

LABORATORIUM 8

Modelowanie oznaczania części mowy

1. Sprawdzić, jak dobrze działa kod z wykładu 9 na własnych przykładach z zestawu danych korpus Browna (na ocenę 3).
2. Sprawdzić za pomocą algorytmu naprzód-wstecz lub Viterbiego więcej niż jedno słowo na raz. (Zamiast prawdopodobieństwa oznaczenia przy danym poprzednim oznaczeniu, znajdowałibyśmy prawdopodobieństwo oznaczenia przy danych dwóch poprzednich oznaczeniach.) (na ocenę 4)
3. Wykorzystać obserwacje danych dziennika serwisu internetowego i oznaczać użytkowników jako będących w określonym stanie (klient, użytkownik, potencjalny klient) i przeanalizować zachowanie użytkowników i ustalić, kiedy przestają być zaangażowani. Dziennik serwisu internetowego sporządzić samodzielnie, wprowadzając wyżej wymienione dane ręcznie (wystarczy 10-20 obserwacji). (na ocenę 5+)