IDEA

מיישמים טכנולוגיות AI בעסקים

Take-Home Task:
Natural language query - SQL DB

Overview: Why Text-to-SQL (or NL-to-SQL)?

Organizations store valuable business data (orders, customers, inventory, financials, etc) in various types of **SQL databases**. However, querying this data typically requires writing SQL queries, which creates a dependency on developers, analysts, or IT staff for even the simplest reports.

Our goal is to empower **non-technical team members** (sales, finance, support, operations) to **ask data questions in natural language** — like:

"Show me the top 10 customers by revenue in the last 6 months"
"What is the current inventory level of item 24567?"
"How many support tickets were opened last month?"

By building a **Text-to-SQL assistant**, we bridge the gap between business intent and technical execution:

- Save time for the data team
- Enable self-serve insights

You will be building a **realistic mini-prototype** of such a system, simulating how a conversational interface can generate SQL queries over a SQL database.

Goal

Build a working prototype of a **text-to-SQL assistant / chatbot** that allows business users to ask natural language questions (e.g. *"Show me the top 10 customers by revenue last year"*) and get correct SQL queries (or query results) over an **SQL Server database**.

sales@idea2mtl.com   www.idea2mtl.com

What You'll Deliver

- A self-contained web app (or API) that:
    - Accepts natural language questions.
    - Returns the translated SQL query, executes it and returns results
    - Supports a demo SQL database schema.
- A clear **readme** explaining your architecture, assumptions, and trade-offs.
- Bonus: UI for non-technical users to test it.

Technical Guidelines

- Use LLM of your choice.
- SQL Database to be used: AdventureWorks Sample Databases - Select the OLTP version of 2022 - AdventureWorks2022.bak.
- Optional: RAG pipeline or vector search for relevant tables schema.
- Note that you will also need to install an SQL Server - you can find an option to download here or you can use MySQL or other.

Suggested Features

- Convert natural language questions to SQL (SELECT, joins, filters, etc.).
- Show the generated SQL to the user.
- Optionally: Execute the SQL (against a real or mock DB).
- Bonus: Feedback loop ("Was this SQL correct?") to simulate real-world learning.

Example of natural language question which should be answered (using the AdventureWorks Sample Database)

- Top 5 Customers by Total Purchase Amount
- Products with No Sales in the Last 6 Months
- List Employees Hired in the Last 5 Years
- Identify Products with No Sales in the Past Year
- Average Order Quantity by Product Category
- Employees with Above-Average Salaries

Time Expectation

Home task should be submitted no later than May 4th (EOD)
Please don't over-engineer! Focus on solving the core problem well, with attention to clarity, modularity, and realistic scope.
Please do not hesitate to contact us for any clarifications required!

Submission

- GitHub link (public or private invite)
- Indicate if vibe coding / no code tools (like Curson, Windsurf, Lovable, etc) were used and in which parts

What We'll Be Evaluating

Category	What We're Looking For
GenAI & ML Engineering	Effective prompt engineering, use of LLMs (OpenAI preferred), schema ingestion, optional RAG
Backend Architecture & Ownership	Thoughtful data pipeline design, use of tools like LangChain or MCP, extensibility
Fullstack & Integration	Working web app or well-documented API, nice if user-friendly (Streamlit, Flask UI, etc.)
Mindset & Independence	Initiative, clarity of documentation, prioritization of effort vs. value

sales@idea2mtl.com   www.idea2mtl.com