

# Assignment 1

## Advanced Machine Learning (CS566)

Date: 30-Jan-2021

**Deadline:10-feb-2021**

### Instructions:

1. The assignments should be completed and uploaded before the deadline.
2. Markings will be based on the correctness and soundness of the outputs. Marks will be deducted in case of plagiarism.
3. Proper indentation and appropriate comments are mandatory.
4. Submission Details :
  - a. You should zip all the files and name the zip file as <all-roll\_no.zip>, eg. 1921cs28\_1921cs29\_1921cs30.zip.
  - b. Submit your code along with images/answers to questions for proper evaluation.
  - c. Submission link : <https://www.dropbox.com/request/c9BxJLhU2oKdYG7tntsg>
5. Make necessary assumptions if required. For further clarification, you can contact the TAs from : <http://172.16.1.5/~asif/CS566-AML/>

The assignment targets document classification of the BBC news dataset. The dataset/files has already been converted to a single csv file for easier processing. The dataset has 2 features, first is the 'Article' which is the input for the task while the second 'Class' features classifies the article into one of 5 classes : business, entertainment, politics, sport and tech . You can find more info at <http://mlg.ucd.ie/datasets/bbc.html>

Download dataset from :

<https://www.dropbox.com/s/acabyI01j43cf9i/bbc.csv?dl=0>

### What you have to do:

1. Implement the LSTM and GRU models for the sequence classification task. Implement the models with minimum hidden layers and also keep all hyperparameters exactly the same for both the models. Run the model for 5 epochs. Use the seed value of '1' for consistent results. Compare the two models on the basis of performance metrics (accuracy, precision, recall) and number of parameters in each model. Write your results in a separate file when submitting. Steps are
  - a. Preprocess/clean the data(remove stopwords etc.)
  - b. The input to LSTM/GRU will be the articles, with tokens represented with a 100 dimension vector each(ie. Embedding size = 100).
  - c. Make sure to pad the input to the maximum article length.
  - d. Use 70:10:20 split for training, validation and testing
  - e. Use of dataloader is optional.

2. The general consensus for LSTM/GRU for classification tasks is to use ReLU for hidden layers and sigmoid/softmax for final output. Run your model(created above) with different activation layers(Relu, softmax/sigmoid, tanh, elu, leaky ReLU) at their appropriate positions and calculate the performance change for each combination. Specify which works best in case of LSTM/GRU. Write your results in a separate file when submitting.
3. Evaluation:
  - a. Perform 3-fold cross-validation
  - b. Compute the overall accuracy for each of the 3-folds (for all the classes together)
  - c. Compute the class-wise accuracy for each of the 3-folds