**  
  
Eksploracja danych tekstowych - laboratoria 9: Topic modelling**

*Do realizacji zadań wykorzystaj środowisko jupyter notebook. Po zakończeniu zadań wyślij odpowiedzi na adres michal.zarnecki@cdv.pl.*

**1. Analiza sentymentu**

1. Oceń czy następujące zdania mają wydźwięk: pozytywny, negatywny czy neutralny.

a) Jestem bardzo szczęśliwy :)

b) NIENAWIDZĘ swojej pracy!!!

c) Skyr to islandzki produkt mleczny o konsystencji podobnej do jogurtu lub twarogu, ale o wyjątkowym smaku i wartości odżywczej.

d) Kocham pizzę!

e) To już trzeci raz w tym miesiącu jak mój samochód się zepsuł :(

f) "Skrrrt" to onomatopeja używana w języku potocznym i w kulturze hip-hopowej jako dźwięk opisujący krótki, szybki odgłos opon samochodowych, gdy pojazd gwałtownie hamuje lub przyspiesza.

2. Postaraj się stworzyć najbardziej pozytywne zdanie jak to tylko możliwe. Przetłumacz je na język angielski (możesz skorzystać ze strony www.deepl.com) oraz oceń sentyment zdania za pomocą kodu w jupyter notebook’u notebooku sentiment\_analysis.ipynb. Czy uda CI się osiągnąć wynik powyżej 0.95 dla współczynnika “compound”?

**2. Modelowanie Tematyczne**

1. Przeprowadziłeś Modelowanie Tematyczne na zbiorze tekstów. Jak mogłyby się nazywać tematy, które zostały przyporządkowane przez algorytm? Oto 10 najczęściej występujących słów dla każdego z 5-ciu wybranych tematów:

a) Technologia, Przyszłość, Haker, Android, Mroczny, Miasto, Komputer, Dystopia, Wirtualność, Transhumanizm

b) Sakura, Tokio, Sushi, Kultura, Kyoto, Tempel, Gejsza, Fuji, Hiroshima, Zen

c) Atom, Energia, Jądrowa, Reaktor, Plazma, Neutron, Deuter, Trit, Ciepło, Fuzja

d) Tradycja, Para, Zaproszenie, Muzyka, Taniec, Alkohol, Rodzina, Ślub, Wujek, Młodzi

e) Detektyw, Zbrodnia, Mroczny, Femme, Gangster, Intryga, Suspens, Śledztwo, Kryminał, Deszcz