January-May 2022 Semester

CS691: Fundamentals of Deep Learning

Programming Assignment II

Date: **8th April, 2024**
Deadline for submission of PDF file of report: **Monday, 29th April, 2024**

Datasets:

(1) Image dataset for classification
(2) Image captioning data set with captions in English
(3) Machine translation data set for translation from English to an Indian language

Task 1: Image classification using a MLFFNN with two hidden layers, and with Deep CNN features for an image as the input to the MLFFNN, with (a) VGGNet as Deep CNN and (b) GoogLeNet as Deep CNN.

Task 2: Image classification using a CNN with CL1, PL1, CL2 and PL2 as the layers. Use kernels of size 3x3, stride of 1 in the convolutional layers. Use the mean pooling with a kernel size of 2x2 and stride of 2 in the pooling layers. Use 4 feature maps in CL1. The number of feature maps in CL2 is a hyperparameter.

Task 3: Image captioning using a CNN with NetVLAD as encoder and a single hidden layer RNN based decoder.

Task 4: Same as Task 3 using the single hidden layer LSTM network based decoder

Task 5: Machine translation with encoder and decoder, each built using a single hidden layer LSTM network.

GloVe representation is to be used as the word representation for English.

IndicBERT representation to be used as the word representation for Indian languages.

Performance of the image captioning system and the machine translation system is to be given as BLEU@k scores with k = 1, 2, 3 and 4

Report should include details of experimental studies, results, observations and analysis of results.