A surge in interest in the evaluation of the quality of chatbot conversation has been observed in recent years. We performed Dialogue Quality (DQ) and Nugget Detection (ND) subtasks in Chinese and English. However, the majority of existing conventional approaches are based on the long short-term memory (LSTM) model. The paper suggests a method for assisting customers in resolving problems. This subtask aims to automatically determine the status of dialogue sentences in a dialogue system’s logs. In conversation tasks, we developed fine-tuning methodologies for the transformer model. To evaluate and show the concept, we created a wide framework for testing and displaying the XLM-RoBERTa model’s performance on conversational texts. Finally, the experimental findings of the two subtasks demonstrated the efficacy of our strategy. The experimental findings for the DialEval-2 task showed that the suggested method’s performance is reasonably equal to that of the LSTM-based baseline model. The main contribution of our study is our suggestion of two crucial elements, namely, tokenization methods and fine-tuning procedures, to increase the conversation quality and nugget identification subtasks in dialogue assessment.