

# COL774 Assignment-4 Report

Name : Ujjwal Mehta | Entry No. : 2020CS10401

Name : Avval Amil | Entry No. : 2020CS10330

## Non- Competitive Part

### 1. Convolutional Neural Network:

In this part of assignment we have created our own convolutional neural network(CNN) in order to classify the genre of the book based on its cover image and in order to do so we have created our own Net class as the child class of the **nn.Module in pytorch** and then we have defined its constructor as well as forward function as described in the assignment and use this finally to train our model. For the preprocessing part, we first load the csv files using pandas and then we have our own `convert_to_tensor` function which reads and converts the image to a tensor from a corresponding image name. The accuracy that we get on **non competitive test dataset** is **14%**. Hyperparameters used are : **Batch Size:1, Epoch : 5, Optimizer : SGD, Learning Rate : 0.001.**

### 2. Recurrent Neural Network:

In this part of assignment we have implemented a recurrent neural network in order to classify the book genre based on its title. In order to implement this RNN we have to first convert the input sentence which is a book title into an embedding which we can give as input to our RNN and that embedding is created during preprocessing via the **Glove Embeddings** which transforms a given word token into 300 size vector i.e. we first do preprocessing of our data inside the `tokenize` function by first tokenizing our sentence then converting it into a matrix using the glove embeddings and finally we do a padding in order to a sentence embedding of size  $10 * 300$ , using this preprocessing we get a 3d tensor of size  $N * 10 * 300$ . Then we implement the `BidirectionalRNN` class by inheriting the **nn.Module in pytorch** and using its `inbid rnn` class, we define the constructor and forward function. Finally we optimize using training data to train our model. The accuracy that we get on **non competitive test dataset** is **43.8%**. Hyperparameters used are : **Batch Size:300, Embedding Size : 10, Epoch : 25, Optimizer : Adam, Learning Rate : 0.0006.**

## Competitive Part

In order to implement our architecture for this part, we tried and tested out various different kinds of model before we arrived at our final accuracy. First we tried using

LSTM in our earlier RNN architecture where after several improvements of our hyperparameters we got an accuracy of 51%. Then tried building our own TextCNN model in order to classify the book title into different genres which didn't prove to be of much use and that got us an accuracy of 47%. After even trying to combine the image CNN with TextCNN and not seeing improvement we finally got improvement using the **Hugging Face Library** where we imported and tested out various transformers models to improve our accuracy. First we tested out its DistilBERT architecture using which we got an accuracy of 63% and after we further tested out and got **RoBERTa Large** as our final architecture upon which we got **65.9%** accuracy. This architecture has 24-layer, 1024-hidden, 16-heads and 355M parameters. Hyperparameters used are : **Batch Size:16, Epoch : 6, Learning Rate : 2e-5.**