

**\*\* To be updated as needed \*\***  
**NTICR-19 MedNLP-CALL**

January 31, 2026: First version

これは、NTICR-19: MedNLP-CALL のサンプルデータを示したものであり、通報内容の緊急度分類（トリアージ）および要約を行う緊急通報タスクに関するものです。現在のサンプルデータセットは日本語で作成されています。

This describes sample data of the NTCIR-19: MedNLP-CALL, tasks about emergency calls for classifying the triage and summarizing the call. The current sample dataset is in Japanese.

### Material

このデータセットは、通報者（Caller）と指令員（Dispatcher）の会話、および対応する緊急度ラベル（非常に緊急, 準緊急, 低緊急）と要約から構成されています。

The dataset consists of a conversation between the caller and the dispatcher, along with the corresponding triage label (Very-emergency, Semi-emergency, and Low-emergency) and summary.

```
"id": 127,  
"conversation": "Dispatcher: もしもし、119 番です.... \n\nCaller: 救急です。....",  
"triage_label": "Very Emergency",  
"summary": "呼吸困難..."
```

### Input: Conversation

対話は、あらかじめ定義された主訴、報告症状、重症度レベルに基づき、LLM を用いて合成的に生成されました。対話の品質は、人間および LLM によって評価されました。

The dialogues were synthetically generated using an LLM with predefined medical complaints, reported symptoms, and severity level. The quality of the dialogue was evaluated by human and LLM.

### Output: Triage Label and Summary

Subtask 1: 緊急度ラベルは、非常に緊急、準緊急、低緊急の3段階で構成されています。緊急度ラベルは、日本の緊急度判定プロトコル (Ver.3) に基づいて専門家が注釈付けしました。

Subtask 1: The triage label consists of Very-emergency, Semi-emergency, and Low-emergency. The label was annotated by the expert based on the [Japanese Emergency Severity Assessment Protocol \(Version 3\)](#).

Subtask 2: 要約は、通話内容に基づいて専門家によって作成されました。

Subtask 2: The summary was provided by the expert based on the call content.

### Evaluation

#### Subtask 1:

トリアージラベルの予測は、F1 スコアおよびアンダートリージとオーバートリージの割合を用いて評価されます。アンダートリージは、実際のラベルよりも低い緊急度として患者を分類してしまう予測と定義されます。一方、オーバートリージは、実際のラベルよりも高い緊急度として患者を予測する場合に発生します。

The prediction of the triage label is evaluated using F1-score and the percentage of undertriage and overtriage. Undertriage is defined as a prediction that classifies a patient at a lower level than the actual label. In contrast, overtriage occurs when a patient is predicted to be more urgent than their actual label.

**Subtask 2:**

生成された要約は、BLEU および ROUGE を用いて評価されます。専門家が作成した要約は参照テキスト（リファレンス）として使用されます。

The generated summary is evaluated using BLEU and ROUGE. The summaries provided by the expert will serve as the reference text.

**Contact**

If you have any questions, please feel free to let us know.

Email: [mednlp@is.naist.jp](mailto:mednlp@is.naist.jp)

Web: <https://sociocom.naist.jp/mednlp-call/>