





Boot

Índice

1

Última Sprint

2

Exploração

3

Pré-processamento

4

Bag of Words

5

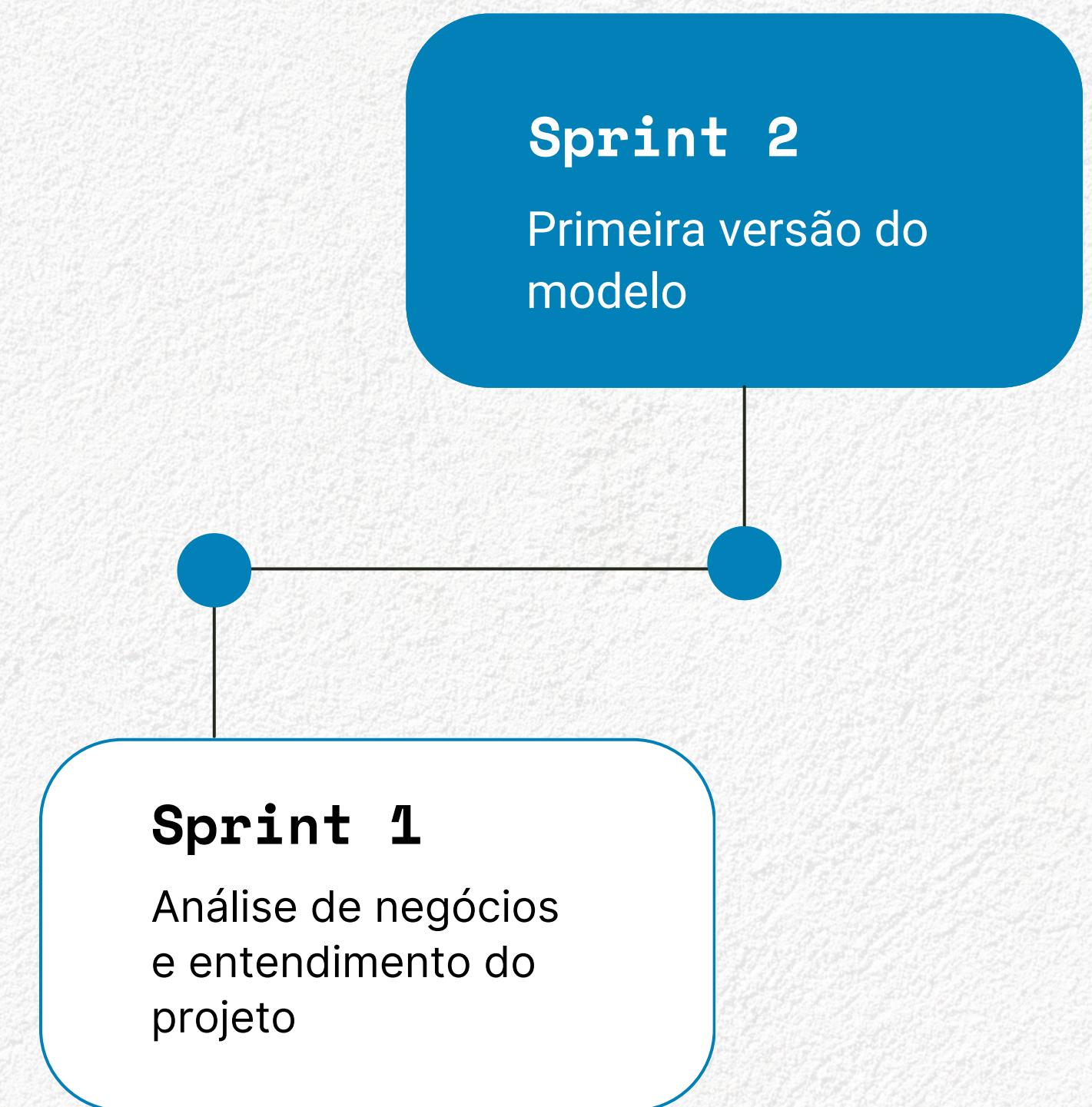
Modelo Inicial
PLN

6

Próximo
Incremento

Boot

1



Boot

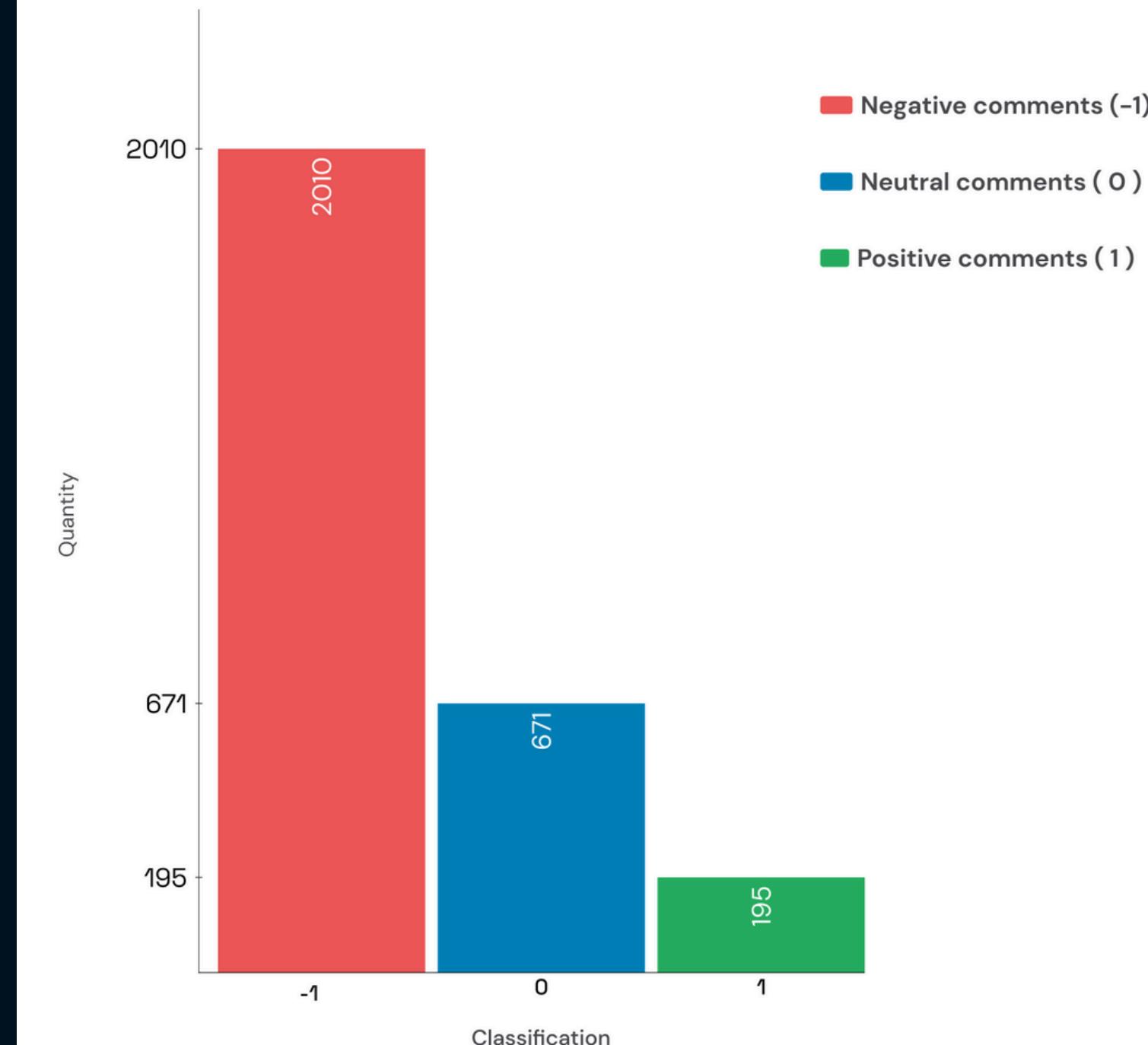
2

Exploração dos dados

Boot

2

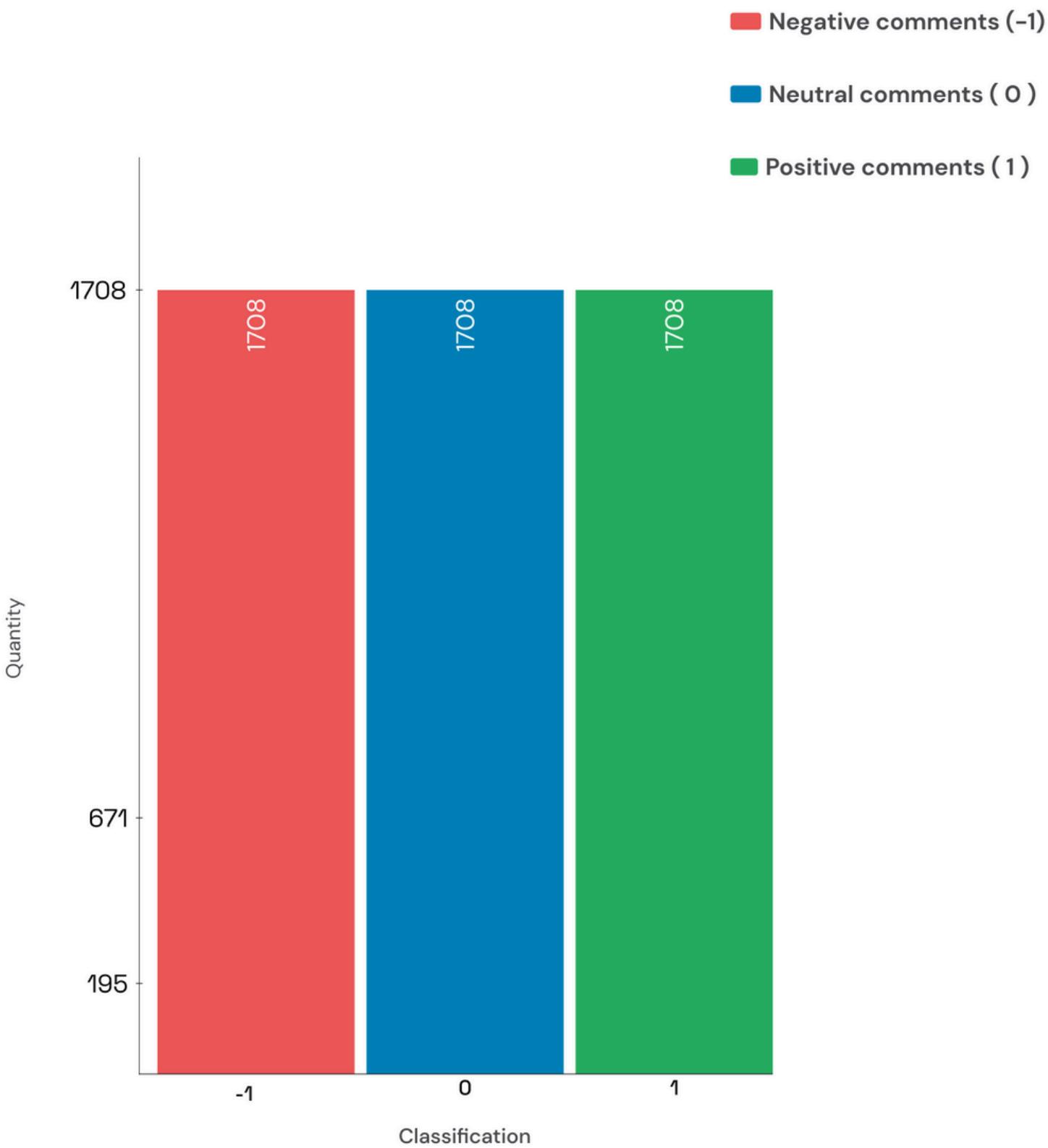
Number of posts classified by negative, positive and neutral sentiment



Boot

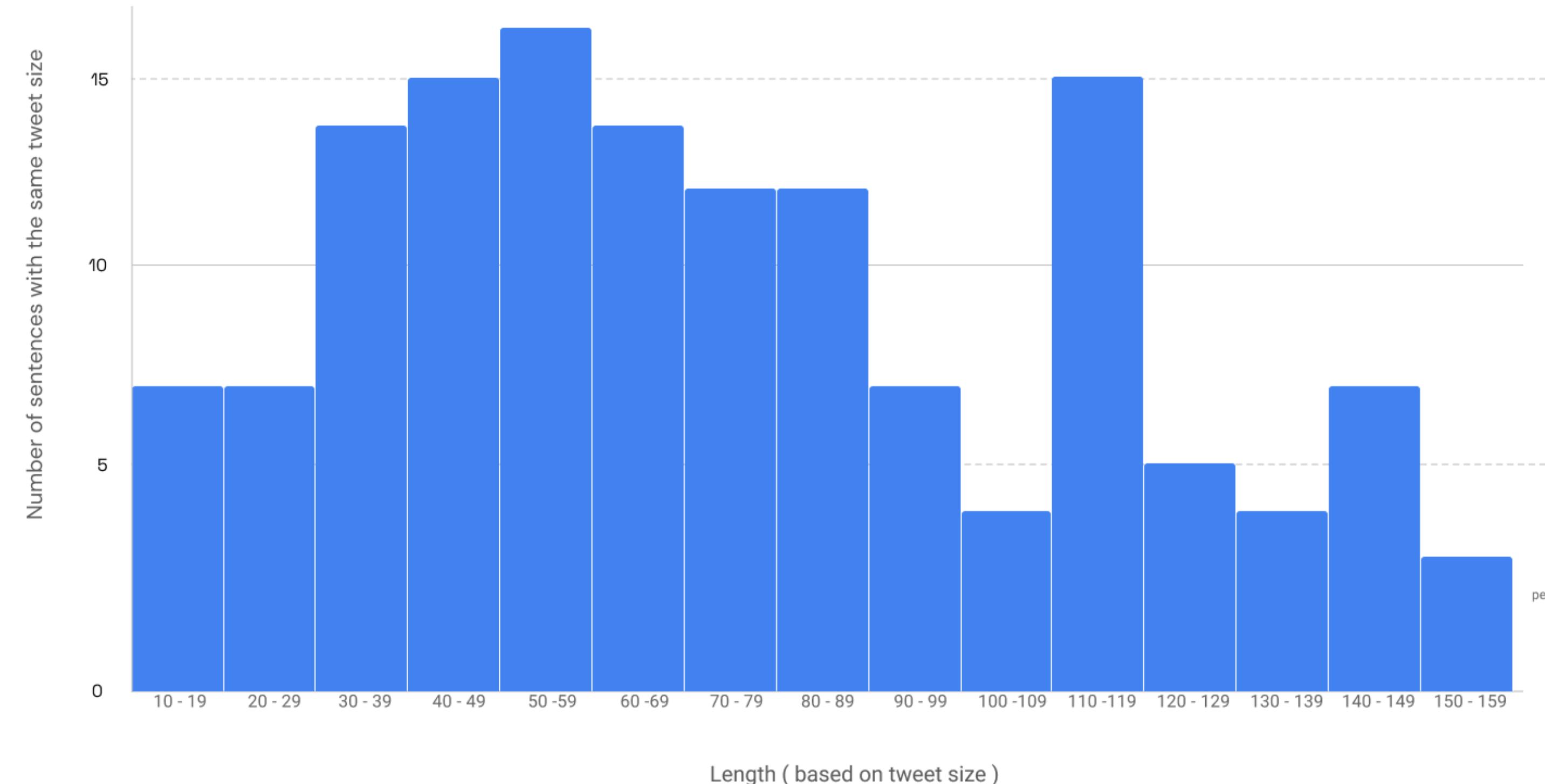
2

Class distribution after SMOTE Oversampling

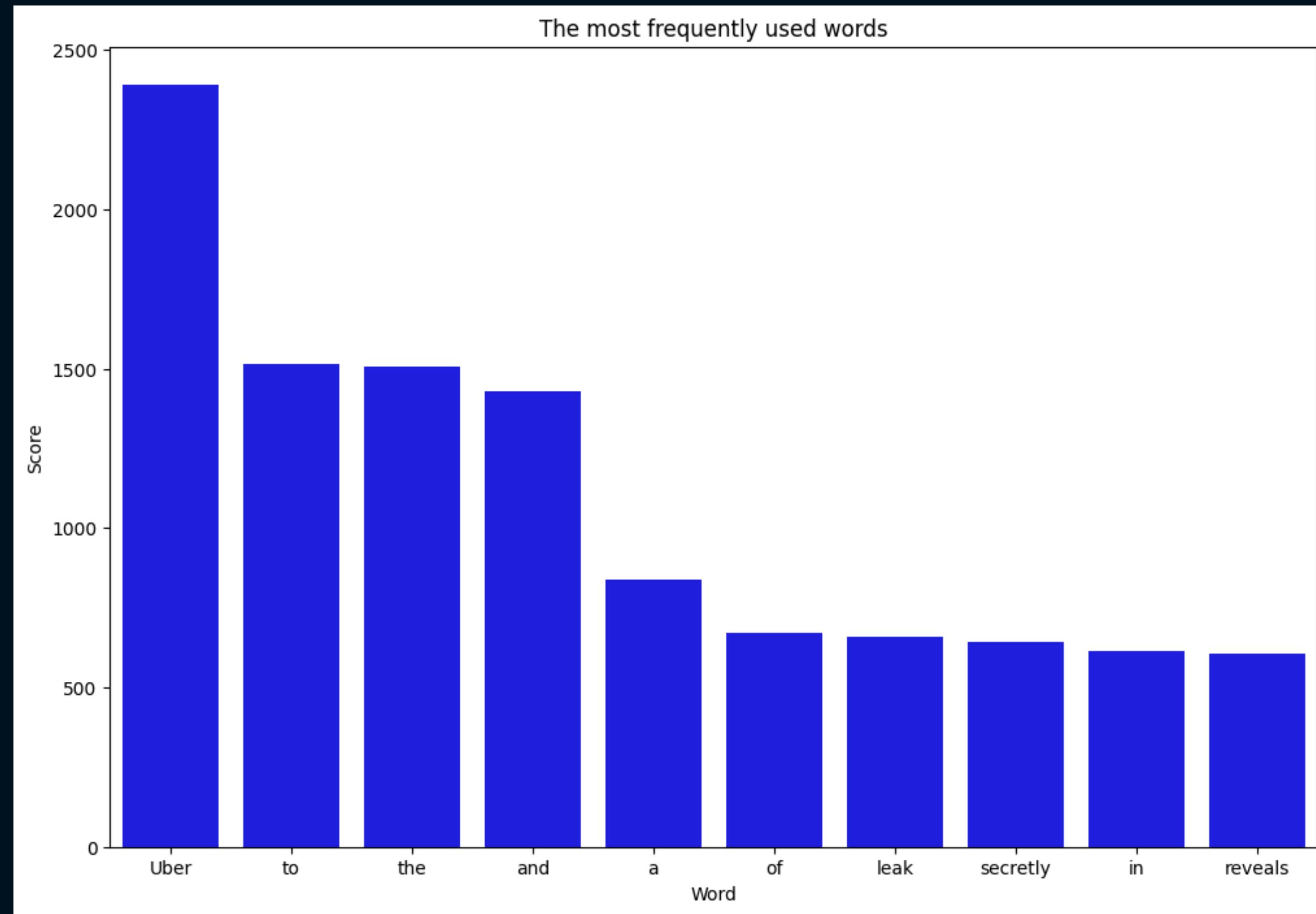


2

Size of sentences with positive classification



2



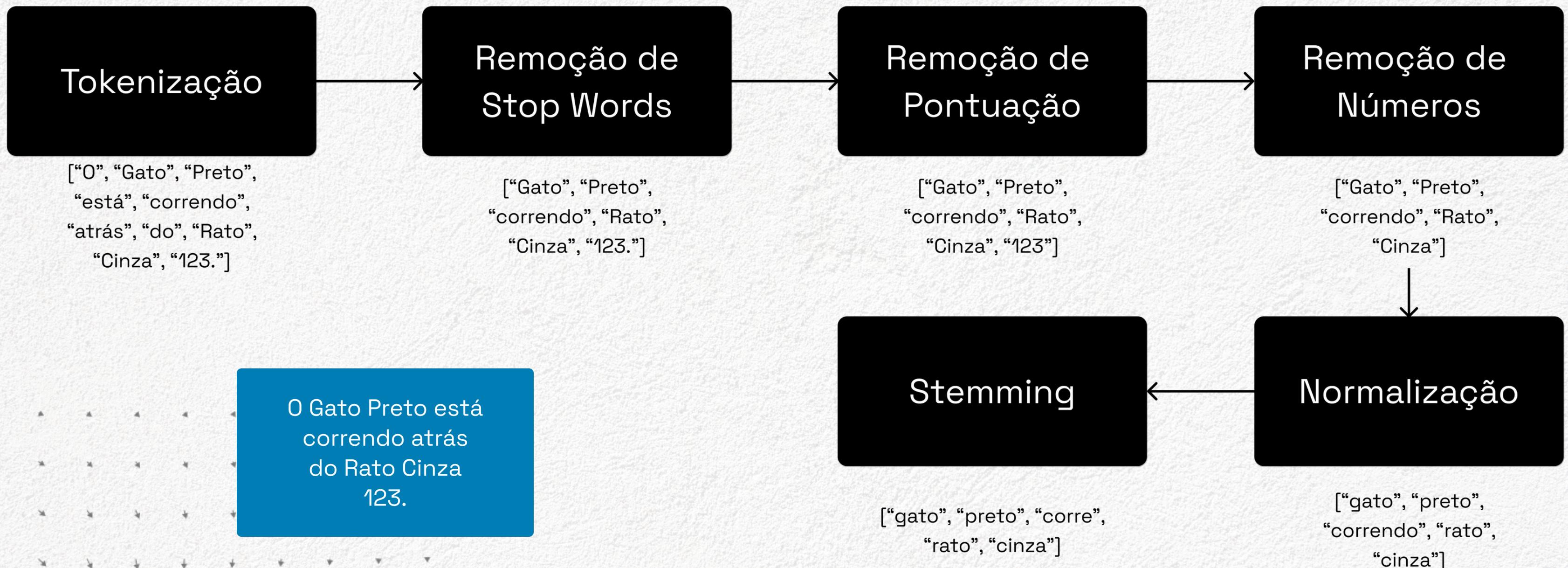
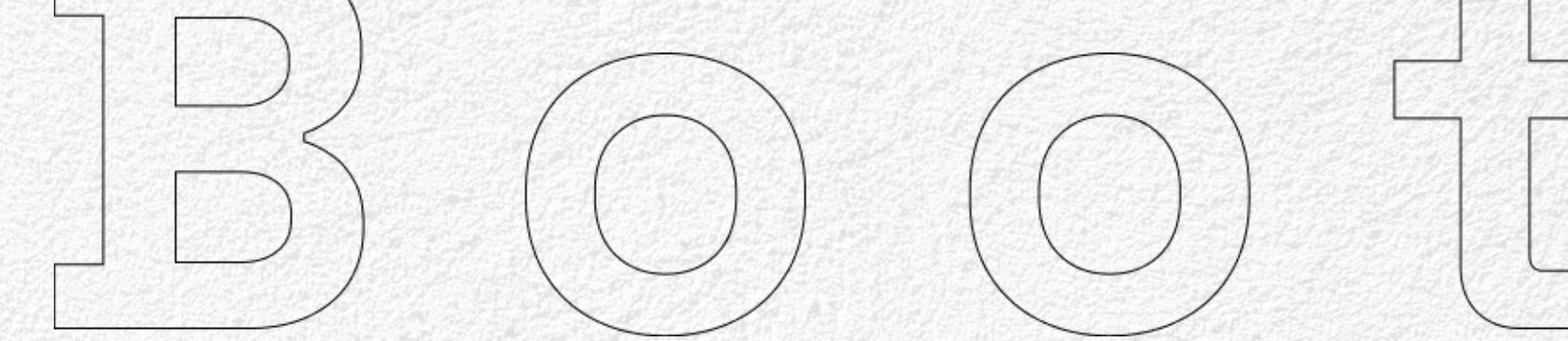
Boot

3

Pré-processamento

Boot

3



4

Bag of Words

Boot

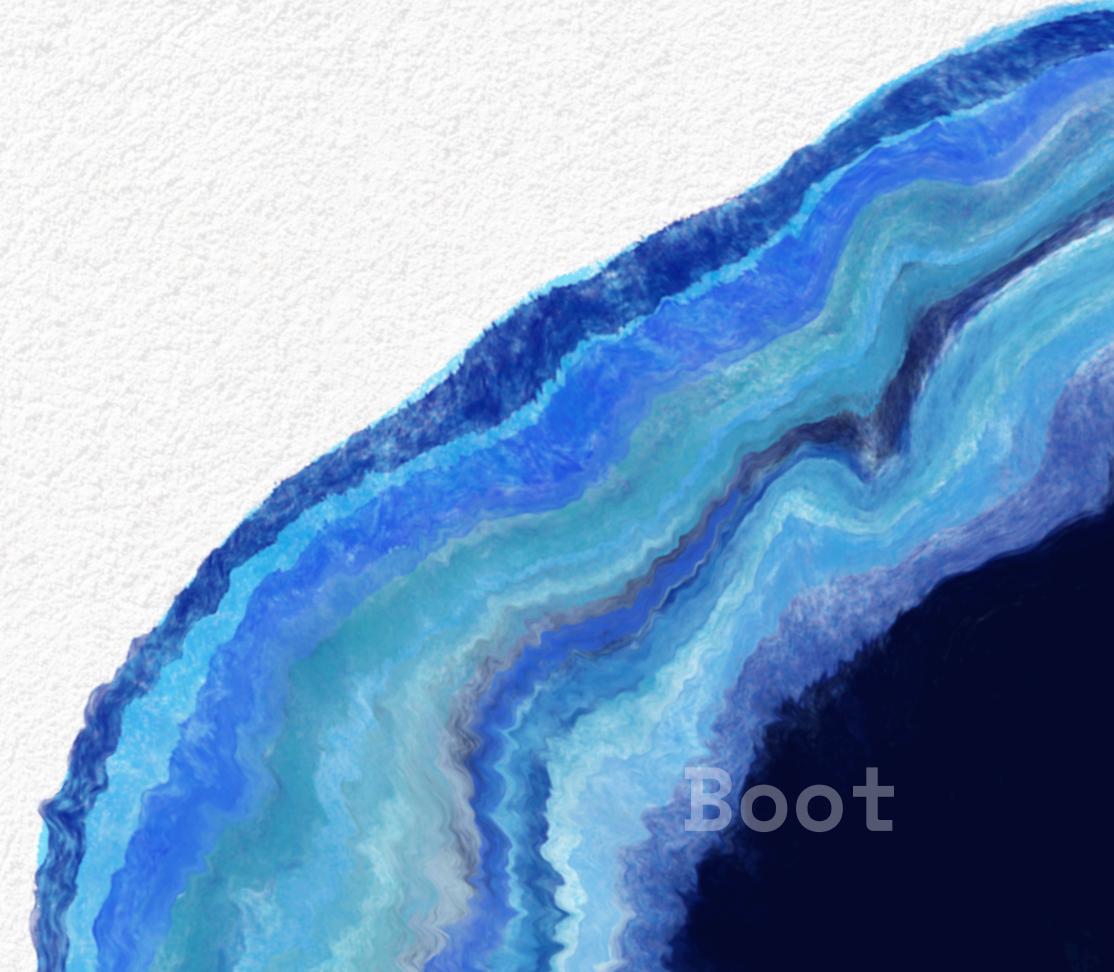
O que é Bag of Words

- Count Vectorizer
- TFIDF Vectorizer
- Hashing Vectorizer

A large, abstract graphic on the left side of the slide features a series of concentric, wavy bands in shades of blue, teal, and white, resembling a cross-section of a planet or a stylized ocean current.

5

Modelo Inicial PLN

A large, abstract graphic on the right side of the slide features a series of concentric, wavy bands in shades of blue, teal, and white, resembling a cross-section of a planet or a stylized ocean current.

Boot

Naive Bayes

- Multinomial ou Normal
- Robustez a irrelevância
- Sensibilidade a frequência

Resultados do Modelo

```
train_naive_bayes(df_countVect_stemmed, "classification")
[140] ✓ 0.0s
...
(GaussianNB(), 0.6482965931863728)
```

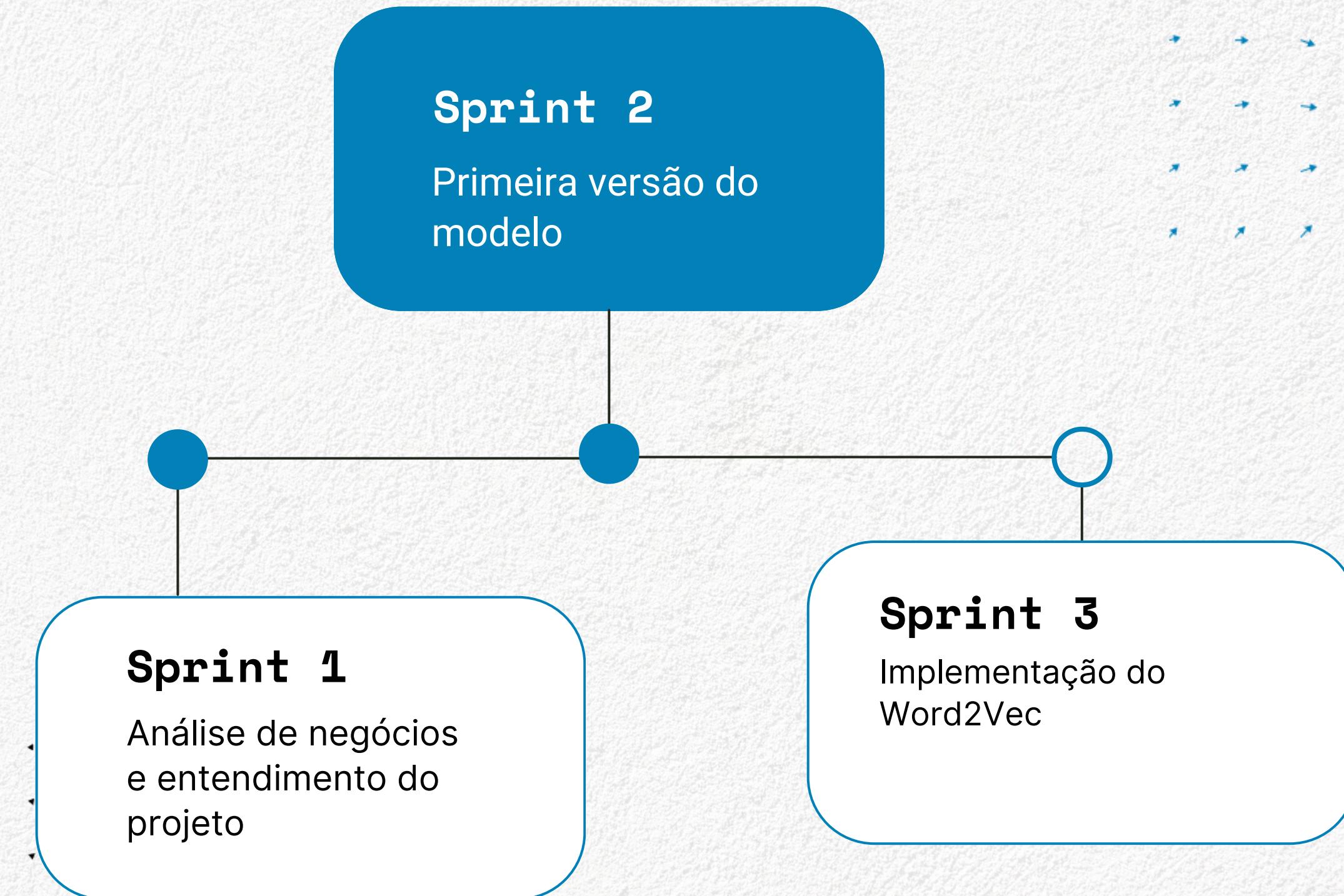
Acurácia: 64,82%

```
train_multinomial_naive_bayes(df_countVect_stemmed, "classification")
[175] ✓ 0.0s
...
(MultinomialNB(), 0.6472945891783567)
```

Acurácia: 64,72%

Boot

I



Boot



Boot



Matheus R.



Henrique Cox



Eduardo HOS



Mateus M.



Marcelo Saasi



Otto



Celine



Boot