

19. Stretnutie 7.4.2020

Agenda

1. Rozpracovane storky
 - a. Daco k anotovaniu? Pozorovania akeho charakteru su fulltext data
 - b. Uprava parserov
 - c. NLP
 - d. ML
 - e. <https://team22-19.studenti.fiit.stuba.sk/jira/browse/CAR-217> - zrusit, neodstranovat diakritiku a velke pismena v db
2. Diskusia k finalnej dokumentacii, `_checkfor_reset` funkcionalita
3. Brainstorming
 - a. Dohladavanie technickych udajov aut ako storka? Vieme potom viac udajov poskytnut pouzivatelovi ale aj lepsie porovnavat auta (hladat podobne) atd <https://www.auto-data.net/en/> <https://www.cars-data.com/>
4. Backlog Grooming

Priebeh stretnutia (hangouts call)

- Diskusia k manualnemu anotovaniu dat
 - Zopar chyb pri vacsine odovzdaných dat - budu opravene pri NLP
- CAR-201: Drobne chyby
- NLP - veronika:
 - Atributy
 - Funkcia ktore pre marketplace ktora zjednoti diakritiku stringov
 - Fukcia normalizacna sa nemoze volat parseroch data su potom deformovane na frontend, pedspracovanie sa musi udiat pred ML
- NLP - ondrej
 - Endpoint je pripraveny
 - Priprava Bazos parsera pre ukladanie vystupu z textPredictora do DB
- ML - triplet loss
 - Velmi by pomohlo dokoncenie zjednotenia parserov
 - Diskusia o možnostiach rozsirenia slovnika
 - Peto posle zoznam slov ktore robia problem vo word2vec ze ich nepozna
- Frontend
 - Diskusia k vzhladu frontendu
- Zdokumentovanie procesu crawlovania a interakcia kontainerov ako cast tasku CAR-201
- CAR-203: rozpracovane
- Upravenie routovania na serveri zmenit URL pre manazera z /carad/ na /predicar/
- Diskusia k dalsiemu sprintu:

- Zobrazovanie statistik na FE
- Priprava a trenovanie modelov podobnosti
-