

Zadanie 2:

Wytrenować i porównać minimum 2 różne modele dla wybranego zadania NLP starając się osiągnąć jak najlepszy wynik.

Dostarczyć trzeba (na platformie MsTeams) program (bez danych) wraz z opisem, który musi zawierać:

- krótką charakterystykę zadania,
- charakterystykę wybranej metody realizacji, w szczególności opis różnic między wariantami (różne zestawy cech, wartości parametrów, architektura sieci ...)
- wyniki dla min. 2 testowanych rozwiązań (np. różne zestawy cech, parametrów lub różne metody); wyniki muszą być podane dla każdej z etykiet/klas oddzielnie
- o ile to możliwe, to porównanie z wynikami uczestników konkursu , lub z innymi znanymi wynikami dla tego zadania
- jakiś komentarz dotyczący zadania czy uzyskanych wyników (dobrych lub złych)

Ponadto trzeba przygotować i przedstawić prezentację projektu na ostatnich zajęciach (max. 10 minut)

I wariant

Realizacja któregoś z zadań konkursów Poleval - dowolny rok, np.. <http://2019.poleval.pl/>, <http://2018.poleval.pl/>, ew. inny własny temat uzgodniony wcześniej

II wariant

Rozpoznawanie i wyrażen temporalnych . Dane (przekonwertowana i nieco uproszczona zawartość KPWr) umieszczone są w MStTeams (grupa wykładowa, katalog 'lab' plik kpwr-time.zip) Dane podzielone są na treningowe i testowe. Znaczenie kolumn:

ORDER_ID - kolejny token

TOKEN_ID - kolejny token w zdaniu (zdania rozdzielone są pustymi liniami)

ORTH - segment

LEMMA - lemat

POS - część mowy

CTAG - pełen tag morfologiczny

LABEL - etykiety 'date', 'time', 'set', 'duration' w podziale na "B" i "I" oraz "O"