

Answers for the Questions and Description:

- The precision, recall and F1 score on the dev data for the **TASK-1** is **precision: 78.10%; recall: 73.53%; F1-Score: 75.74%** respectively.
- The precision, recall and F1 score on the dev data for the **TASK-2** is **precision: 88.40%; recall: 89.82%; F1-Score: 89.10%** respectively.

Brief Explanation of the solution:

Hyperparameters used in model architecture and training are:

- embedding dim = 100
- number of LSTM layers = 1
- LSTM hidden dim = 256
- LSTM Dropout = 0.33
- Linear output dim = 128
- momentum=0.9
- lr = 0.1
- tag_size
- batch_size
- criterion
- optimizer
- epochs

Solution Description:-

Firstly, I Loaded and prepared the dataset for all train, test and dev data from the provided dataset files. Then I created the Word Dictionary by setting the word as Key and its index in the corpus as the Value and created the training, dev and the test dataset sentence vectors. I have used custom Collate functionality to determine how individual samples are combined into batches during training or testing, since each sample in the dataset may have different sizes or shapes, they cannot be directly combined into batches. Thus, we pad all samples/sentences to a fixed length and combines them into a tensor. After, for Task-1, I create BiLSTM network with the provided hyperparameter values and train it using the training dataset provided and use it on the dev and test data to get the accuracy, precision, recall and f1-score on the dev data. We follow the same procedure for the Task-2 and in addition build an embedded matrix for the GloVe word embeddings along with maintaining the case sensitivity of the words in the embedded matrix, and use it in the training, test and dev dataset to get the accuracy, precision, recall and f1-score on the dev data