NARASARAOPETA ENGINEERING COLLEGE (AUTONOMOUS)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

2025-2026

Batch Number	BB-6
Team Members	K . Mani Kanta(22471A05A6) G. Srinivas (22471A0594) G. Balaji (22471A0593)
Guide	V. Karuna Kumar M.Tech
Title	**Semantic-Augmented Prompt-Guided Sketch Filling for Text-to-SQL Generation**
Domain/Technology	DEEP LEARNING
Base Paper Link	**https://ieeexplore.ieee.org/document/10711192**
Dataset Link	**https://www.kaggle.com/datasets/shahrukhkhan/wikisql**
Software Requirements	Browser: Any latest browser like Chrome Operating System: Windows 7 Server or later Python (COLAB)
Hardware Requirements	SystemType: Intel Core i5 or above RAM: 8 GB Number of cores:5 Number of Threads: 4
Abstract	This paper proposes a deep learning model for the Text-to-SQL task using a **sketch filling approach** built on top of a **BERT-Large encoder**. The model addresses two common challenges: (1) **insufficient information** in user queries, where column names or values may be missing, and (2) **semantic gaps** between the natural language query and the database content. To handle these, the model first preprocesses queries using **named entity tagging** and applies **data augmentation** by randomly swapping entities, which strengthens the learning of query templates. During encoding, the model introduces relevant **table content** and uses a **table-aware attention mechanism** to improve semantic understanding. Experimental results on benchmark datasets (like WikiSQL) demonstrate that the proposed method, along with ablation studies, confirms the benefits of both content enhancement and data augmentation in improving prediction accuracy.

Signature of the student(s) **Signature of the Guide** **Signature of the project coordinator**