



WEB SYSTEM FOR DATA ANNOTATION

RÓZSA, HROMÁDKA, POSPÍŠILOVÁ, SCHWARZOVÁ



ZADANIE

- Pre mnohé úlohy strojového učenia je potrebné získať označované dáta. Vašou úlohou bude vytvoriť systém, ktorý umožní ich značkovanie. Anotátor odpovedá len na: otázka - odpoveď - hodnotenie (vlastná akcia). Aby sa zabránilo typickým problémom pri značkovaní je potrebné implementovať niekoľko akcií, ktoré napr. pridávajú šum, opakujú otázky, ... Na základe výsledkov, by mali byť prezentované základné štatistiky.
- Boli sme patricne zmateni

V PRAXI

- Web System For Data Annotation slúži na anotáciu dát. Dáta sa načítajú z textového súboru, program ich rozdelí do balíkov na anotovanie po 100 slovách. Administrátor určí prístupové práva jednotlivých užívateľov k balíkom na anotovanie. Po prihlásení užívateľ uvidí jemu prístupné balíky a môže začať s anotáciou. Má na výber označiť či slovo patrí do kategórie, nepatrí do kategórie alebo sa jedná o error. Taktiež má možnosť v prípade preklepu navrátiť sa o jedno slovo vzad a opakovane ho označkovat'. Takto označkové dáta sa potom uložia do výstupu vo formáte xml. Sú prístupné takisto štatistiky k označkovým dátam a jednotlivým anotátorom.

OD ZACIATKU

- Na porade sme si zvolili python a javascript, pretoze nam to prislo najjednoduchsie a jozi s tym uz robil, co sa mu potom moc neoplatilo :D

NAHRAVANIE SUBOROV

- Sexy screen nahravania suborov

UKAZKA NAHRAVANEHO TEXTAKU

- Sexy screen textaku



LOGIN

- Sexy screen loginu

ROZDELENIE BALIKOV

- Sexy screen rozdelenia balikov

SAMOTNE ANOTOVANIE

- Sexy screen anotovacej stránky

VYSTUP DAT

- Sexy screen vystupu dat

STATISTIKY

- Sexy screen statistik ak nejake spravime

VSETKO JE SAMOZREJME NA GITE

- Lebo sme mudri a pouzivali ho a teraz to mozete najst pod <https://github.com/Euklidex/WebSystemForDataAnnotation>

SEXY SCREEN GITU MOZNO

- Lebo sme sexy

ĎAKUJEME ZA POZORNOST

- Priestor na otázky