

Rubric for Assignment Evaluation

The assignment will be evaluated based on the following criteria. Each task consists of multiple components, and the total marks are distributed accordingly.

Task 1

- **Viva (Conceptual Understanding and Explanation):** 10 marks
- **Evaluation based on F1-score:** 10 marks
 - Tag-level : 5 marks
 - If score ≥ 90 give 5 marks
 - If score ≥ 85 and < 90 give 2.5 marks
 - Else 0 marks
 - Chunk-level : 5 marks
 - If score ≥ 70 give 5 marks
 - If score ≥ 65 and < 70 give 2.5 marks
 - Else 0 marks
 - **Scores won't be rounded off**
- **Implementation of BIO Encoding:** 5 marks
- **Implementation of RNN with FastText and GloVe embeddings:** 2.5 marks
- **Implementation of GRU with FastText and GloVe embeddings:** 2.5 marks
- **Inclusion of training and validation plots:** 5 marks

Task 2

- **Viva (Conceptual Understanding and Explanation):** 10 marks
- **Evaluation based on Accuracy Score:** 10 marks
 - If score ≥ 80 give 10 marks
 - If score ≥ 70 and < 80 give 5 marks
 - Else 0
 - **Scores won't be rounded off**
- **Model creativity:** 5 marks
- **Proper dataset formatting and preprocessing:** 2.5 marks
- **BERT, RoBERTa, BART fine-tuning:** 5 marks
- **Inclusion of training and validation plots:** 2.5 marks

Task 3

- **Viva (Conceptual Understanding and Explanation):** 10 marks
- **Fine-tuning of SpanBERT-CRF for question-answering:** 5 marks
- **Fine-tuning of SpanBERT for question-answering:** 5 marks
- **Inclusion of training and validation plots:** 5 marks
- **Evaluation on the validation set:** 5 marks

- If score ≥ 50 give 5 marks
 - Else 0
-

Penalties

- **Failure to submit the report:** -10 marks
- **Failure to implement the testing function:** -10 marks
- **Failure to submit a saved model:** -5 marks
- **Padding token considered while calculating F1 in Task 1:** -10 marks
- **Failure to submit any of the required dataset files (train_task_1.json, val_task_1.json, train_task_2.json, val_task_2.json):** -5 marks per missing file