

Prepis hovorenej reči, podpora slovenského jazyka

Autor: Samuel Baran

Vedúci: RNDr. Erik Bruoth, PhD.

Motivácia

- ◇ Možnosť zorientovať sa v problematike rozpoznávania reči
- ◇ Získanie prehľadu o existujúcich nástrojoch na spracovanie prirodzeného jazyka
- ◇ Prehĺbenie poznatkov v oblasti hlbokého učenia (Deep learning)
- ◇ Osvojenie schopnosti aplikovať algoritmy hlbokého učenia
- ◇ Príprava východísk pre spracovávanie hovorenej podoby slovenského jazyka

Ciele

1. Analyzovať existujúce prístupy STT

- ◇ Získať prehľad v už existujúcich prístupoch STT konverzie
- ◇ Porovnať dané prístupy

Ciele

2. Vytvoriť tréningovú sadu pre úlohu pomocou dostupných zdrojov

- ◆ Datasets iniciatívy Mozilly Common Voice (je pripravovaný zber údajov pre slovenský jazyk, nevieme kedy bude sprístupnený dataset)
- ◆ Preprocessing audiokníh (rozdelenie audioknihy na menšie jednotky, napr. vety, priradenie príslušných textových reprezentácií viet)

Ciele

3. Identifikovať open source implementácie modelov (min. 2) rozpoznávania reči (STT)
4. Navrhnuť stratégiu tréningu modelov a ladenia hyperparametrov
5. Základné porovnanie modelov na základe vybraných metrík resp. popis hlavných problémov

Najbližšie kroky

- ◆ Naštudovať aktuálne prístupy spracovávanie prirodzeného jazyka s použitím hlbokého učenia
- ◆ Naštudovať open source implementáciu STT algoritmu z dielne Mozilly

Literatúra

- ◇ Dan Jurafsky and James H. Martin. Speech and Language Processing (3rd ed. draft https://web.stanford.edu/~jurafsky/slp3/edbook_oct162019.pdf)
- ◇ Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning (<http://www.deeplearningbook.org>)
- ◇ Jacob Eisenstein. Natural Language Processing
- ◇ <http://web.stanford.edu/class/cs224n/>
- ◇ <https://github.com/espnet/interspeech2019-tutorial>
- ◇ <https://github.com/mozilla/DeepSpeech> open source

Ďakujem za pozornosť