Sprint 4

- Sistem de recomandare pentru feed -

Sprint 11.12.2023 - 22.12.2023 (+ cateva zile din vacanta de iarna aprox 29.12.2023)

Overview of the team and their planned capacity:

- Robert, Stefan si Petre au lucrat la proiect

- Radu si Cosmin au fost ocupati cu proiecte la alte materii si nu au putut contribui

Sprint backlog (at the start of the sprint):

- (item) cautat cum se poate face un sistem de recomandare

- (acceptance criteria) sa fie un sistem de recomandare fezabil pentru proiectul nostru

- (item) cautat ce inseamna preferinte in cazul nostru

- (acceptance criteria) feature-uri relevante care pot fi folosite pentru a clasifica postarile

- (item) extragere de caracteristici din postari

- (acceptance criteria) un model de machine learning care sa poata extrage caracteristici din postari

- (item) clasificarea postarilor in functie de preferinte

- (acceptance criteria) sa apara pentru fiecare postare o categorie de preferinte

- (item) adaugat sistem de recomandare la feed

- (acceptance criteria) feed-ul sa arate cu prioritate postarile care sunt in concordanta cu preferintele utilizatorului

- (item) adaugat posibilitatea de a alege preferintele

- (acceptance criteria) un mod prin care utilizatorul sa isi poata alege preferintele

Status of planned items and their status:

- (item) cautat cum se poate face un sistem de recomandare

- (solved) am ajuns la concluzia ca langchain este cea mai buna solutie pentru proiectul nostru

- (item) cautat ce inseamna preferinte in cazul nostru

- (solved) am clasificat postarile in functie de o lista de hashtag-uri prestabilite

- (item) extragere de caracteristici din postari

- (solved) am folosit un language model pentru a extrage caracteristici din descrierile postarilor

- (item) clasificarea postarilor in functie de preferinte

- (solved) am folosit descrierea postarilor pentru a clasifica postarea in 3 hashtag-uri

- (item) adaugat sistem de recomandare la feed

- (solved) cu cat o postare are mai multe matched hashtags cu preferintele utilizatorului, cu atat va aparea mai sus in feed

- (item) adaugat posibilitatea de a alege preferintele

- (solved) utilizatorul isi poate alege hashtag-urile preferate

- (extra item) am creat un server de python care sa extraga hashtag-urile

Sprint review:

- Am reusit sa implementam un sistem de recomandare care sa afiseze postarile in functie de preferintele utilizatorului

- Am reusit sa implementam un sistem de extragere de caracteristici din postari

- Am reusit sa implementam un sistem de clasificare a postarilor in functie de preferinte

- Ar fi fost bine daca postarea ar fi fost clasificata si dupa poza, nu doar dupa descriere (dar nu era fezabil)

Sprint retrospective:

- Destul de ambiguu cum se foloseste langchain (pentru ChatGPT merge bine, dar pentru LLama nu are unele functionalitati)

- A trebuit sa gasim un trade-off intre cat de bun sa fie sistemul de recomandare si cat de mult sa dureze clasificarea postarilor + costul

- Avem nevoie de foarte multe date pentru a putea face un sistem de recomandare bun (un set de date cu poze + descriere si ca label hashtag-urile)

- A fost greu de gasit un prompt bun pentru langchain (pentru descrieri scurte merge prost)

- Initial voiam sa folosim un sistem de recomandare bazat pe colaborative filtering, dar platforma noastra nu are suficienti utilizatori pentru a putea face asta

- Daca nu foloseam langchain, puteam sa folosim un language model mai mic pe care sa il rulam local