Deep Learning CSA 250

Assignment 3

Aalo Majumdar

SR 16116

M.Tech Res, CSA

Drive link for Model parameters:

https://drive.google.com/open?id=1xRdwb6mzIfsbsP4eJOi85nZarhRsHQZ6

STEP 1:

Used sklearn's Logistic Regression for Step 1

Extracted TF-IDF features of premise and hypothesis and vectorised them separately.
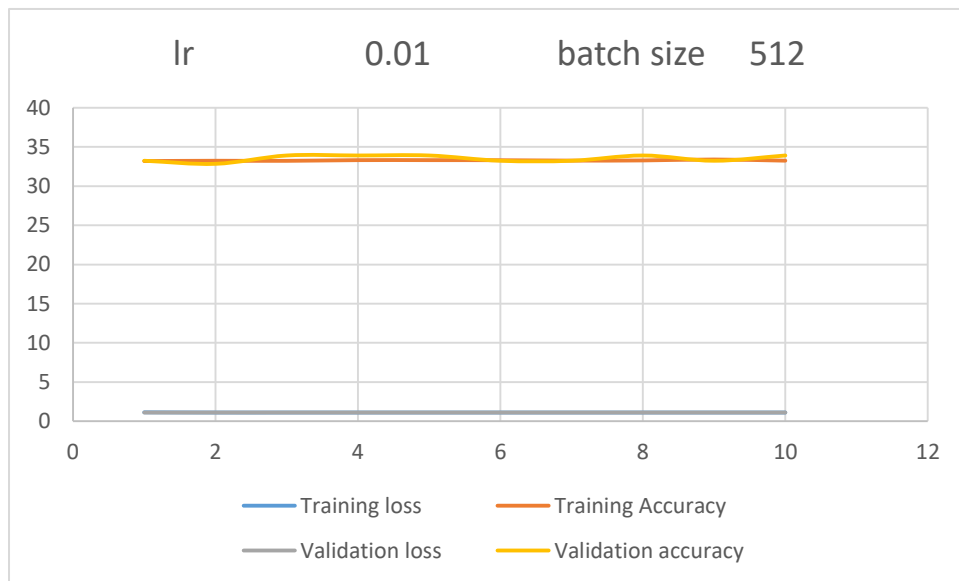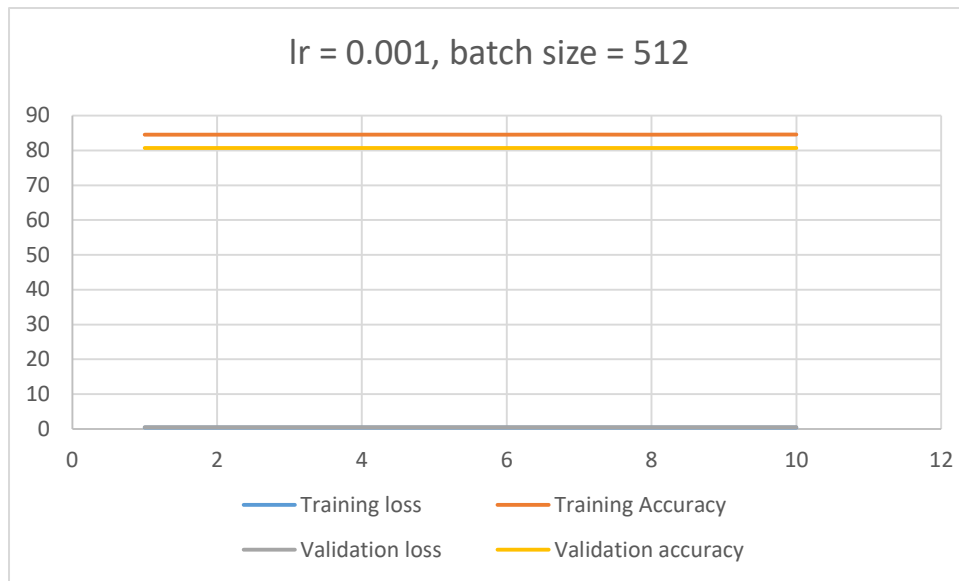
Then fitted Logistic regression model on the tf-idf vectors.
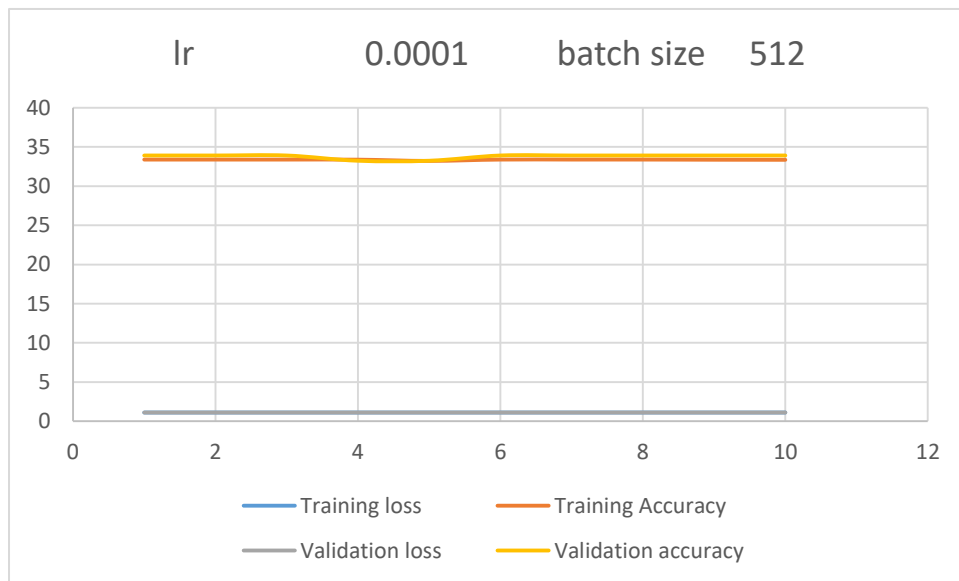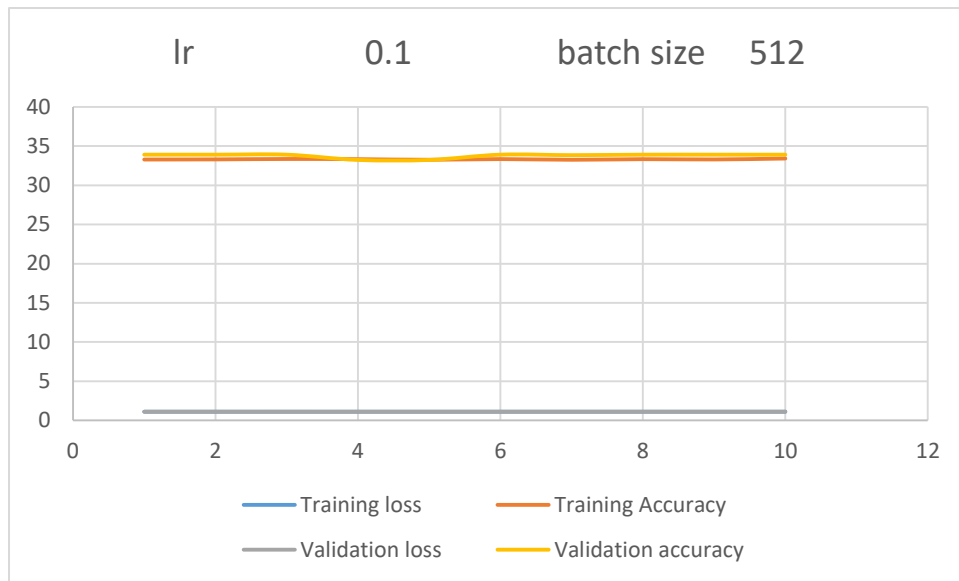

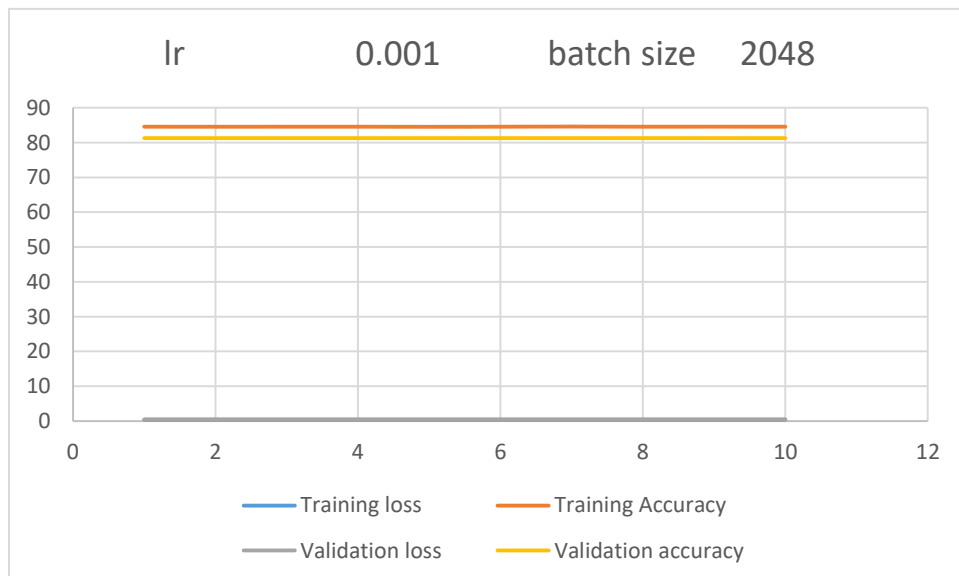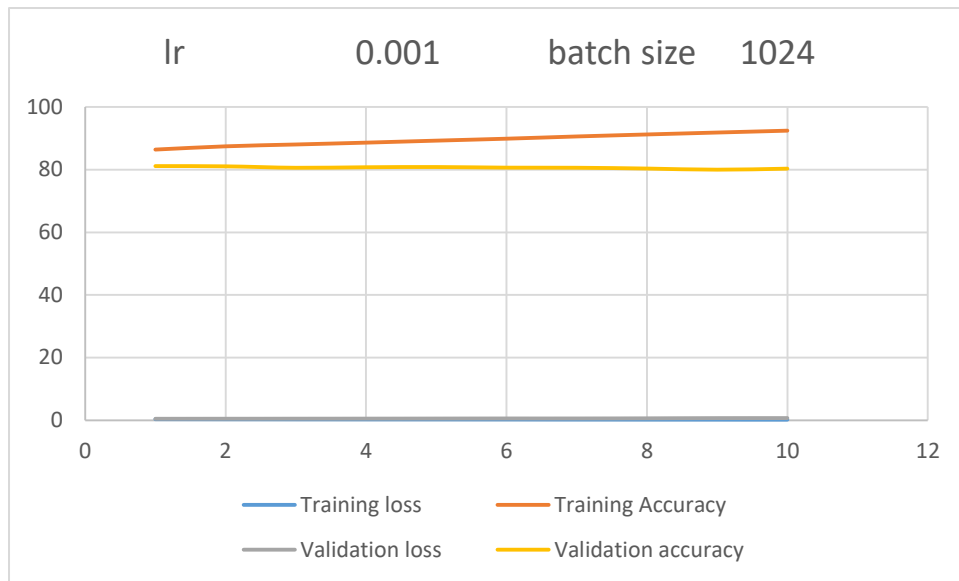Training accuracy = 63.086

Test accuracy = 62.958

STEP 2:

Used BiLSTM for Step 2

Tuned learning rate (lr) and batch size and evaluated the graphs of training loss, training accuracy, validation loss, validation accuracy vs number of epochs

lr = 0.001, batch size = 512

Legend: Training loss, Training Accuracy, Validation loss, Validation accuracy

lr   0.01   batch size   512

Legend: Training loss, Training Accuracy, Validation loss, Validation accuracy

lr   0.1   batch size   512

lr   0.0001   batch size   512

lr   0.001   batch size   1024

lr   0.001   batch size   2048

Got the best results with batch size = 2048 and learning rate = 0.001

Test loss = 0.488

Test accuracy = 80.93