

Proposal Requirement

- **Problem Statement and Task Definition**

In speech recognition and transcription, the resulting text often lacks punctuation, leading to blocks of text that are difficult to read and understand. This poses a significant challenge in scenarios where accurate and comprehensible transcriptions are needed, such as in legal documentation, medical records, and educational materials. The task at hand is to develop a system that can accurately insert punctuation into such unpunctuated text, enhancing readability and maintaining the original spoken content's integrity.

- **Description of the challenges**

- Speech recognition systems convert speech to unpunctuated text, which is difficult to read and comprehend.
- The punctuation restoration process is complex, involving understanding context, grammar, and the nuances of spoken language.

- **Input/Output Behavior with Concrete Examples**

- Input

Text without punctuation, like transcribed speech from open data platforms.

```
1|hello how are you I'm fine thank you and you I'm doing well too
```

- Output

The same text with appropriate punctuation is restored.

```
1|Hello, how are you? I'm fine, thank you. And you? I'm doing well too.
```

- **Related works**

- Yi, J., Tao, J., Bai, Y., Tian, Z., & Fan, C. (2020). Adversarial transfer learning for punctuation restoration. arXiv preprint arXiv:2004.00248.
- Cho, E., Niehues, J., & Waibel, A. (2012). Segmentation and punctuation prediction in speech language translation using a monolingual translation system. In Proceedings of the 9th International Workshop on Spoken Language Translation: Papers (pp. 252-259).
- Tilk, O., & Alumäe, T. (2016, September). Bidirectional Recurrent Neural Network with Attention Mechanism for Punctuation Restoration. In Interspeech (Vol. 3, p. 9).

- **Methodology**

- Approach

- 1. RT-Based Approach

- Employing the BERT model as a foundational method for punctuation restoration.

- BERT's strong suit in understanding context and sentence structures makes it ideal for predicting punctuation in unpunctuated text.

- 2. T5 and BART Models

- In parallel, using T5 and BART models is known for their effectiveness in text-to-text tasks like punctuation restoration.

- Fine-tuning these models on specific datasets like BNC-corpus, adjusting for sentence length and contextual cues.

- Why This Method

- These models have shown promising results in NLP tasks, including punctuation restoration.

- Challenges and Phenomena

- Balancing training data quantity and quality, sentence boundary detection.

- Algorithms and Trade-offs

- Exploring the trade-off between the accuracy of punctuation prediction and the model's efficiency.

- **Evaluation Metrics**

- Mainly using the F1-score to measure the performance of the restored punctuation against a ground truth set.

- Considering the balance of precision and recall in the model's predictions.

- **Baselines**

- Simple rule-based methods or classifiers.

- Compared with existing models like a non-fine-tuned T5 or BERT model.

Work Plan

- **Time Schedule**

- Literature review and baseline model setup.
- Data collection and preprocessing from selected datasets.
- Model fine-tuning and initial testing.
- Further model refinement and evaluation.
- Final testing, analysis, and report writing.

- **Discussion**

<https://hackmd.io/@userwei/BJteGHB1R>

- **Repo**

https://github.com/chou-ting-wei/NYCU_AI-Final-Project