

# 17. Stretnutie 25.3.2020

## Agenda

1. Standup
2. Priprava dalsich taskov

## Priebeh stretnutia (hangouts call)

- Standup
  - Max
    - Kompresia pomocou RAR - vieme uložit 1GB do priblizne 50MB. Nedostatok miesta uz nebude problem
    - Task automaticka kompresia - opravenie chyby pri spustani skriptu
  - Veronika
    - Problem s api requestom v pythone - nejde to co v postmane ide
      - Riesenie postman generuje code snippet, ktory by mal fungovat. V pripade ze to nepojde napise do generalu
    - Blstakove NLP
      - Momentalne mame 20K inzeratov z bazosu a potrebujeme asi 6 hodin na spracovanie vsetkych API od Blastaka. Treba napisat mail ci by nam povolil viac znakov alebo mensi casovy rozostup
  - Brano
    - Uprava DB refill - momentalne je 4GB subor nahrany do databazy za 4 hodiny
    - Este su v systeme nejake issues
  - Ondrej
    - Alternativny stemmer - 6 jodin nieje az tak vela + nieje iste ako bude pracovat iny stemmer so slovincinou
  - Peter
    - Vektory vytvorené one-hot encodingom su prilis "sparse"
      - Trenciansky kraj a Trencianský kraj sú iné vekory
      - Na takej prázdnej vektorovej reprezentácii auta nieje možné trenovať model
      - Bolo navrhnuté použiť NLP na embedovanie. To by pravdepodobne zaradilo bratislavsky a trenciansky kraj do jednej vektorovej podoby
    - Rozhodli sme sa preto prioritne pracovat na NLP tasku s tym ze snad budu casti z toho pouzitelne aj pri ML

- Riso
  - Praca docstringoch
- Tomas
  - Zmeny na frontenede a API suvisiace s taskami alebo odsrtranovanim bugov su skoro hotove
- Zajtra stvrtok 15:00 call ku NLP (aj v suvislosti s ML, dat vediet Mirovi, keby sa chcel pripojit)
- Hrubý plan co pojde do dalsieho sprintu - Max to pridal, rozpravame co okrem NLP ML veci chceme robit
  - CAR 199-204