

TEXT TO SQL

Hiya Jain - 60009210182

Aryann Tated - 60009210194

Yash Thakar - 60009210205

Introduction

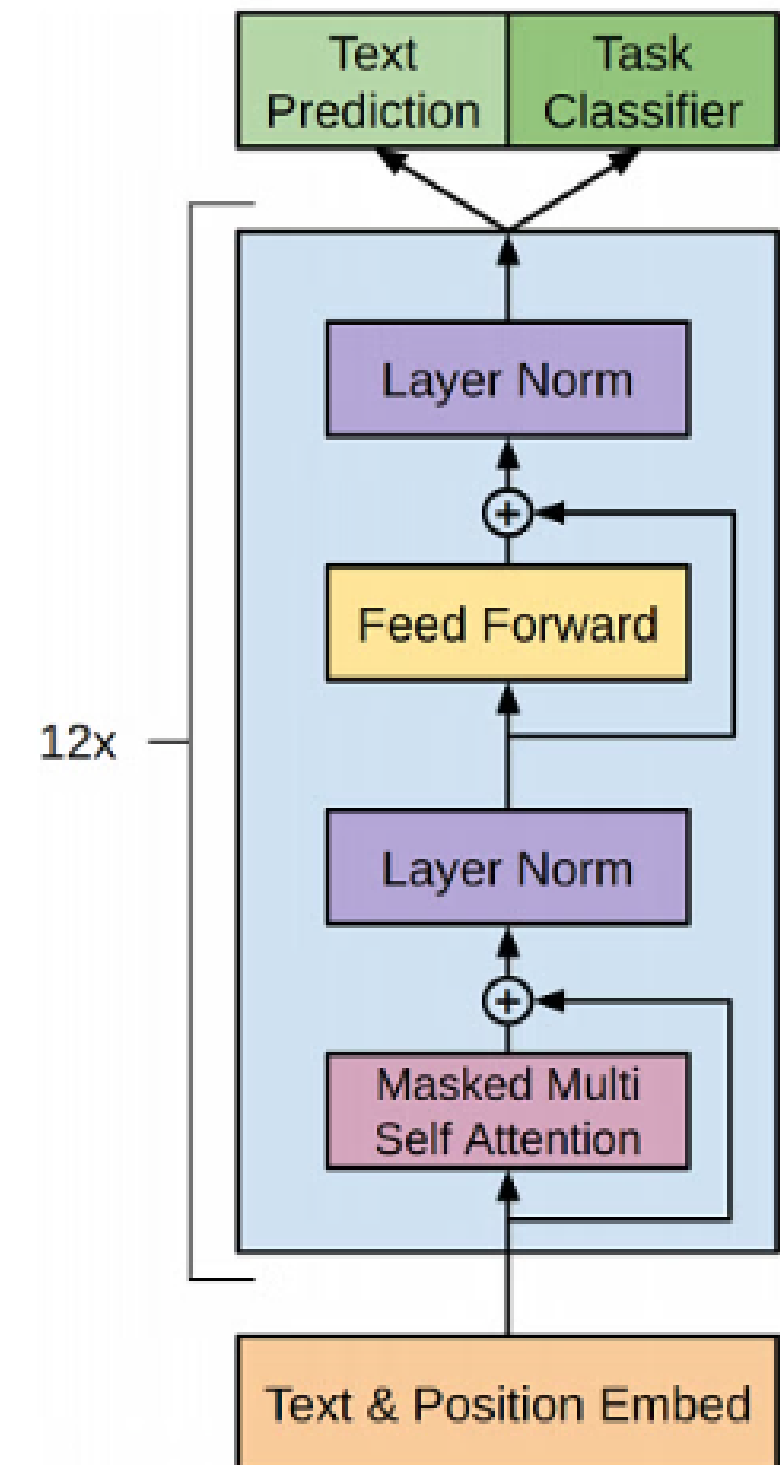
Text-to-SQL converter bridges the gap between non-technical users and databases, enabling anyone to query data without SQL expertise.

Dataset and Model used

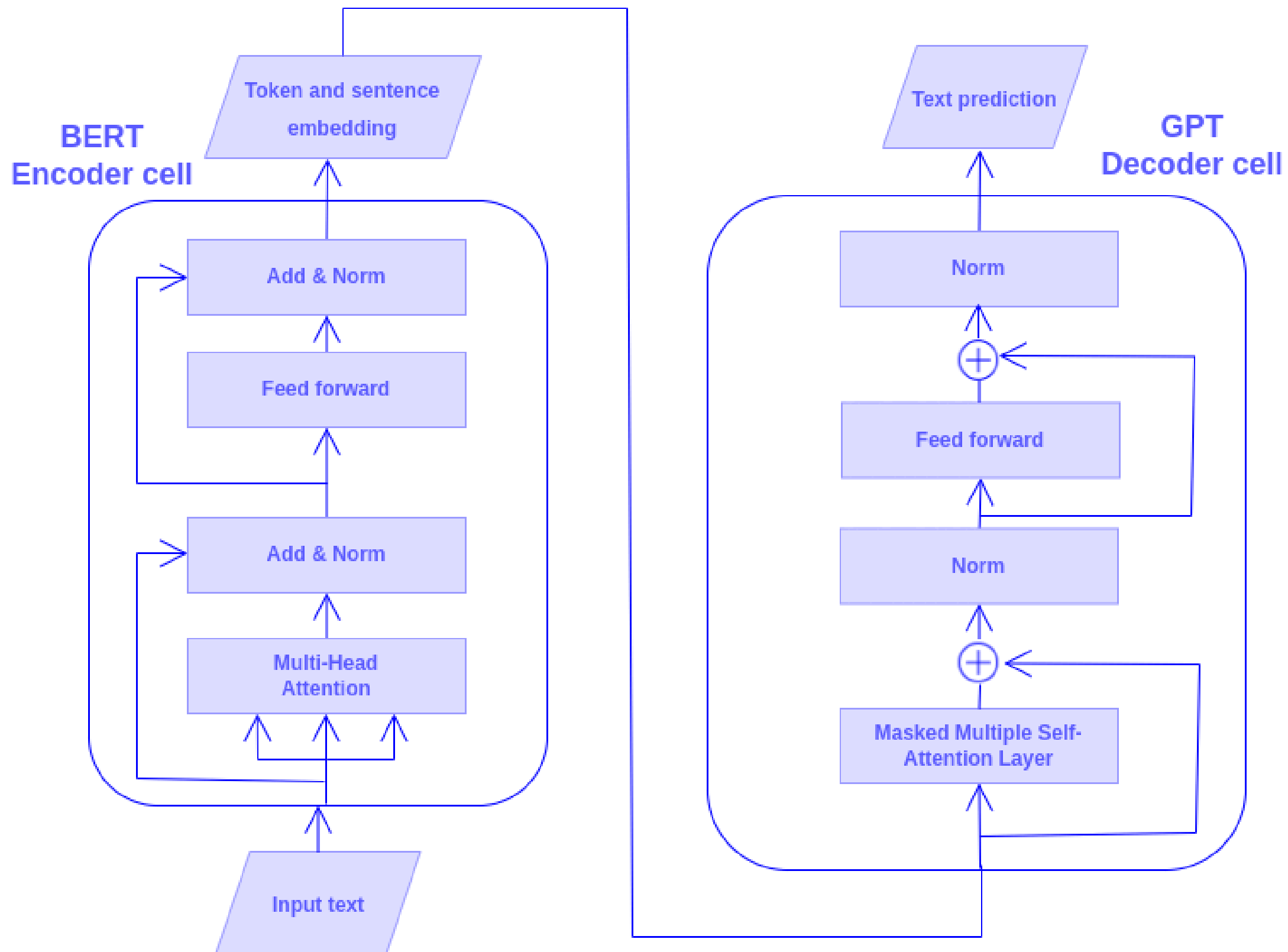
- Built upon WikiSQL and Spider datasets, with 78,577 examples of natural language queries paired with their SQL queries.
- SQLGlot parses queries to infer column types from operators (>, <) and functions (MIN, MAX, AVG, SUM), defaulting to VARCHAR if uncertain.
- The table contains 3 columns - Questions, Answer and Context.

BART

- Combines bidirectional encoder (like BERT) with autoregressive decoder (like GPT)
- BART employs corruption methods - token masking, deletion, text infilling, sentence permutation, and document rotation, training the model to reconstruct the original text by learning both content and structural patterns.



Architecture



Performance Evaluation



Fine Tuned Model Deployed on Hugging Face - [link](#)

Compared performance evaluation using BLEU score.

$$\text{BLEU} = BP \cdot \exp \left(\sum_{n=1}^N w_n \log P_n \right)$$

$$BP = \begin{cases} 1 & \text{if } c > r \\ e^{(1-r/c)} & \text{if } c \leq r \end{cases}$$