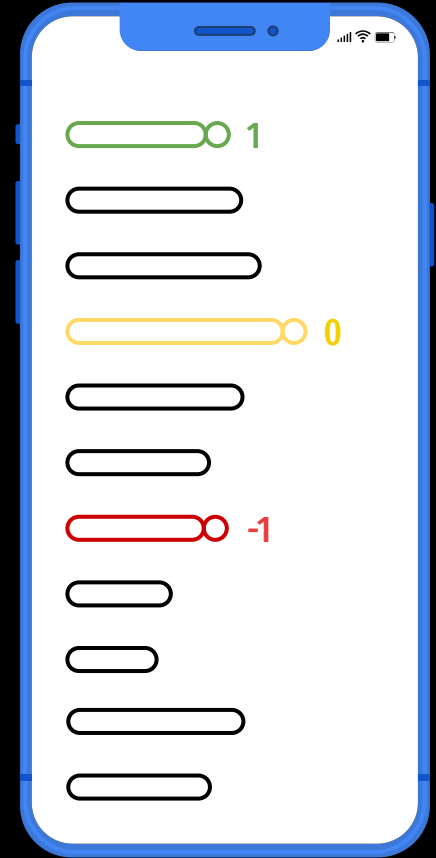


Desafio

Criar um modelo para
análise de sentimentos

Objetivo da Sprint

Criação de um modelo BoW



Personas



BENÍCIO SILVA

Analista de mídias sociais

Explicar a persona



MAYA SMITH

Gerente de mídias sociais

Explicar a persona



Personas



GIOVANNI SILVA

Cientista de dados

Explicar a persona

XX DRIVER LICENSE XX



User Stories



Etapas do pré-processamento

1. Segmentação

- Divisão do texto em frases menores

2. Tokenização

- Divisão das frases em tokens (palavras)

3. Stop-Words

- Remoção de palavras irrelevantes

4. Lematização

- Redução de palavras à sua forma base

5. NER

- Classificação de palavras

6. BoW

- Representação simplificada de um texto em uma matriz



1. Segmentação

All the time I spent learning to read French is worth it when I can read the 'details' about #EmanuelMacron and his Uber deal! Macron is a #SmoothOperator !

All the time I spent learning to read French is worth it when I can read the 'details' about #EmanuelMacron and his Uber deal!

Macron is a #SmoothOperator !



Etapas do pré-processamento

1. Segmentação

Too many taxi drivers have told me about Uber cutting access to the app and withholding earned wages - for months at a time - due to spurious customer complaints. Driver evidence made no difference. A big data leak shows more reasons not to use them. <https://t.co/DjcGHX9fYx>

i had a bad drive . i want my refund

Westin happen?? This could never be Abuja. My Uber trips in a month no reach like dis.



2. Tokenização

All the time I spent learning to read French is worth it when I can read the 'details' about #EmanuelMacron and his Uber deal!

All the time I spent learning
to read French is worth it
when I can read the 'details' about
#EmanuelMacron and his Uber deal !

Macron is a #SmoothOperator !

Macron is a
#SmoothOperator !



3. Stop Words

All the time I spent learning
to read French is worth it
when I can read the 'details' about
#EmanuelMacron and his Uber deal !

Macron is a
#SmoothOperator !

and: 1429

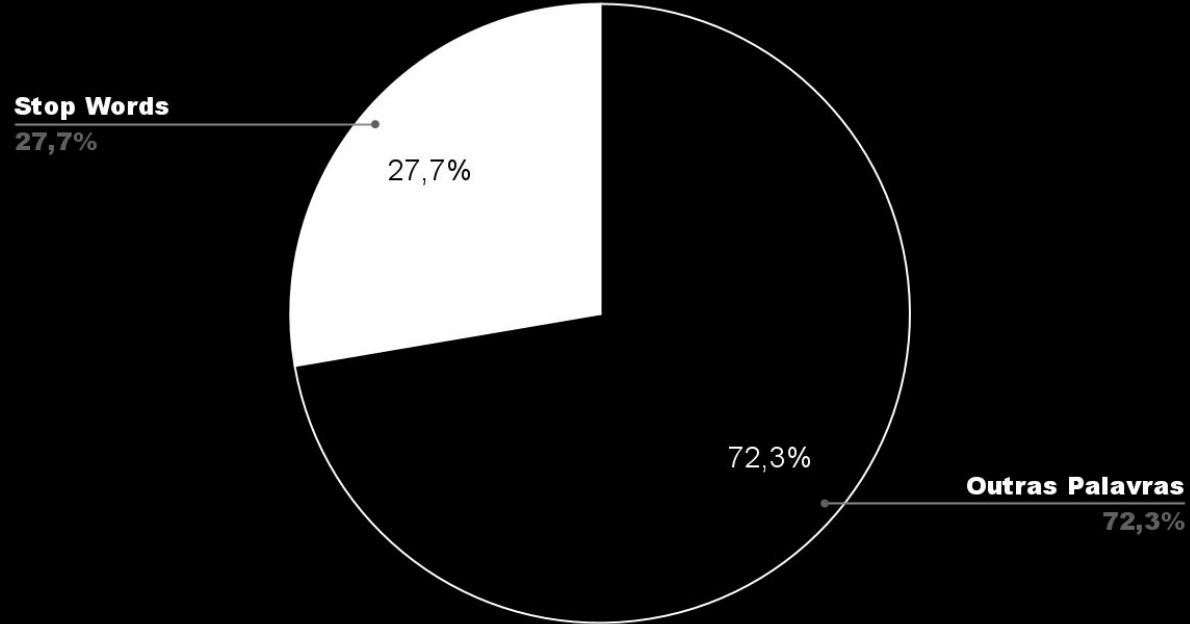
the: 1507

to: 1514



3. Stop Words

Distribuição de palavras



3. Stop Words

All the time I spent learning
to read French is worth it
when I can read the 'details' about
#EmanuelMacron and his Uber deal !

Macron is a
#SmoothOperator !

and: 1429

the: 1507

to: 1514



3. Stop Words

~~All~~ ~~the~~ time ~~it~~ spent learning
~~to~~ read French ~~is~~ worth ~~it~~
~~when~~ ~~it~~ ~~can~~ read ~~the~~ 'details' about
~~#EmanuelMacron~~ and ~~his~~ Uber deal ~~:~~

Macron ~~is~~ ~~a~~
~~#SmoothOperator~~ ~~:~~

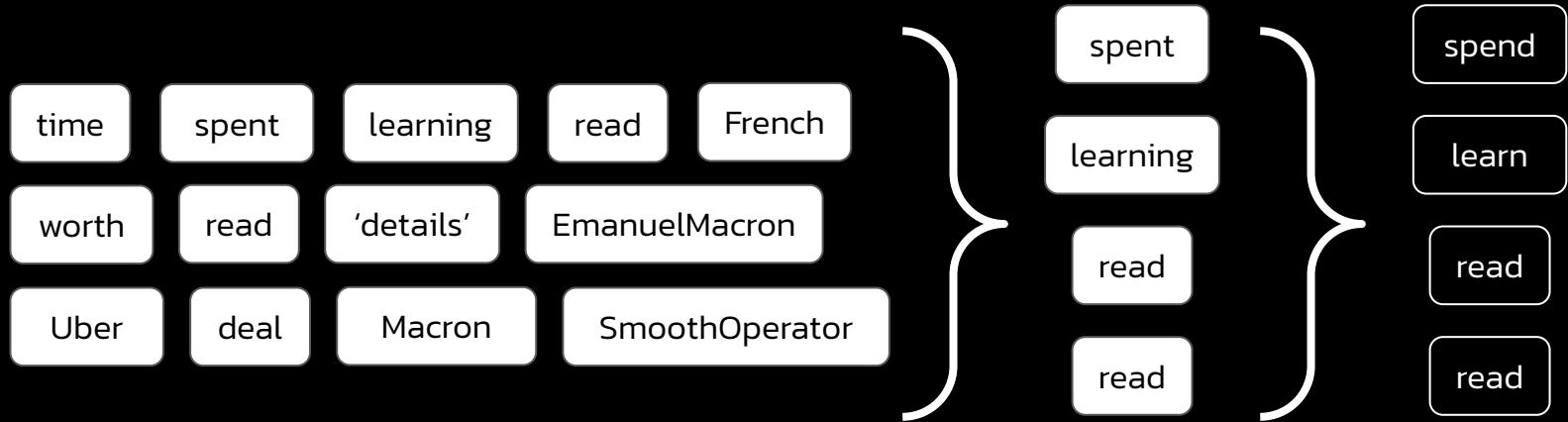
and: 1429

the: 1507

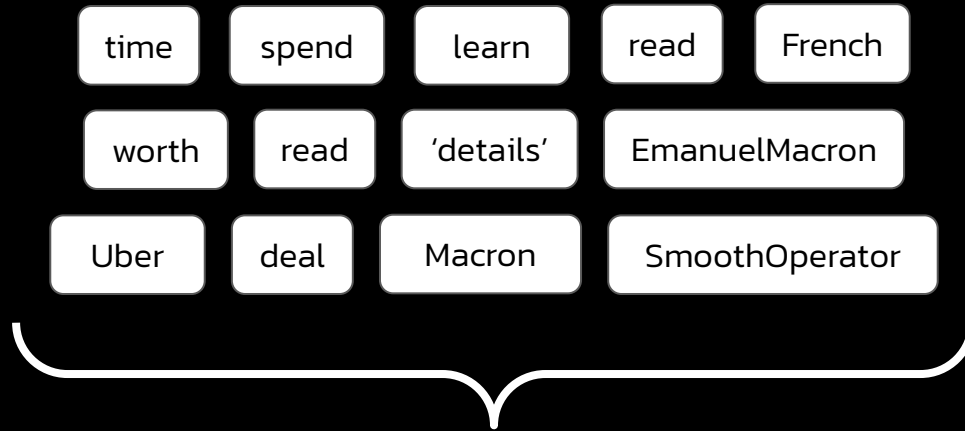
to: 1514



4. Lematização



5. NER



```
('French', 'LANGUAGE')  
( 'EmanuelMacron', 'PERSON')  
( 'Macron', 'PERSON')
```

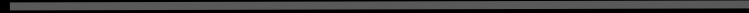


6. Aplicar BoW

learn	time	spend	Macron	detail	read	French	worth
1	1	1	1	1	1	1	1
0	0	0	1	0	0	0	0



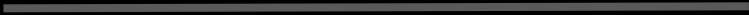
Primeiro modelo (Naive Bayes)



Triângulo



Círculo



Quadrado



Primeiro modelo (Naive Bayes)



Triângulo



Círculo



Quadrado

Acurácia: 0.854942871336314

```
[[1264 123 5]
 [ 88 394 0]
 [ 31 45 63]]
```



Primeiro modelo (Naive Bayes)



Triângulo



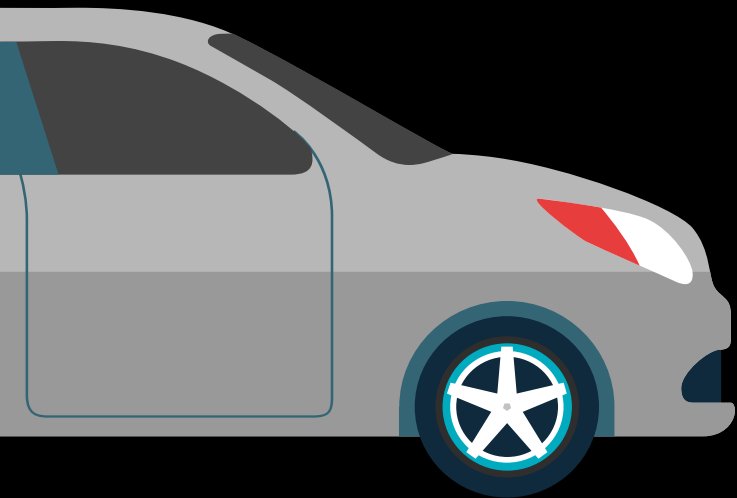
Círculo



Quadrado



Canvas Proposta de Valor



Tarefas do usuário:

Ler relatórios mensais para entender os sentimentos de seus clientes no último mês e gerar ações a partir disto

Dores:

Falta de escalabilidade nas análises atuais e existência de relatórios mensais

Ganhos:

Análises com resultados mais precisos e presença de alertas personalizados



Próximos passos

-

Word2Vec





Ana Martire



Daniel Zular



Davi Motta



Enzo Boccia



Erik Batista

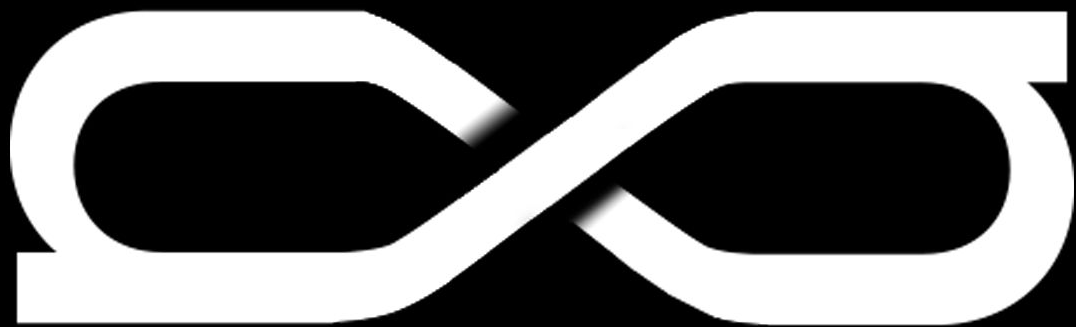


Lucas Barbosa



Paulo Octavio





nexus