Analysis of Nested Named Entity Recognition Solutions

Muhammed Esad Mazi

m.mazi@innopolis.university

1 Personal and University Information

- **Name and Surname:** Muhammed Esad Mazi

- **University Email:** m.mazi@innopolis.university

- **Nickname in CodaLab:** esadmazi

- **GitHub Repository Link:** `https://github.com/esadmazi/IU-NLP-2024`

2 Introduction

This report documents the process and methodology employed in the development of a Named Entity Recognition (NER) system, detailing various approaches tested and their respective performance metrics. The objective was to maximize the F1 score by experimenting with different configurations and model architectures.

3 Approaches and Methodologies

3.1 Initial Approach

Our initial model was a straightforward implementation using the BERT model pre-trained on a multilingual corpus. The setup included:

- Pre-processing the data to fit BERT's input requirements.

- Fine-tuning BERT on the NER task with default parameters.

This model provided a baseline F1 score upon which further improvements were based.

3.2 Improved Model Configurations

After establishing a baseline, various configurations were explored:

- **Adjusting Token Length:** The token length was modified to analyze its impact on the F1 score, optimizing computational efficiency and model performance.

- **Selective Fine-Tuning:** Instead of fine-tuning the entire BERT model, only the classifier layers were optimized to adapt to the NER task more specifically.

3.3 Final Model

The final model configuration that yielded the best F1 score involved:

- Truncated input sequences to 128 tokens to balance between computational efficiency and context retention.

- Training only the classifier layer of the BERT model to specialize its predictions for the NER task without altering the pre-trained contextual embeddings.

4 Results and Discussion

The results demonstrate that the selective training of the classifier layers and adjusting the token length provided the best performance.

5 Conclusion

This exercise showcased the importance of model configuration and fine-tuning in achieving high performance in NER tasks. Future work could explore more complex combinations of model architectures and training strategies to further enhance the accuracy and efficiency of the NER system.