Song Genre AI

Dedukowanie gatunków muzycznych
na podstawie tekstu piosenki

Paweł Gościak, Marceli Grad, Jan Krawczyk

Spis treści

Wprowadzenie

Celem naszego projektu jest stworzenie sztucznej inteligencji, która będzie w stanie rozpoznawać gatunek muzyczny na podstawie tekstu piosenki. Muzyka od zawsze była nieodłączną częścią naszego życia, a różnorodność gatunków muzycznych jest ogromna. Nasza sztuczna inteligencja ma za zadanie automatycznie analizować teksty piosenek i przypisywać im odpowiedni gatunek muzyczny, uwzględniając różnorodne cechy takie jak tematykę, styl oraz ilość słów.

Dane

Baza danych

Dane, które wykorzystaliśmy do uczenia naszej sztucznej inteligencji zostały pobrane ze strony www.kaggle.com. Znajduje się tam duża baza artystów wraz z gatunkami wykonywanych przez nich utworów oraz baza piosenek, które zawiera teksty. Po połączeniu obu baz oraz odrzuceniu nieinteresujących nas danych dokonaliśmy lematyzacji pozostałych tekstów.

Lematyzacja

Lematyzacja to proces sprowadzania słów do ich podstawowej formy, zwanej lematem, poprzez usunięcie odmiany fleksyjnej oraz uwzględnienie znaczeniowych i gramatycznych właściwości danego słowa.

```
I feel so unsure
As I take your hand and lead you to the dance floor
As the music dies, something in your eyes
Calls to mind a silver screen
And all those sad goodbyes

I'm never gonna dance again
Guilty feet have got no rhythm
Though it's easy to pretend
I know you're not a fool

Should've known better than to cheat a friend
And waste the chance that I've been given
So I'm never gonna dance again
The way I danced with you

Time can never mend
The careless whispers of a good friend
To the heart and mind
Ignorance is kind
There's no comfort in the truth
Pain is all you'll find
```

```
I feel so unsure
as I take your hand and lead you to the dance floor
as the music die , something in your eye
call to mind a silver screen
and all those sad goodbye

I be never go to dance again
guilty foot have get no rhythm
though it be easy to pretend
I know you be not a fool

should 've know well than to cheat a friend
and waste the chance that I 've be give
so I be never go to dance again
the way I dance with you

Time can never mend
the careless whisper of a good friend
to the heart and mind
ignorance be kind
there be no comfort in the truth
Pain be all you will find
```

zdj.1 “Careless whisper” przed i po lematyzacji

Dalsza obróbka tekstu

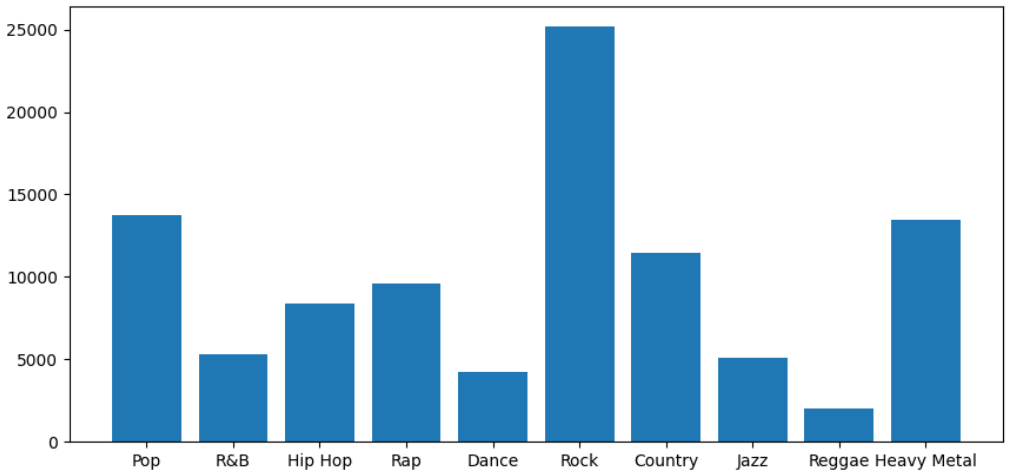
Następnym krokiem, który wykonaliśmy było odrzucenie słów, które z racji swojego znaczenia nie mogły jednoznacznie wskazywać na dany gatunek muzyczny oraz odrzucenie wszystkich znaków interpunkcyjnych i symboli muzycznym.

```
['the', 'as', 'i', 'be', 'a', 'you', 'to', 'and', 'it', 'not', 'do', 'in', 'my', 'us', 'of',
 'your', 'know', "'", 'so', 'love', 'but', 'no', 'yes', '?', 'he', 'she', 'we', 'make', 'if',
 "'ve", 'want', '!', 'well', '"', 'could', 'from', 'would', "'s", 'at', '...', 'her', 'his',
 'all', 'around', 'then', 'when', 'they', 'them', 'into', 'an', ':', 'their', 'those', 'these',
 'this', 'mine', 'too', 'through', 'who', 'how', 'why', 'until', 'unless', 'that', 'with', 'on',
 'or', 'will', "won't", "can't", "haven't", "isn't", 'have', 'what', 'by', 'there', 'here',
 'which', 'whom', 'whose', 'some', 'than', 'like', 'also', 'because', '!', 'each', 'during', '(',
 ')', '[', ']', u"\u2122", 'soon', 'although', 'however', 'let', 'get', 'go', 'come', 'can',
 'take', 'our', '.', '..', '*', '_', '+', '/', 'a1', 'a2', 'a3', 'a4', 'a5', 'a6', 'a7', 'b1', 'b2',
     'b3', 'b4', 'b5', 'b6', 'b7', 'c1', 'c2', 'c3', 'c4', 'c5', 'c6', 'c7', 'd1', 'd2', 'd3', 'd4', 'd5',
     'd6', 'd7', 'e1', 'e2', 'e3', 'e4', 'e5', 'e6', 'e7', 'f1', 'f2', 'f3', 'f4', 'f5', 'f6', 'f7', 'g1',
     'g2', 'g3', 'g4', 'g5', 'g6', 'g7', 'h1', 'h2', 'h3', 'h4', 'h5', 'h6', 'h7']
```

zdj.2 słowa usunięte z tekstów

Produkt końcowy

Na końcu naszej obróbki bazy danych otrzymaliśmy następujące dane:



zdj.3 rozłożenie tekstów według gatunków

```
Pop  =  13759
R&B  =  5309
Hip Hop  =  8412
Rap  =  9589
Dance  =  4252
Rock  =  25177
Country  =  11432
Jazz  =  5124
Reggae  =  1990
Heavy Metal  =  13496
```

zdj.4 rozłożenie tekstów według gatunków

Wykonanie

Naiwny klasyfikator Bayesa

Do wykonania naszej sztucznej inteligencji wykorzystaliśmy naiwny klasyfikator Bayes, który wykorzystując twierdzenie Bayes, wyznacza prawdopodobieństwo należenie piosenki do konkretnego gatunku muzycznego na podstawie częstotliwości występowania słów w konkretnych gatunkach muzycznych oraz częstotliwości występowania gatunku muzycznego w naszej bazie.

Np. prawdopodobieństwo, że piosenka o treści “love her” będzie należała do gatunku pop wynosi:

$$P("pop") = P("pop\ gerne") \times P("love") \times P("her")$$

gdzie:

$$P("love") = \frac{"love"}{"suma\ słów\ w\ piosenkach\ typu\ "pop"}$$

$P("pop\ gerne")$ - prawdopodobieństwo piosenki bycia piosenką pop wśród bazy treningowej
$P("love")$ - prawdopodobieństwo wystąpienia słowa “love” wśród wszystkich piosenek typu pop w bazie trenigowej
$P("her")$ - prawdopodobieństwo wystąpienia słowa “her” wśród wszystkich piosenek typu pop w bazie trenigowej.

W celu implementacji tego algorytmu stworzyliśmy klasę, która zawierała następujące pola:

```
Naive bayes algorithm


Class deciding which genre song belongs to base
on popular words in different genres.


train_data: data used for training, type: pandas.core.frame.DataFrame
categories: list of different music genre, type: list
genre probability: probability of occurrence of different music genre in base, type: dict
word_genre_dictionaries: number of occurrences of words in different genres, type: dict
total_genre_words: total number of words in the genre, type: dict
"""
```

zdj.5 opis klasy

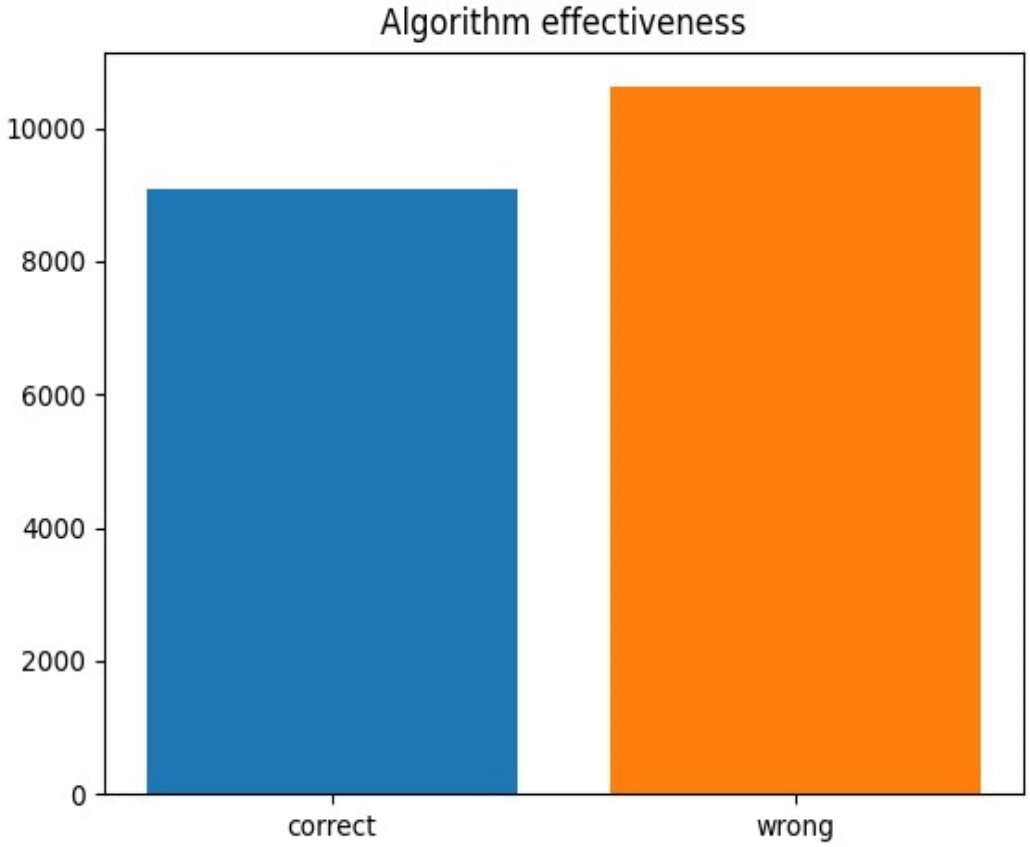
Następnie korzystając z biblioteki counter zliczyliśmy ilość wystąpień wszystkich słów w danych gatunkach muzycznych.

Do procesu uczenia naszego algorytmu wykorzystaliśmy 80% przygotowanej przez nas bazy, podczas gdy pozostałe 20% przeznaczyliśmy na testy.

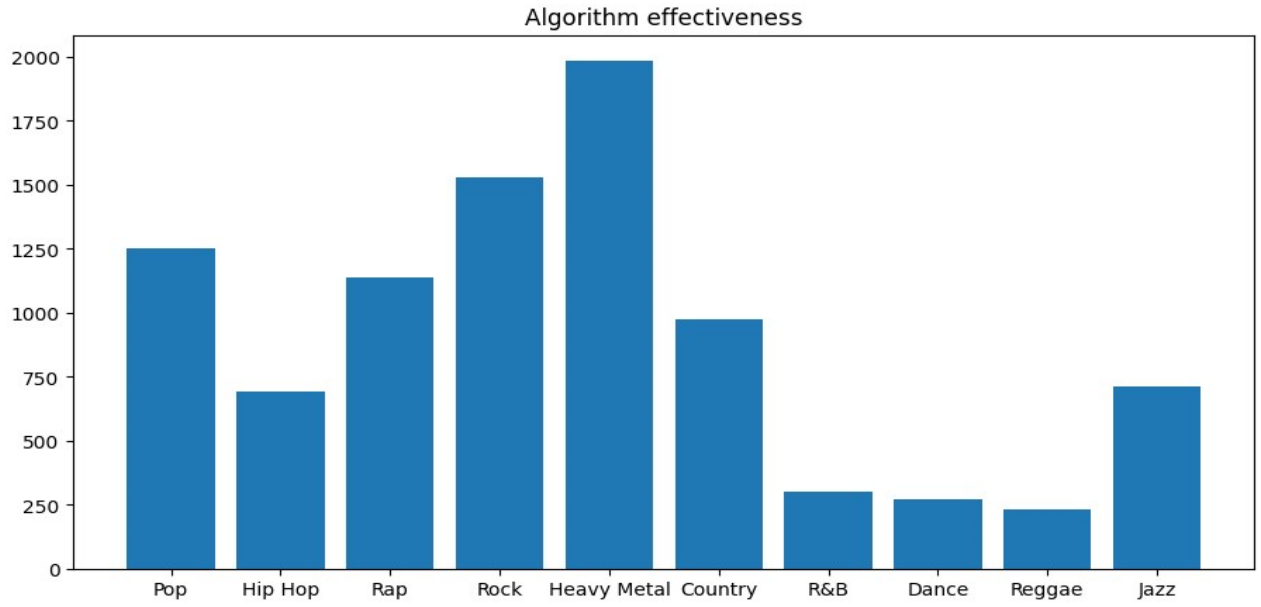
Wyniki

Otrzymany wynik

Po implementacji naiwnego klasyfikatora Bayesa uzyskana przez nas skuteczność wynosiła około 46%. Wynik ten uznaliśmy za zadowalający ze względu na fakt, że wybór losowej odpowiedzi spośród podanych zapewniał nam jedynie około 10% trafień.



zdj.6 graficzna reprezentacja skuteczności algorytmu



zdj. 7 graficzna reprezentacja skuteczności algorytmu według gatunków muzycznych

Potencjalne sposoby ulepszenia

Mimo, że nasz algorytm osiągnęła wyniki, które uznaliśmy za zadowalające, istnieją sposoby, żeby ulepszyć potencjalnie jego działanie takie jak:

- powiększenie bazy danych
- skorzystanie z języka programowania oferującego większą dokładność operacji na liczbach zmiennoprzecinkowych
- skorzystanie z bazy danych, która przyporządkowuje gatunek do utworu, a nie do artysty

Podsumowanie

Podsumowując naiwny klasyfikator Bayesa jest jednym z najlepszych algorytmów do zastosowania w przypadku naszego projektu, ponieważ jego skuteczność znacznie przewyższa losowe dobieranie gatunków do tekstów piosenek.

Pozyskana wiedza

Podczas wykonywania tego projektu przestudiowaliśmy wiele zagadnień dotyczących obróbki tekstu, lematyzacji, twierdzenia bayesa oraz serializacji instancji klasy. Poznaliśmy wiele niezbędnych bibliotek, takich jak pandas - do obróbki plików csv, pickle - do serializacji instancji klasy czy tkinter - do wykonywania interfejsów graficznych.

Pokonane problemy

W trakcie implementacji napotkaliśmy kilka problemów, które udało nam się pokonać, w trakcie wykonania projektu. Przykładami takich problemów są:

- ograniczenia plików pickle - ponieważ nasza klasa przechowywała bardzo dużo słowników, to nie byliśmy w stanie zserializować ich wszystkich przy pomocy samej klasy pickle. Z tego powodu skorzystaliśmy również z plików w rozszerzeniu json.
- wybór biblioteki do wykonania lematyzacji - ponieważ proces lematyzacji dla naszej początkowej bazy danych zajął nam 72 godziny, musieliśmy przetestować wiele bibliotek oraz wybrać taką, która najlepiej balansuje wydajność i skuteczność.

Bibliografia

- https://betterprogramming.pub/predicting-a-songs-genre-using-natural-language-processing-7b354ed5bd80 - inspiracja
- https://www.kaggle.com/datasets/neisse/scrapped-lyrics-from-6-genres?select=lyrics-data.csv - dane
- https://looka.com/ - logo