

4. Korzystając z wbudowanego w Gensim [mechanizmu wykrywania wyrażen wielosegmentowych](#) wykryj frazy składające się z maksymalnie 3 słów, korzystając z domyślnych ustawień biblioteki.

```
def detect_phrases():
    global sentencesAll, trigramSentences
    print("detect_phrases")
    sentences = judgmentsText.split('.')
    for sentence in sentences:
        sentencesAll.append(sentence.strip().split())
    phrases = Phrases(sentencesAll)
    bigram = Phraser(phrases)
    trigram = Phrases(bigram[sentencesAll])
    for sent in sentencesAll:
        trigramSentences.append(trigram[bigram[sent]])
```

5. Dla tak przygotowanego zbioru danych wytrenuj model word2vec o następujących parametrach:

```
model = Word2Vec(trigramSentences, sg=0, window=5, size=300, min_count=3)
```

7. Znajdź 3 najbardziej podobne wyrażenia (słowa i zwroty) dla następujących wyrażen:

nie znalazło mi wyrażenia „szkoda majątkowa”.

Sąd_Najwyższy

```
[('Sąd_Najwyższy', 0.8129898905754089), ('Trybunał_Konstytucyjny', 0.6764004230499268), ('Trybunał', 0.6716378927230835)]
```

Trybunał_Konstytucyjny

```
[('Trybunał', 0.93972247838974), ('Trybunał_Konstytucyjny', 0.8215391635894775), ('Trybunał', 0.813628077507019)]
```

kodeks_cywilny

```
[('Kodeks_pracy', 0.7980093955993652), ('Kodeks_cywilny', 0.7941955327987671), ('kodeks_postępowania_cywilnego', 0.7539807558059692)]
```

kpk

```
[('kpk', 0.8113726377487183), ('kpk', 0.7931006550788879), ('kk', 0.7911224365234375)]
```

sąd_rejonowy

```
[('sąd_apelacyjny', 0.7189524173736572), ('sąd', 0.7032870650291443), ('sąd_właściwy', 0.6818022131919861)]
```

szkoda

```
[('krzywda', 0.7928489446640015), ('szkoda', 0.7827692031860352), ('Szkoda', 0.7716143727302551)]
```

wypadek

```
[('przypadek', 0.7776857614517212), ('wypadek', 0.7367169857025146), ('Wypadek', 0.6562826633453369)]
```

kolizja

[('dysproporcja', 0.7396142482757568), ('relacja', 0.7084054350852966), ('luka', 0.7073505520820618)]

rozwód

[('rozwód', 0.7571244835853577), ('separację', 0.7037941813468933), ('przez_rozwód', 0.6236369013786316)]

nieszczęście

[('480;', 0.7080917954444885), ('360;', 0.7026132941246033), ('męskie', 0.7011219263076782)]

8. Znajdź wypadkową operacji na słowach (5 najbliższych wyników)

i. Sąd Najwyższy - kpc + konstytucja

[('Konstytucja', 0.49939578771591187), ('Naczelny_Sąd_Administracyjny', 0.49430277943611145), ('Trybunał_wielokrotnie', 0.4905400276184082), ('Trybunał_Konstytucyjny', 0.4838156998157501), ('Trybunał', 0.4765018820762634)]

i. pasażer - mężczyzna + kobieta

[('pasażerka', 0.7400874495506287), ('obywatelka', 0.6115729808807373), ('żona_powoda', 0.6009453535079956), ('opiekunka', 0.5974345207214355), ('ofiara', 0.5924198627471924)]

i. samochód - droga + rzeka

[('telefon_komórkowy', 0.6609154939651489), ('pistolet', 0.6444844007492065), ('samochód_marki', 0.6388851404190063), ('Samochód', 0.6273850798606873), ('ciągnik', 0.6146700382232666)]