ISSN 2616-5775 Vol. 8, Issue 5: 38-45, DOI: 10.25236/AJCIS.2025.080504

Research on Multi-Turn Chinese NL2SQL Methods Based on Semantic Rewriting

Shuaike Guo*, Yongzheng Yang

School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin, Guangxi, China
guoxiaoke_hd@163.com

Abstract: *In the era of data intelligence, Natural Language to SQL (NL2SQL) technology serves as a core interface for human-machine data interaction, making its performance optimization in multi-turn Chinese dialogue scenarios highly valuable for research. In addressing the issue where reference resolution and semantic omission in the Chinese context can lead to gaps in understanding user intent, this paper proposes a model, RW-T5, integrated with a semantic rewriting mechanism. This model is based on the pre-trained T5 architecture and utilizes hierarchical modeling of dialogue history along with turn-aware encoding to accurately parse semantic unit segmentation and temporal dependencies in multi-turn interactions. It features an innovative design for a global context injection and bidirectional cross-attention fusion module, enabling the capture of both the overall semantic focus and fine-grained word-level semantic details. Utilizing a sequence optimization strategy based on multi-dimensional semantic feature fusion, the model effectively performs explicit resolution of implicit reference relationships and logical completion of omitted semantics in multi-turn dialogues, providing a semantically complete and structurally standardized input for subsequent SQL statement generation. Experimental validation on the large-scale Chinese multi-turn dialogue benchmark dataset, CHASE, shows that this model significantly outperforms other advanced NL2SQL parsing methods, fully validating the effectiveness of the dynamic semantic rewriting mechanism and hierarchical modeling approach, and offering an effective solution for the engineering implementation of intelligent data interaction systems in Chinese multi-turn dialogue scenarios.*

Keywords: *Natural Language Processing; Multi-Turn Dialogue Understanding; Semantic Rewriting; NL2SQL*

1. Introduction

In the digital age, data has become an important basis for decision-making in enterprises and organizations. As the amount of data continues to grow and data structures become increasingly complex, efficiently extracting valuable information from massive data sets has become a key issue. Traditional database querying methods rely on Structured Query Language (SQL); however, the complexity of SQL syntax makes it difficult for non-technical users to use it directly for data queries. Natural Language to SQL (NL2SQL) technology has emerged to address this challenge, aiming to convert users' natural language queries into corresponding SQL statements, thereby enabling natural interaction between users and databases, lowering the barriers to data querying, and improving data utilization efficiency.[1]

NL2SQL technology in multi-turn dialogue scenarios further expands the flexibility of user interactions with databases. In complex data query requirements, users often find it challenging to express all their intents clearly in a single turn. Multi-turn dialogue allows users to gradually refine their query requirements; the system generates accurate SQL queries based on each user input and the previous dialogue history.[2] For example, in the context of analyzing enterprise sales data, a user may first ask, "What was the total sales amount last month?" and then follow up with, "What is the sales proportion for each region?" Through multi-turn dialogue, the system can better understand the user's complex needs and generate the corresponding SQL queries to retrieve accurate data.

However, NL2SQL technology in the Chinese context faces numerous challenges. The semantic ambiguity in Chinese means that the same word or phrase can have different meanings in different contexts; for example, “apple” can refer to either the fruit or Apple Inc. This complicates semantic understanding and SQL generation. The issue of reference resolution is also prominent in Chinese; in

ISSN 2616-5775 Vol. 8, Issue 5: 38-45, DOI: 10.25236/AJCIS.2025.080504

multi-turn dialogues, users may use pronouns to refer to entities previously mentioned, requiring the system to accurately identify these reference relationships to generate correct SQL queries. The context dependency in multi-turn dialogues requires the system to effectively integrate historical dialogue information to understand the user's coherent intent, yet existing methods often suffer from information loss or insufficient integration when handling contextual information.

Semantic rewriting technology offers a new approach to address these issues. Semantic rewriting dynamically modifies and optimizes the semantic representation of user queries, allowing for better capture of the user's true intent and disambiguation of semantics. In multi-turn dialogues, semantic rewriting can adjust the semantic representation of the current turn based on contextual information, ensuring that the generated SQL aligns more closely with the user's overall needs. Therefore, researching multi-turn Chinese NL2SQL methods based on semantic rewriting has significant theoretical and practical importance, with the potential to enhance the accuracy and robustness of SQL generation in multi-turn interactions and advance the development of natural language and database interaction technologies.

To address the aforementioned issues, we propose the RW-T5 model, which is based on semantic rewriting for multi-turn Chinese NL2SQL generation. By dynamically calculating the semantic weights of each dialogue turn, the model achieves hierarchical aggregation of key information from historical dialogue, emphasizing the core context. It further captures semantic associations between turns through cross-turn semantic alignment, addressing the unique challenges of word order flexibility and polysemy in Chinese. By combining Conditional Random Fields (CRF) for sequence labeling optimization, the model explicitly completes ambiguous references and omitted expressions into specific semantics, providing precise semantic input for subsequent SQL generation and significantly improving the semantic completeness and intent resolution accuracy of multi-turn dialogues.

2. Model Design

2.1 Model Overview

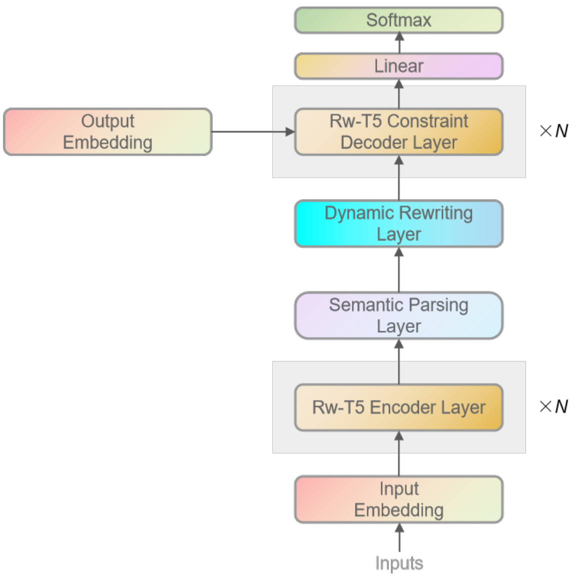

Figure 1 Model Architecture

The architecture of the multi-turn Chinese NL2SQL method based on semantic rewriting proposed in this paper is shown in Figure 1. It is based on the mT5 [3] model and is divided into four stages: "Dialogue Encoding – Semantic Parsing – Dynamic Rewriting – Constraint Decoding." The aim is to effectively handle natural language queries in multi-turn dialogues and accurately convert them into SQL statements.

First, a pre-trained language model is used to encode the input data, such as natural language questions, serializing different data items with delimiters. Secondly, a semantic parsing module is introduced into the model to analyze the encoded semantic vectors using syntactic analysis techniques. This module dynamically evaluates the semantic importance of each turn in the dialogue by calculating

ISSN 2616-5775 Vol. 8, Issue 5: 38-45, DOI: 10.25236/AJCIS.2025.080504

turn weights to highlight key contextual information, achieving hierarchical aggregation of the core semantics of multi-turn dialogues and constructing a cross-turn semantic interaction matrix to capture the semantic associations between historical dialogues and the current statement, further enhancing the model's ability to understand cross-turn dependencies. Next, a dynamic rewriting module is introduced, which maps contextual semantic features and the features of the current statement into a unified semantic space. Through multi-dimensional similarity measurements, the module accurately assesses the semantic correlation between the two, providing a quantitative basis for semantic completion. Finally, decoding is performed in conjunction with relevant decoding strategies to generate SQL statements in the specified format. The following sections will provide a detailed introduction to each module.

2.2 Dialogue Coding

The context encoding module employs the encoder strategy of the mT5 model to fully capture the complex information in multi-turn dialogues. In practical applications, the historical dialogue sequence contains the query intentions gradually expressed by the user during multi-turn interactions, as well as the SQL query statements obtained in previous turns. The current query represents the specific question posed by the user in this turn, while the database schema (including table names, column names, data types, etc.) provides structural information for the query. These pieces of information are input into a hierarchical encoder, where the historical dialogue sequence is first encoded. Using a self-attention mechanism, the model can automatically learn the relationships between different dialogue turns, capturing topic shifts and continuities. For the current query, it is transformed into a semantic vector so that it resides in the same semantic space as the encoded vectors of the historical dialogues. The information from the database schema is encoded into vector representations, which are then fused with the encoded vectors of the historical dialogues and the current query. Ultimately, we obtain the encoded vectors corresponding to these inputs.

2.3 Semantic Parsing

After encoding, we obtain the encoded vectors of the historical dialogue sequences and the encoded vector of the current incomplete statement, where $h_{\mathrm{i}} = \{h_1, h_2, h_3, \cdots, h_{n-1}\}$ represents the dialogue turn, h_n Li is the turn length, and i is the hidden layer dimension. We then perform average pooling on each historical turn to obtain the historical turn vectors $h_i^{avg} \in R^d$, which are then weighted and fused using an attention mechanism:

$$\alpha_i = soft\max(q^T W_a h_i^{avg}),\quad i < n \tag{1}$$

$$c_{gloabal} = \sum_{i=1}^{n-1} \alpha_i h_i^{avg} \tag{2}$$

The query vector $q = h_n^{avg}$ a is the average pooling representation of the current round of utterances, and $W_a \in R^{d\times d}$ represents the learnable parameters. Subsequently, the global context is injected by concatenating the global context vector $c_{gloabal}$ to each word vector in the feature vector h_n of the current utterance:

$$X_{enhanced} = [x_n; c_{global} \cdot 1_N] \in R^{N\times 2d} \tag{3}$$

In the next step, the semantic segments of the context and the current utterance are explicitly aligned. A fine-grained interaction model is established between the context and the current incomplete statement by introducing a bidirectional attention mechanism. The word-level representation of the context $C = [H_1, H_2, \cdots, H_{n-1}] \in R^{M\times d}$ is used to calculate the attention from the context to the current utterance $M = \sum_{i=0}^{n-1} L_i$, retrieving relevant semantics from the context through cross-attention:

ISSN 2616-5775 Vol. 8, Issue 5: 38-45, DOI: 10.25236/AJCIS.2025.080504

$$X_{ctx2cur} = Attention(X_{enhanced}W_Q, CW_K, CW_V) \tag{4}$$

Calculate the impact of the current utterance on the context in reverse and filter key segments:

$$X_{cur2ctx} = Attention(CW_Q^{'}, X_{enhanced}W_K^{'}, X_{enhanced}W_V^{'}) \tag{5}$$

Subsequently, the results of the bidirectional attention are fused using dynamic gating signals, and a residual connection with the original representation is established to obtain the feature representation of the current incomplete statement:

$$g = \sigma(W_g[X_{ctx2cur}; C_{cur2ctx}]) \tag{6}$$

$$X_{final} = H_n + g \odot X_{ctx2cur} + (1-g) \odot C_{cur2ctx} \tag{7}$$

Among them, $X_{final} \in R^{N \times d}$

To further enhance the measurement capability of the semantic association between the context and the current statement, a similarity calculation module based on a twin network is introduced. This module inputs the current statement features X_{final}, after being fused through cross-attention, and the global context vector $c_{gloabal}$ into the two branches of the twin network [4], forming a contrastive learning input pair:

$$f_{context,i} = Siamese(c_{gloabal,i}) i \in [1, M] \tag{8}$$

$$f_{current,j} = Siamese(X_{final,j}) j \in [1, N] \tag{9}$$

The twin network includes convolutional layers and fully connected layers, learning the similarity measure of semantic features through a parameter-sharing mechanism. The similarity calculation employs a multidimensional fusion strategy, integrating cosine similarity, dot product similarity, and the nonlinear similarity computed by the multilayer perceptron (MLP) to ultimately obtain the similarity matrix $F_{i,j}$:

$$F_{i,j} = \gamma_1 \cdot \cos(f_{\text{context,i}}, f_{\text{current,j}}) + \gamma_2 \cdot \text{dot}(f_{\text{context,i}}, f_{\text{current,j}}) + \gamma_3 \cdot \text{MLP}(f_{\text{context,i}}, f_{\text{current,j}}) \tag{10}$$

Among them, γ_1, γ_1, and γ_3 are learnable weight coefficients used to dynamically adjust the contributions of different similarity calculation methods. Cosine similarity measures the directional consistency of feature vectors, while dot product similarity reflects the magnitude correlation of feature vectors. The MLP captures nonlinear semantic relationships. Through multidimensional fusion, this module can measure the semantic correlations between context and the current statement from different perspectives, providing a more comprehensive decision-making basis for subsequent semantic area localization and completion.[5]

2.4 Dynamic Rewriting

In the semantic rewriting task, the core function of dynamic rewriting is to locate and identify the areas in the input text that require semantic modifications. This is achieved through an improved U-Net [6] architecture combined with Conditional Random Field (CRF) sequence optimization, facilitating the precise localization of semantic modification areas and the efficient prediction of editing types.

The original U-Net employs an encoder-decoder structure with skip connections to merge multi-scale features, but traditional downsampling operations can lead to a loss of spatial resolution, negatively impacting the capture of detail information. To address this, an improved U-Net architecture is adopted, incorporating dilated convolutional blocks and dense connection blocks to enhance the accuracy and efficiency of semantic area segmentation. Additionally, the introduction of dense connection blocks enables inter-layer feature concatenation, enhancing feature reuse capabilities. During the decoding phase of the improved U-Net, bilinear interpolation upsampling and convolution operations are used to progressively restore the resolution of the feature maps. Finally, a feedforward neural network maps each feature vector to a probability distribution over three editing types (insertion,

ISSN 2616-5775 Vol. 8, Issue 5: 38-45, DOI: 10.25236/AJCIS.2025.080504

replacement, retention), generating a probability matrix $\mathbf{Y}$.

To further leverage the sequential dependencies between editing types (such as specific entity vocabulary being more likely to follow an "insertion" operation), a Conditional Random Field (CRF) layer is introduced for sequence constraint optimization. The goal of the CRF layer is to solve for the label distribution $\mathbf{Y}^*$ that maximizes probability given the input matrix $\mathbf{X}$. Its energy function is defined as follows:

$$\mathbf{Y}^* = \arg\min_Y E(\mathbf{Y}, \mathbf{X}) \tag{11}$$

$$E(\mathbf{Y}, \mathbf{X}) = \sum_{i=1} \psi_u(y_i) + \lambda \sum_{i=1} \sum_{j>i} \psi_p(y_i, y_j) \tag{12}$$

The unary potential function $\psi_u(y_i)$ is computed using Equation 11, which measures the compatibility between a single position's label y_i and the observed data. This is done by reinforcing the selection of high-confidence labels through negative log probabilities. In Equation 12, the pairwise potential function $\psi_p(y_i, y_j)$ uses a Gaussian kernel function to measure the consistency between adjacent labels:

$$\psi_u(y_i) = -\log P(y_i) \tag{13}$$

$$\psi_p(y_i, y_j) = \begin{cases} 0 & y_i = y_j \\ \exp\left(-\dfrac{\| \mathbf{p}_i - \mathbf{p}_j \|^2}{2\theta^2}\right) & y_i \neq y_j \end{cases} \tag{14}$$

Here, p_i and p_j represent the character position coordinates, and $\theta = 1$ is the width parameter of the kernel function. This function enhances the local smoothness of segmentation results by penalizing inconsistent labels at non-continuous positions.

Based on the predicted editing type results from the segmentation layer, the semantic rewriting process is divided into two key steps: first, using a connected region segmentation algorithm (such as 4-neighborhood connectivity detection) to identify continuous editing areas, generating a minimal bounding rectangle to locate the text segments that need modification; second, generating the rewritten statement $h_n^{'}$ based on the editing type labels (e.g., "insertion" corresponds to supplementing referring entities, while "replacement" corresponds to disambiguation correction), in combination with the contextual semantics.

2.5 Constraint Decoding

The SQL statement generation decoding is based on the core modules of the mT5 model, such as masked self-attention, cross-attention, and feedforward networks. By introducing constraint logic and a dynamic syntax monitoring mechanism, a hierarchical constraint decoding framework is constructed to ensure that the generated SQL statements possess both syntactic correctness and consistency with the database schema. The specific mechanisms include:

In the masked self-attention module, a clause type bias matrix is overlaid, dynamically adjusting the attention weight distribution based on the semantic type of the currently generated clause (e.g., SELECT, FROM, WHERE) [7], guiding the model to focus on the semantic space of the current clause. The cross-attention module introduces a schema-enhanced bias generated from database schema embeddings, strengthening the semantic association with the corresponding nodes in the database schema when generating table names or column names, thus reducing field reference errors.A finite state machine is integrated to construct a state transition table that predefines a valid order for clause generation and updates the syntax state in real-time during decoding, dynamically filtering out illegal token generations. Building on the traditional cross-entropy loss, additional losses for syntax violations and schema alignment are introduced, forming a multi-objective optimization system. This dual constraint approach from both the generation process and loss function ensures the legality and accuracy of SQL statements.

ISSN 2616-5775 Vol. 8, Issue 5: 38-45, DOI: 10.25236/AJCIS.2025.080504

3. Experiments and Analysis

3.1 Experimental Setup

3.1.1 Dataset

CHASE [8] is a large-scale dataset designed for the Chinese multi-turn dialogue NL2SQL task, consisting of 5,459 question sequences, 280 databases, and 17,940 annotated data entries. It focuses on context-dependent complex queries, with the proportion of context-dependent questions reaching 64.7%, significantly surpassing the English datasets SParC (52.5%) and CoSQL (31.8%). Additionally, the SQL complexity distribution is more balanced, with simple queries constituting 28% of the dataset.The dataset is divided into CHASE-C (a manually constructed high-difficulty subset, with "difficult" and "very difficult" SQL queries making up over 44%) and CHASE-T (based on improved translations from SParC). The annotation system covers multi-dimensional information such as schema linking methods and types of context dependency. The database schema is stored in JSON format, including table structures and primary-foreign key relationships.

3.1.2 Evaluation Metrics

To scientifically evaluate the model's performance in the multi-turn Chinese NL2SQL task, this study employs Question Matching (QM) and Interaction Matching (IM) as evaluation metrics. QM is based on exact matches of SQL clause sets and measures the accuracy of single question parsing, counting as correct only when all clauses (SELECT, FROM, WHERE, etc.) are fully consistent with the standard SQL. IM assesses the semantic coherence of multi-turn dialogues, requiring that all questions in the dialogue satisfy QM=1 to be considered a successful match, effectively testing the model's ability to inherit semantics across turns.

3.2 Comparison of Experimental Results

The CHASE dataset, as the first large-scale context-aware natural language to SQL dataset aimed at the Chinese context in recent years, provides an important benchmark for research in semantic parsing in multi-turn dialogue scenarios. This study selected EditSQL [9], IGSQL [10], and RatSQL+Concat [11] as comparison models for the experiments. The above models were initially designed primarily for English datasets. This study adapted them to accommodate the Chinese language characteristics and multi-turn dialogue annotation standards of the CHASE dataset by improving Chinese tokenization strategies, adjusting cross-modal semantic alignment mechanisms, and optimizing database schema encoding.

The results of the comparison experiments are shown in Table 1. A thorough analysis of the experimental results reveals that the proposed RW-T5 model outperforms all other models across various metrics. The EditSQL model focuses on capturing semantic correlations between the current turn and historical turns, while IGSQL emphasizes encoding historical database schema items. In contrast, the RW-T5 model preprocesses the input data through semantic rewriting, allowing the model to better perceive semantic changes across turns.

Table 1 Comparison of Experimental Results

Model	CHASE				CHASE-C				CHASE-T	
	Dev		Test		Dev		Test		Dev	
	QM	IM	QM	IM	QM	IM	QM	IM	QM	IM
EditSQL	37.7	17.4	37.8	14.7	33.6	8.4	32.6	8.7	41.6	21.6
IGSQL	41.4	20.0	40.4	15.6	31.4	10.8	32.6	9.3	43.3	26.3
Rat-Con	35.1	14.6	32.5	9.8	24.6	5.4	23.9	4.5	43.7	21.6
RW-T5	55.6	30.2	50.8	28.4	42.3	25.3	40.8	23.2	54.2	39.1

In terms of the Interaction Matching (IM) metric, the RW-T5 model significantly outperforms the comparison models on the dataset. This is attributed to the RW-T5 model's ability to ensure semantic consistency by aligning natural language questions with database schema items, whereas EditSQL and IGSQL only encode natural language and database schemas separately. The RW-T5 model represents the relationships between question words and schema items and inputs them into the model for unified interactive encoding, thereby effectively enhancing the accuracy of semantic processing in multi-turn dialogues.

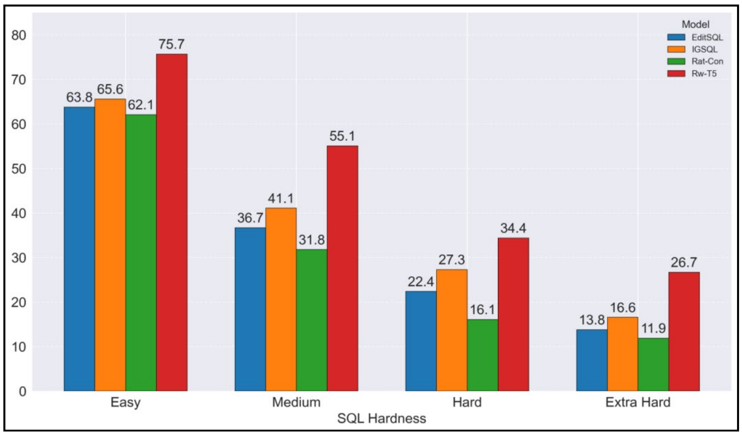

Figure 2 QM Values of the Model at Different Difficulty Levels

To further analyze the effectiveness of the proposed model in handling contextual relevance, the QM values of different difficulty SQL statements for each model were compared on the validation set of the CHASE dataset. The SQL statements were classified into four levels based on the complexity of keywords and SQL structures: simple, medium, difficult, and very difficult. The experimental results are shown in Figure 2, indicating that the RW-T5 model outperforms other models across all difficulty levels of SQL statements. As the complexity of the SQL statements increases, the difficulty of the model's predictions correspondingly rises. Currently, for very difficult SQL statements, the generation performance of the model still needs improvement, which remains a significant challenge in this field.

Table 2 QM Values Across Different Turns

Model	1	2	3	4	>=5
EditSQL	55.8	38.1	25.7	19.2	16.1
IGSQL	59.1	42.3	29.4	24.5	20.5
Rat-Con	51.4	37.5	24.8	18.4	14.5
RW-T5	70.5	59.4	51.8	45.7	40.3

A comparison of the average QM values for each sequence position was also conducted, and the experimental results are shown in Table 2. As the number of turns increases, the amount of historical information to be considered grows, leading to more complex semantic changes and greater prediction difficulty. This fully demonstrates that the semantic processing method proposed in this paper for multi-turn NL2SQL problems can generate SQL statements more effectively when there are contextual dependencies between turns.

3.3 Ablation Experiments

Table 3 Ablation Experiments

Model	Model configuration	QM(%)	IM(%)
Baseline	Complete model	55.6	30.2
ExperimentalA	Remove the semantic rewrite module	43.9	17.1
ExperimentalB	Use the mT5 decoder	46.1	23.5
ExperimentalC	Remove the semantic rewrite module and use the mT5 decoder	39.5	12.8

To verify the effectiveness of the core components of the RW-T5 model, this study conducted ablation experiments based on the CHASE dataset (which covers scenarios such as single-table, multi-table JOINs, and nested queries). Four comparative groups were set up focusing on key modules: the baseline model includes all core modules, experimental group A removed the semantic rewriting module, experimental group B used the decoder from the mT5 model, and experimental group C removed both the semantic rewriting module and used the decoder from the mT5 model. The experimental results show that the baseline model significantly outperforms the ablation models in terms of Question Matching (QM) and Interaction Matching (IM). The semantic rewriting module addresses the ambiguity in natural language, while the constraint decoder ensures the syntactical validity of the generated SQL. The functional modules exhibit a significant synergistic enhancement effect, where the absence of any module leads to a notable decline in model performance, thereby validating the indispensability of each component in the multi-turn Chinese NL2SQL task and

ISSN 2616-5775 Vol. 8, Issue 5: 38-45, DOI: 10.25236/AJCIS.2025.080504

providing a clear technical pathway for subsequent model optimization. (Table 3)

4. Conclusion and Outlook

This paper presents the RW-T5 model, which effectively addresses the problems of reference resolution and complex schema parsing in Chinese multi-turn dialogues through hierarchical semantic modeling and schema parsing techniques. Experimental results demonstrate that the model has achieved a new technical level in multi-turn semantic coherence and complex SQL generation capabilities, while the engineering system has validated its practical application value. Future research will focus on the following directions: (1) optimizing the logical reasoning process of SQL generation using reinforcement learning; (2) exploring fusion modeling methods for cross-modal data (such as charts and voice); and (3) researching few-shot learning frameworks to reduce the model's dependence on large-scale labeled data. These studies will further promote the deep application of NL2SQL technology in fields such as intelligent data platforms and business intelligence analytics.

References

[1] Androutsopoulos I, Ritchie G D, Thanisch P. Natural Language Interfaces to Databases - An Introduction[J]. Natural Language Engineering, 1995, 1(1): 29-81.DOI:10.1017/S135132490000005X.
[2] Lafferty J, McCallum A, Pereira F. Conditional random fields: Probabilistic models for segmenting and labeling sequence data[C]//Icml. 2001, 1(2): 3.
[3] Raffel C, Shazeer N, Roberts A, et al. Exploring the limits of transfer learning with a unified text-to-text transformer[J]. Journal of machine learning research, 2020, 21(140): 1-67.
[4] Bromley J, Guyon I, LeCun Y, et al. Signature verification using a "siamese" time delay neural network[J]. Advances in neural information processing systems, 1993, 7(4):737-744.
[5] LiuQ, Chen B, Lou J G, et al. Incomplete Utterance Rewriting as Semantic Segmentation[C]//Conference on Empirical Methods in Natural Language Processing, 2020: 2846-2857. DOI:10.18653/v1/2020.emnlp-main.227.
[6] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. Springer international publishing, 2015: 234-241.
[7] Lin K, Bogin B, Neumann M, et al. Grammar-based Neural Text-to-SQL Generation[J]. 2019. Available from: https://arxiv.org/pdf/1905.13326.
[8] Guo J, Si Z, Wang Y, et al. Chase: A large-scale and pragmatic Chinese dataset for cross-database context-dependent text-to-sql[C]//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 2021: 2316-2331.
[9] Zhang R, Yu T, Er H Y, et al. Editing-based SQL query generation for cross-domain context-dependent questions[C]//International joint conference on natural language processing; Conference on empirical methods in natural language processing. 2019: 5337-5348.
[10] Cai Y, Wan X. IGSQL: Database schema interaction graph based neural model for context-dependent text-to-SQL generation[J]. 2020. Available from: https://doi.org/10.48550/arXiv. 2011.05744.
[11] Wang B, Shin R, Liu X, et al. Rat-sql: Relation-aware schema encoding and linking for text-to-sql parsers[C]//Annual Meeting of the Association for Computational Linguistics. 2020: 7567-7578.