Spis treści:

1. Wstęp
   1. Podstawowe definicje
   2. Teza pracy
   3. Motywacja
2. Przegląd literatury
   1. Historia uczenia głębokiego i sieci neuronowych
   2. Przegląd zastosowania uczenia głębokiego do przetwarzania języka naturalnego
   3. Publikacje dotyczące wektoryzacji tekstu i nowe metody wektoryzacji
   4. Nowe architektury sieci neuronowych służących do PJN
   5. metody pomiaru efektywności algorytmów przetwarzających język naturalny
3. Zbiory danych
   1. Sarcasm detection
   2. distaster tweets calsification
   3. question answering
   4. tłumaczenie maszynowe (jeśli zdążę)
   5. klasyfikacja dla języka polskiego (jeśli zdążę)
4. Metody
   1. sposoby przygotowania i czyszczenia danych tekstowych
   2. techniki wektoryzacji tekstu
      1. tokenizacja, bag-of-words, one-hot matrix
      2. Word2Vec
      3. GloVe
      4. ELMO
   3. głębokie sieci neuronowe:
      1. podstawowe informacje dotyczące SN
      2. CNN,
      3. RNN,
      4. LSTM,
      5. BERT
5. Eksperyment
   1. opisy metod prepocesingu danych dla każdego zbioru
   2. opis sposobów wektoryzacji dla każdego zbioru
   3. opis architektur sieci dla każdego sposobu
6. Wyniki
7. Wnioski
8. Podsumowanie
9. Bibliografia

Wstęp

Teza