Database Systems
CMPT 308

NL2SQL Project - 100 points

Goal

To learn about the exciting field of Natural Language (NL) interfaces to SQL-based systems (NL2SQL). We've spent the first half of our semester querying relational databases with SQL, but not everybody has the pleasure of taking this class. For those unfortunate people, and countless others who do not know SQL yet seek to turn data into information, a natural language interface may be the answer.

Instructions

Research the current state of NL2SQL technologies and create a summary presentation. Here are some papers and web sites you might find useful:
- Natural Language to SQL: Where are we today?
- TypeSQL: Knowledge-based Type-Aware Neural Text-to-SQL Generation
- Relation-aware Schema Encoding for NL2SQL
- Relation-aware Parsing for NL2SQL
- WikiSQL

Deliverables

Create a slide presentation as if you were going to teach this to our class. (Don't worry, you're not going to, as I crave the attention myself.) But put yourself in that mindset as you develop your slides. You never know truly something at a deep level until you teach it to someone else.)

- Slide 0 - Title, your name, the date, a cool graphic or logo or image

- Slide 1 - What does an NL2SQL system look like? Show a small example.

- Slide 2 - Early History

- Slide 3 - Recent History, including LLMs.

- Slide 4 - Current Systems and Capabilities - what they do

- Slide 5 - Current Systems and Capabilities - how they do it

- Slide 6 - What's next (features coming soon)

- Slide 7 - Research Opportunities

- Slide 8 - How would **you** approach this problem?

- Slide 9 - References

Resources

- The references in this assignment and on our web site.
- Your own research

Submitting

Submit your work as a PDF. (Only PDF files will be accepted.) Push it to your GitHub repository **before** the date at which it is due (see syllabus). Remember, neatness counts. So does accuracy, spelling, grammar, correctness, and essentially everything else. It only has to be perfect. Make it so.