

Kopija __Roadmap

Šiame dokumente aprašomi pagrindiniai etapai, būtini semantinės paieškos sistemos kūrimui. Sistemos tobulinimo metu gali prireikti ne kartą sugrįžti prie tam tikrų ankstesnių etapų, siekiant tobulinti paieškos veikimą.

1. Duomenų parengimas

Užduotys:

- Sintetinių duomenų generavimas: paslaugų aprašymų semantinei paieškai ir paieškų pavyzdžių kūrimas naudojant didelius kalbų modelius (LLM).
- Train/test duomenų skirstymas: paieškos duomenų atskyrimas į mokymosi (train) ir testavimo (test) rinkinius siekiant užtikrinti modelio tikslumą.

Definition of Done:

- Sugeneruotas paslaugų CSV failas, turintis šiuos papildomus laukus:
 - Sugeneruotas paslaugos aprašymas, skirtas semantinei paieškai.
 - Paslaugos paieškų frazių masyvas, paskirstytas į train/test dalis.

2. Testavimo procedūros parengimas

Užduotys:

- Apibrėžti vertinimo metrikas.
- Nustatyti ir paruošti Python funkciją, kuri įvertina paieškos efektyvumą.
- Apskaičiuoti baseline rezultatus su neadaptuotu modeliu.

Definition of Done:

- Python funkcija, kuri vertina paieškos modelio tikslumą.
- Pamatotas baseline, pateikiant vertinimo metrikų rezultatus.

3. Geriausio modelio parinkimas

Užduotys:

- Testuoti kosinuso panašumo metodą su daugiakalbiais modeliais, pritaikytais semantinei paieškai.
- Atrinkti geriausią modelį tolimesniam modelio tobulinimui.

Definition of Done:

- Sudarytas modelių sąrašas su jų vertinimo metrikomis pagal testavimo procedūrą.

4. Modelio adaptavimas - Finetuning

Užduotys:

- Paruošti ir adaptuoti modelį, atsižvelgiant į geriausius paieškos metodus.
- Įvertinti modelio adaptavimą naudojant iš anksto nustatytas metrikas.

Definition of Done:

- Sukurtas ir pritaikytas adaptuotas modelis.
- Apskaičiuoti ir užfiksuoti tikslumo metrikos rezultatai.

5. Infrastruktūros parengimas

Užduotys:

- Vektorinės duomenų bazės įdiegimas ir konfigūravimas.
- Sukurti REST API, skirtą paslaugų paieškai pagal vartotojų užklausas.
- Docker konteinerizavimas ir paruošimas diegimui.
- Diegimas į dev/test aplinką.

Definition of Done:

- Sukurtas konteinerizuotas sentence-transformers API, kuris grąžina TOP k paslaugų.

6. Paieškos tobulinimas - papildomos savybės

Užduotys:

- Sukurti mechanizmą, kuris leidžia susieti paieškos frazes su konkrečiomis paslaugomis.
- Įdiegti svorio nustatymo mechanizmą, kad administratorius galėtų įtakoti paieškos rezultatus.

Definition of Done:

- Administratorius gali susieti paieškos frazę su konkrečia paslauga, užtikrinant tikslesnius paieškos rezultatus.
- Administratorius gali reguliuoti paslaugų svorius, keisdamas paieškos rezultatų rūšiavimą pagal svarbą.