Problem Set 9

Femi Adeniran

May 5, 2023

Abstract

1 Project Breakdown

1.1 Purpose

A way to query numerical databases with natural language powered by LLMs.

1.2 Problem

Dynamic Information: Search engines and AI like ChatGPT may struggle to provide accurate answers for real-time, constantly changing data. Search Engine Limitations: Search engines rely on published content, which may not always be up-to-date or contain real-time information. ChatGPT Limitations: ChatGPT has a knowledge cutoff date and cannot access real-time data, limiting its ability to provide current information.

1.3 Method and Tools

We employ the GPT 3.5 Turbo API as the base for natural language processing . Work is done in understanding properly it's one shot capabilites A detailed explanation of the database is given followed by a few one shot examples that show what should be done and how.

And I began prompting and refining the prompt to make sure the LLM had decent coverage over the database and also good results. The logic behind the whole project was to create a translation bridge between natural language query and a programming language that can query an actual database.

I used a language I was familiar with to test it, and did the code execution in python and Dataframe in Pandas

1.4 Results

After a long period of brainstorming I finally stumbled on the right piece of information to solve the project. The results have been impressive and querying databases never felt easier.

It understands queries that dont have columns for them, it creates connections like nicknames, and club slogans etc

It has a big future use in all kinds of databases and going forward I dont htink there will be a different way of querying database