

# Execício\_Aula\_08

November 6, 2023

## 1 Base de conhecimento “Music Dataset: Song Information and Lyrics”

## 2 Preparação de ambiente

```
[1]: !pip install spacy --upgrade
    !python -m spacy download en_core_web_sm

import spacy
import en_core_web_sm
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
import re
import random
from google.colab import drive
from sklearn.model_selection import train_test_split
from sklearn import tree
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score, confusion_matrix
```

Requirement already satisfied: spacy in /usr/local/lib/python3.10/dist-packages (3.6.1)

Collecting spacy

Downloading

spacy-3.7.2-cp310-cp310-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (6.6 MB)

6.6/6.6 MB

19.6 MB/s eta 0:00:00

Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.10/dist-packages (from spacy) (3.0.12)

Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.0.5)

Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.10/dist-packages (from spacy) (1.0.10)

Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.10/dist-packages (from spacy) (2.0.8)

Requirement already satisfied: preshed<3.1.0,>=3.0.2 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (3.0.9)

Requirement already satisfied: thinc<8.3.0,>=8.1.8 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (8.1.12)

Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (1.1.2)

Requirement already satisfied: srsly<3.0.0,>=2.4.3 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (2.4.8)

Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (2.0.10)

Collecting weasel<0.4.0,>=0.1.0 (from spacy)

  Downloading weasel-0.3.4-py3-none-any.whl (50 kB)

  50.1/50.1 kB

6.0 MB/s eta 0:00:00

Requirement already satisfied: typer<0.10.0,>=0.3.0 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (0.9.0)

Requirement already satisfied: smart-open<7.0.0,>=5.2.1 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (6.4.0)

Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (4.66.1)

Requirement already satisfied: requests<3.0.0,>=2.13.0 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (2.31.0)

Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (1.10.13)

Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages  
 (from spacy) (3.1.2)

Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-  
 packages (from spacy) (67.7.2)

Requirement already satisfied: packaging>=20.0 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (23.2)

Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in  
 /usr/local/lib/python3.10/dist-packages (from spacy) (3.3.0)

Requirement already satisfied: numpy>=1.19.0 in /usr/local/lib/python3.10/dist-  
 packages (from spacy) (1.23.5)

Requirement already satisfied: typing-extensions>=4.2.0 in  
 /usr/local/lib/python3.10/dist-packages (from  
 pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (4.5.0)

Requirement already satisfied: charset-normalizer<4,>=2 in  
 /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy)  
 (3.3.1)

Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-  
 packages (from requests<3.0.0,>=2.13.0->spacy) (3.4)

Requirement already satisfied: urllib3<3,>=1.21.1 in  
 /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy)  
 (2.0.7)

Requirement already satisfied: certifi>=2017.4.17 in  
 /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy)  
 (2023.7.22)

```
Requirement already satisfied: blis<0.8.0,>=0.7.8 in
/usr/local/lib/python3.10/dist-packages (from thinc<8.3.0,>=8.1.8->spacy)
(0.7.11)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in
/usr/local/lib/python3.10/dist-packages (from thinc<8.3.0,>=8.1.8->spacy)
(0.1.3)
Requirement already satisfied: click<9.0.0,>=7.1.1 in
/usr/local/lib/python3.10/dist-packages (from typer<0.10.0,>=0.3.0->spacy)
(8.1.7)
Collecting cloudpathlib<0.17.0,>=0.7.0 (from weasel<0.4.0,>=0.1.0->spacy)
  Downloading cloudpathlib-0.16.0-py3-none-any.whl (45 kB)
                                45.0/45.0 kB
```

5.4 MB/s eta 0:00:00

```
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.10/dist-packages (from jinja2->spacy) (2.1.3)
Installing collected packages: cloudpathlib, weasel, spacy
  Attempting uninstall: spacy
    Found existing installation: spacy 3.6.1
    Uninstalling spacy-3.6.1:
      Successfully uninstalled spacy-3.6.1
```

ERROR: pip's dependency resolver does not currently take into account all  
the packages that are installed. This behaviour is the source of the following  
dependency conflicts.

en-core-web-sm 3.6.0 requires spacy<3.7.0,>=3.6.0, but you have spacy 3.7.2  
which is incompatible.

```
Successfully installed cloudpathlib-0.16.0 spacy-3.7.2 weasel-0.3.4
2023-11-06 22:33:14.062024: E
tensorflow/compiler/xla/stream_executor/cuda/cuda_dnn.cc:9342] Unable to
register cuDNN factory: Attempting to register factory for plugin cuDNN when one
has already been registered
2023-11-06 22:33:14.062118: E
tensorflow/compiler/xla/stream_executor/cuda/cuda_fft.cc:609] Unable to register
cuFFT factory: Attempting to register factory for plugin cuFFT when one has
already been registered
2023-11-06 22:33:14.062174: E
tensorflow/compiler/xla/stream_executor/cuda/cuda_blas.cc:1518] Unable to
register cuBLAS factory: Attempting to register factory for plugin cuBLAS when
one has already been registered
2023-11-06 22:33:14.075956: I tensorflow/core/platform/cpu_feature_guard.cc:182]
This TensorFlow binary is optimized to use available CPU instructions in
performance-critical operations.
To enable the following instructions: AVX2 FMA, in other operations, rebuild
TensorFlow with the appropriate compiler flags.
2023-11-06 22:33:17.030870: W
tensorflow/compiler/tf2tensorrt/utils/py_utils.cc:38] TF-TRT Warning: Could not
```

```

find TensorRT
Collecting en-core-web-sm==3.7.0
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_sm-3.7.0/en_core_web_sm-3.7.0-py3-none-
any.whl (12.8 MB)
                                12.8/12.8 MB
37.9 MB/s eta 0:00:00
Requirement already satisfied: spacy<3.8.0,>=3.7.0 in
/usr/local/lib/python3.10/dist-packages (from en-core-web-sm==3.7.0) (3.7.2)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (1.0.10)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (2.0.8)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (3.0.9)
Requirement already satisfied: thinc<8.3.0,>=8.1.8 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (8.1.12)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (1.1.2)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (2.4.8)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (2.0.10)
Requirement already satisfied: weasel<0.4.0,>=0.1.0 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (0.3.4)
Requirement already satisfied: typer<0.10.0,>=0.3.0 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (0.9.0)
Requirement already satisfied: smart-open<7.0.0,>=5.2.1 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (6.4.0)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in
/usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-
sm==3.7.0) (4.66.1)

```

Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (2.31.0)

Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (1.10.13)

Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (3.1.2)

Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (67.7.2)

Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (23.2)

Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (3.3.0)

Requirement already satisfied: numpy>=1.19.0 in /usr/local/lib/python3.10/dist-packages (from spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (1.23.5)

Requirement already satisfied: typing-extensions>=4.2.0 in /usr/local/lib/python3.10/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (4.5.0)

Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (3.3.1)

Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (3.4)

Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (2.0.7)

Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (2023.7.22)

Requirement already satisfied: blis<0.8.0,>=0.7.8 in /usr/local/lib/python3.10/dist-packages (from thinc<8.3.0,>=8.1.8->spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (0.7.11)

Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.10/dist-packages (from thinc<8.3.0,>=8.1.8->spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (0.1.3)

Requirement already satisfied: click<9.0.0,>=7.1.1 in /usr/local/lib/python3.10/dist-packages (from typer<0.10.0,>=0.3.0->spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (8.1.7)

Requirement already satisfied: cloudpathlib<0.17.0,>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from weasel<0.4.0,>=0.1.0->spacy<3.8.0,>=3.7.0->en-core-web-sm==3.7.0) (0.16.0)

Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from jinja2->spacy<3.8.0,>=3.7.0->en-

```

core-web-sm==3.7.0) (2.1.3)
Installing collected packages: en-core-web-sm
  Attempting uninstall: en-core-web-sm
    Found existing installation: en-core-web-sm 3.6.0
    Uninstalling en-core-web-sm-3.6.0:
      Successfully uninstalled en-core-web-sm-3.6.0
Successfully installed en-core-web-sm-3.7.0
Download and installation successful
You can now load the package via spacy.load('en_core_web_sm')

```

```

[57]: nlp = spacy.load('en_core_web_sm')

drive.mount('/content/drive', force_remount=True)
songs = pd.read_csv('/content/drive/MyDrive/PLN/movie.original.csv',
                    header=0,
                    skiprows=1,
                    names=['Texto', 'Classificacao'], encoding='latin1')

nlp

```

Mounted at /content/drive

```
[57]: <spacy.lang.en.English at 0x7e976b846770>
```

```
[58]: songs
```

```

[58]:

```

	Texto	Classificacao
0	When I put this movie in my DVD player, and sa...	0
1	Why do people who do not know what a particula...	0
2	Even though I have great interest in Biblical ...	0
3	Im a die hard Dads Army fan and nothing will e...	1
4	A terrible movie as everyone has said. What ma...	0
...	...	...
39994	"Western Union" is something of a forgotten cl...	1
39995	This movie is an incredible piece of work. It ...	1
39996	My wife and I watched this movie because we pl...	0
39997	When I first watched Flatliners, I was amazed...	1
39998	Why would this film be so good, but only gross...	1

[39999 rows x 2 columns]

### 3 Manipulação da Base de conhecimento

```

[77]: x = songs.iloc[:,0].values
      y = songs.iloc[:,1].values

x,_, y,_ = train_test_split(x, y, test_size=0.98)

```

```
len(x), len(y)
```

```
[77]: (799, 799)
```

```
[74]:
```

```
[74]: False
```

```
[32]: def preprocessing(sentence):  
    sentence = re.sub(r'@[A-Za-z0-9]+', ' ', str(sentence))  
  
    sentence = sentence.lower()  
    tokens = [token.text for token in nlp(sentence) if not (token.is_stop  
                                                            or token.like_num  
                                                            or token.is_punct  
                                                            or token.is_space  
                                                            or len(token) == 1)]  
  
    tokens = ' '.join(tokens)  
  
    return tokens
```

```
[78]: x_songs_cleaned = [preprocessing(songs) for songs in x]  
      x_teste_cleaned = [preprocessing(songs) for songs in x]
```

## 4 Treinamento de árvore

```
[37]: x_songs_cleaned[0:5]
```

```
[37]: ['watching dick tracy years result vital life childhood seeing adult good movie  
dark satiric incredibly misunderstood thing said oscar nomination pacino  
received rarely discussed fuss came out.<br /><br />pacino good effect clearly  
having loads fun beatty great dick tracy camera manages capture atmosphere film  
noir comic book better film possibly seen taking look scene film breathtaking  
lighting velvet overtones smog smoke combine create great effect.<br /><br  
>there funny cameos including dustin hoffman mumbles think flaws terms acting  
mandatory kid character far better expected.<br /><br />overall fine movie  
misunderstood years release incredibly underrated average imdb critics reviews  
positive check rottentomatoes.com seeing film hard perfect example capturing  
essence comic book style eccentricity.<br /><br />highly recommended stars',  
'enjoyable fast paced br /><br />there way earth actor playing mat main thing  
act convincingly credit audition convinced cast soon accepted naive young  
country boy.<br /><br />while best performance engaging particular interplay  
policemen natural balanced worked well.<br /><br />it minutes long plot complex  
key style thing slick vibrant backdrops atmospheric especially fact colours  
extremely rich gangland identifiable foreign audiences manages distinctly  
australian',  
'filmmaker wanted movie having story tell awful jumble unlikely unexplained
```

coincidences unidentifiable plot line character clear motivation.<br /><br />we clichã snapshots instead characters particular diminutive beautiful crime boss projects overdone tough guy persona casts cartoonish shadow intimidation actual tough guys brought work startling look film shot boys hit road carries tiny suitcase smallest complete american tourist set bright sky blue explanation apology standard visually exception compelling shot beautiful bridge ct',  
 'true gothic horror trash classic!<br /><br />uhm actually horrible movie best thing rosalba neri erected naked nipples intensely suckable material.<br /><br />oh sure rosalba neri fine lady heard flick damn fine looking lady honestly nipples scene movie abruptly ends naked erected total dream nipples man okay sorry getting little carried away <br /><br />basically got expected lady frankenstein sleazy horrible flick big stupid ugly looking frankenstein monster couple naked tits got castle disappointed dragged certain places resurrection creature kept going ages daddy frankenstein experimenting talking trying fun imbecilic homage original frankenstein instead throwing little girl lake mongoloid franky picked naked chick threw river laughing.<br /><br />some friends raved tad bit flick maybe expecting bit rosalba erected nipples sure worth aw crap need stop mentioning nipples think end user comment',  
 'got ta fan little man found burlesque carmen dull unimaginative totally funny.<br /><br />chaplin retelling story carmen plays big role don jose story men women love unclear love woman carmen playing men other.<br /><br />as said think chaplin dozens dozens better movies leave shelve.<br /><br />oh curious difference movie version imdb mistake?<br /><br />on']

```
[38]: len(y)
```

```
[38]: 1249
```

```
[79]: from sklearn.feature_extraction.text import TfidfVectorizer
```

```
vectorizer = TfidfVectorizer()
x_songs_tfidf = vectorizer.fit_transform(x_songs_cleaned)
```

```
[40]: x_songs_tfidf.toarray().shape
```

```
[40]: (1249, 19979)
```

```
[43]: def preprocessing_lemma(sentence):
      tokens = [token.lemma_ for token in nlp(sentence)]
      tokens = ' '.join(tokens)

      return tokens
```

```
[80]: x_songs_cleaned_lemma = [preprocessing_lemma(songs) for songs in
      ↪ x_songs_cleaned]
      x_songs_tfidf = vectorizer.fit_transform(x_songs_cleaned_lemma)
```



```
x_teste_cleaned_lemma = [preprocessing_lemma(songs) for songs in x_teste_cleaned]
x_teste_tfidf = vectorizer.transform(x_teste_cleaned_lemma)

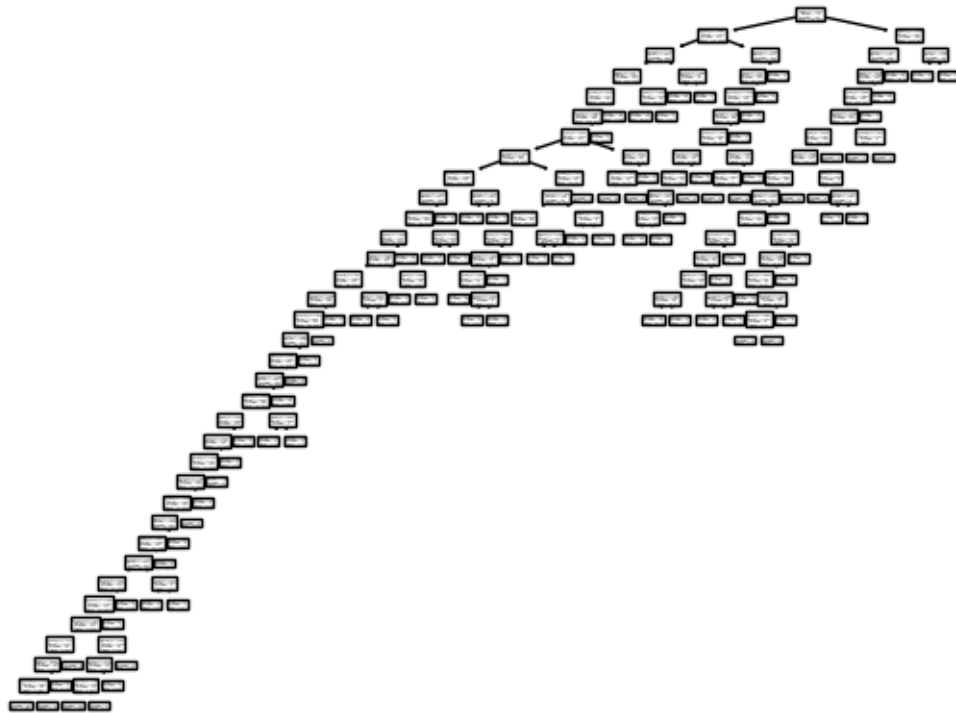
x_songs_tfidf.shape, x_teste_tfidf.shape
```

```
[80]: ((799, 13071), (799, 13071))
```

```
[81]: model = tree.DecisionTreeClassifier(criterion='entropy')
model.fit(x_songs_tfidf, y)
```

```
[81]: DecisionTreeClassifier(criterion='entropy')
```

```
[82]: plt.show(tree.plot_tree(model))
```



```
[83]: predictions = model.predict(x_teste_tfidf)
predictions
```

```
[83]: array([0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1,
        0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 1, 1,
        0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 0, 0,
        1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1,
```

```

0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1,
0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0,
1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1,
1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 0, 1,
0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1,
0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1,
1, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1,
1, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0,
0, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0,
0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0,
1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1,
0, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 1,
0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0,
0, 1, 1, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1,
1, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0,
1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 1,
1, 0, 1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0,
0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0,
0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1,
0, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 0, 1, 1,
0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1,
0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0,
0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1,
0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1,
0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1,
0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1,
0, 0, 0, 1, 1, 0, 1])

```

```
[84]: accuracy_score(y, predictions)
```

```
[84]: 1.0
```

```
[85]: cm = confusion_matrix(y, predictions)
      cm
```

```
[85]: array([[396,  0],
           [  0, 403]])
```

```
[89]: def avalue(sentence):
      x_teste_cleaned.append(sentence)
```

```

x_teste_cleaned_lemma = [preprocessing_lemma(sentence) for sentence in
↪x_teste_cleaned]
x_teste_tfidf = vectorizer.transform(x_teste_cleaned_lemma)

predictions = model.predict(x_teste_tfidf)
score = predictions[-1]

print(f'Classificação: {score}')

if score == 0:
    return 'frase negativa'

return 'frase positiva'

```

```

[91]: frase_usuario = input('Digite uma frase: ')

avaliacao = avalie(frase_usuario)

print(f'A frase "{frase_usuario}" é uma {avaliacao}'.title())

```

```

Digite uma frase: this is a beautiful movie, i love it
Classificação: 1
A Frase "This Is A Beautiful Movie, I Love It" É Uma Frase Positiva

```