

Praksis hos Kartverket - RAG-søk for produktspesifikasjoner

Praksis hos Kartverket - RAG-søk for produktspesifikasjoner

Hos Kartverket har vi i løpet av høsten 2024 sett på RAG-modeller for produktspesifikasjoner og standarder hos GeoNorge. Oppgaven bygget videre på et tidligere bachelorprosjekt og handlet om å utvikle en RAG-modell (Retrieval Augmented Generation) for å muliggjøre en chatbot. Chatboten ville være nyttig for ansatte og eksterne brukere for å skaffe oversiktlig informasjon angående produktspesifikasjoner av kart og geodata. Oppgaven la opp til utforskning og testing av ulike teknologier, i samarbeid med KartAi-prosjektet.

I tiden hos Kartverket har vi holdt på med mulighetsstudier for å finne ut hva slags teknologi-muligheter som kan være aktuelle for å kunne utvikle RAG-søk for produktspesifikasjoner. Vi har sjekket ut forskjellige typer språkmodeller (LLMs), vektordatabaser og embeddingmodeller, som alle er avgjørende elementer i hvordan en RAG-modell fungerer. Vi har også fått prøve å utvikle i et mer profesjonelt miljø på eksterne servere og har møtt på flere utfordringer – som f.eks. å bruke jupyter notebook i kombinasjon med Miniconda for å kjøre/teste koden. Dette krevde og krever fortsatt mye prøving og feiling, men vi kommer stadig nærmere målet. Kartverket har også gitt oss jevnlig faglig påfyll og satt oss i kontakt med eksterne samarbeidspartnere for å kunne gi oss best mulig utgangspunkt for å jobbe med oppgaven.

Vi har lært hvordan vi setter opp en RAG-modell med en vektordatabase, chatbot og brukergrensesnitt, samt hvordan embedding brukes for å vektorisere PDF-filer som data. Vi har fått innsikt i ulike relevante teknologier som KI og LangChain, og evaluert vår RAG-modell ved bruk av RAGAS. Vi har også undersøkt hvordan man kan sette opp forskjellige chatbotter, som OpenAI og Llama. Erfaring med arbeid i programmeringsmiljø og kode-dokumentasjon har også vært sentrale læringspunkter.

Praksisoppholdet har vært både utrolig gøy og lærerikt. Vi har blitt fulgt godt opp av ansatte, og praksisperioden har gitt oss en unik erfaring innen teknologi som det er vanskelig å finne andre steder.

Fullført oppgave

Leverings- 2024 - Høst
tidspunkt:
Samarbeid: Statens Kartverk

Fagområder

- Datateknologi

Fakultet

- Fakultet for samfunnsvitenskap

Emnekoder

- IS-302, praksisprosjekt i Infosys, bachelor

Deltakere

- Andrine Celine Flatby
- Christine Nilsen
- Martin Steiro
- Tobias Molland

Del på sosiale medier:   