

Deep Learning Project 3

Name: Ahmed Sanin MV

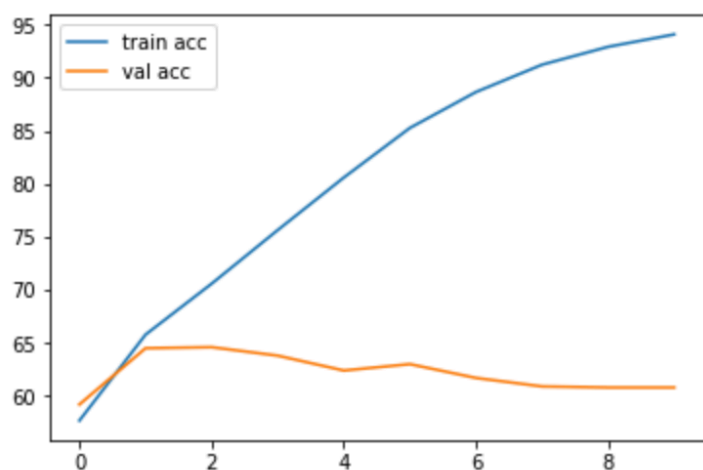
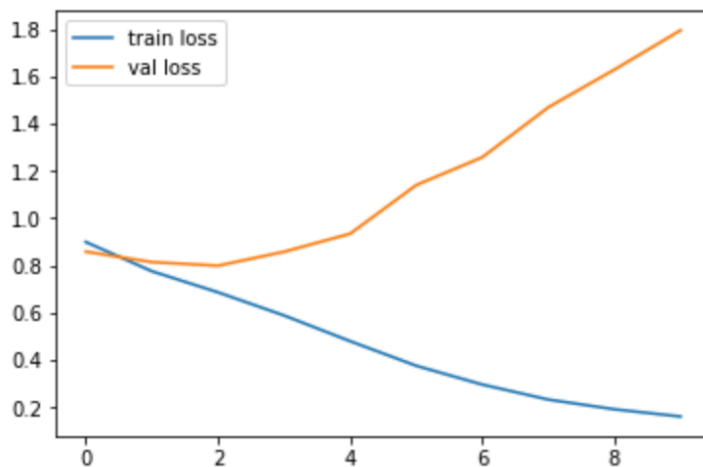
SrNo: 14165

1. For the first part I used simple logistic regression using tf-idf feature for the SNLI data set
Following preprocessing where done . conversion to lower case ,stop words removal,stemming,tokenisation,special character removal,HTML tags processing,Including UNK token to handle out-of-vocabulary words texting accuracy was found to be 48.9 %

2. Implementation of deep neural network model
Different architectures where

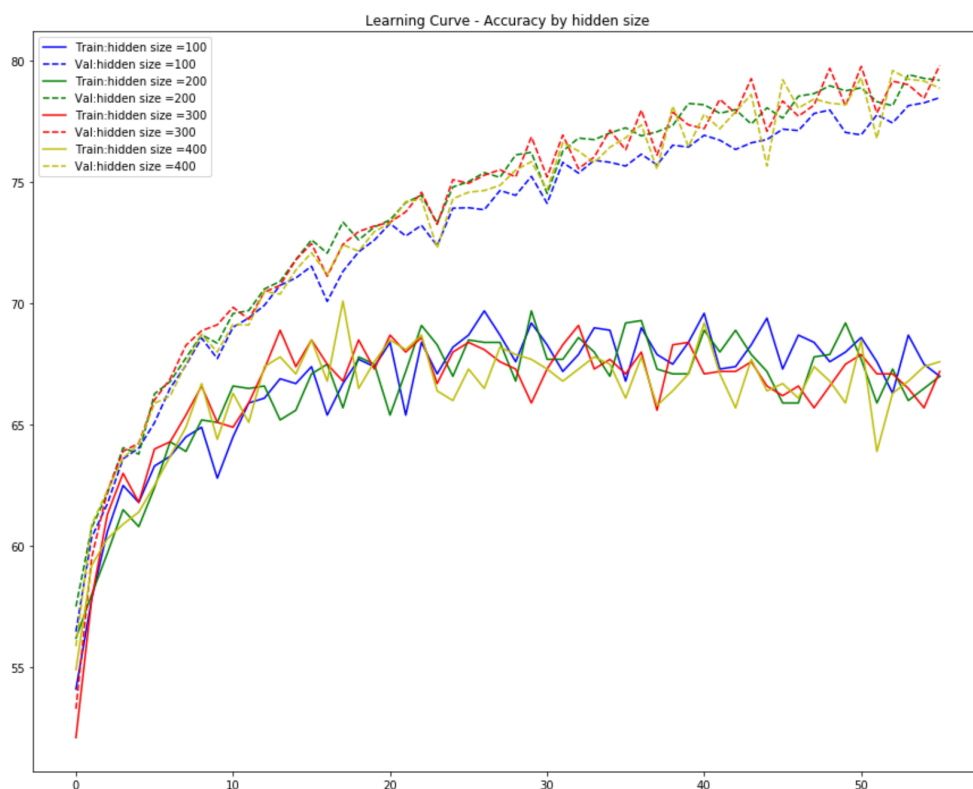
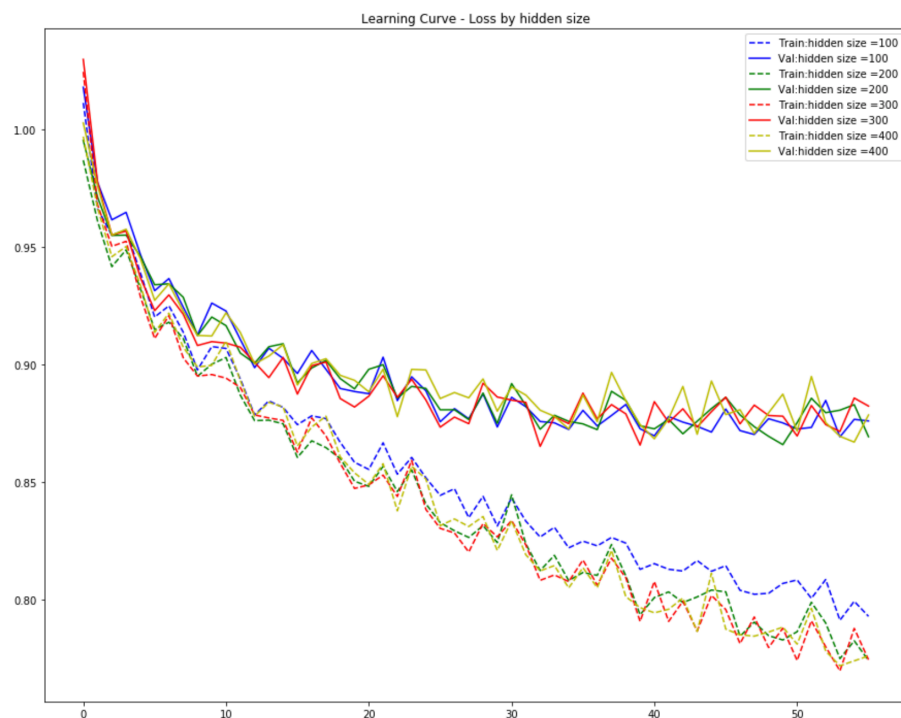
A) Architecture 1

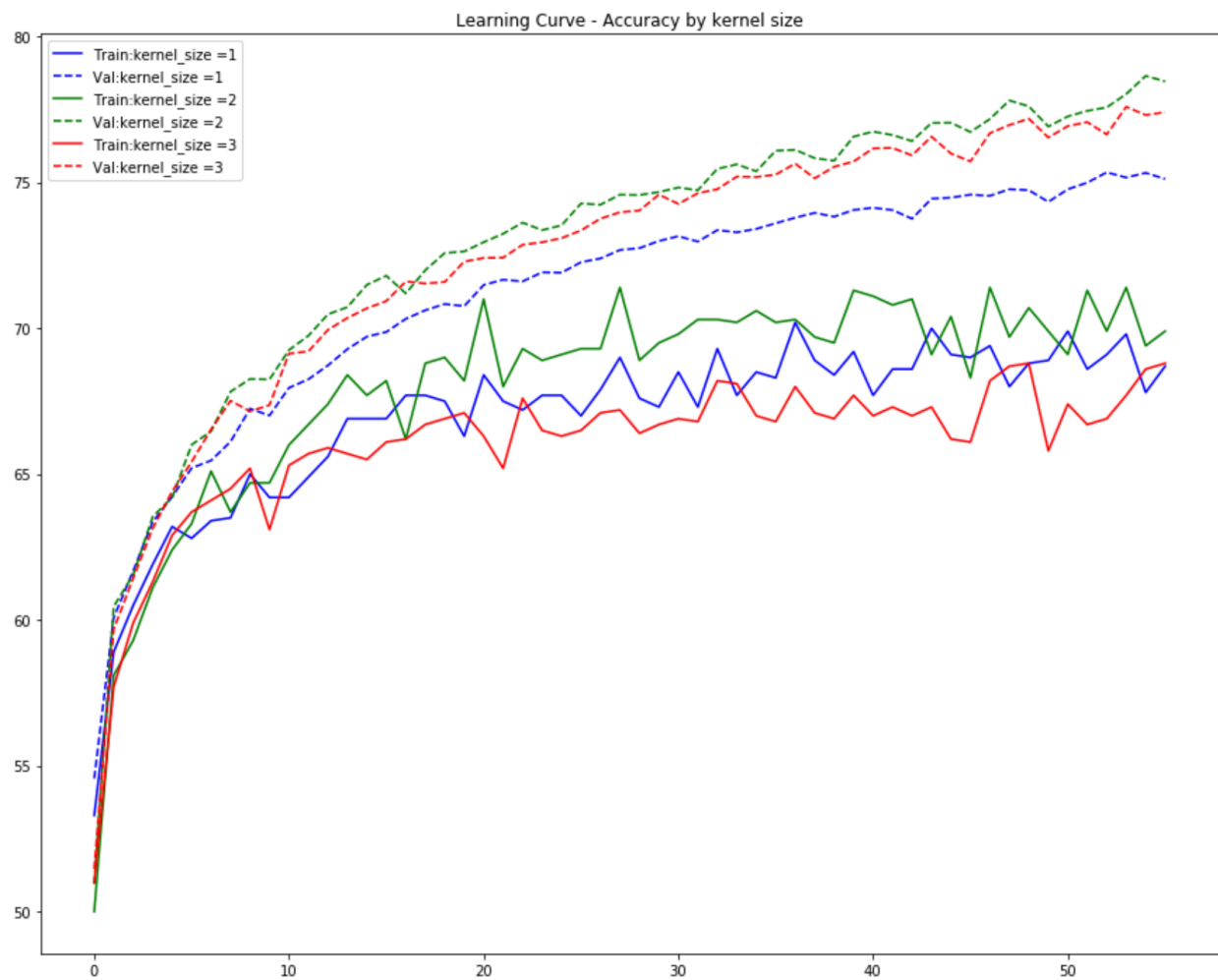
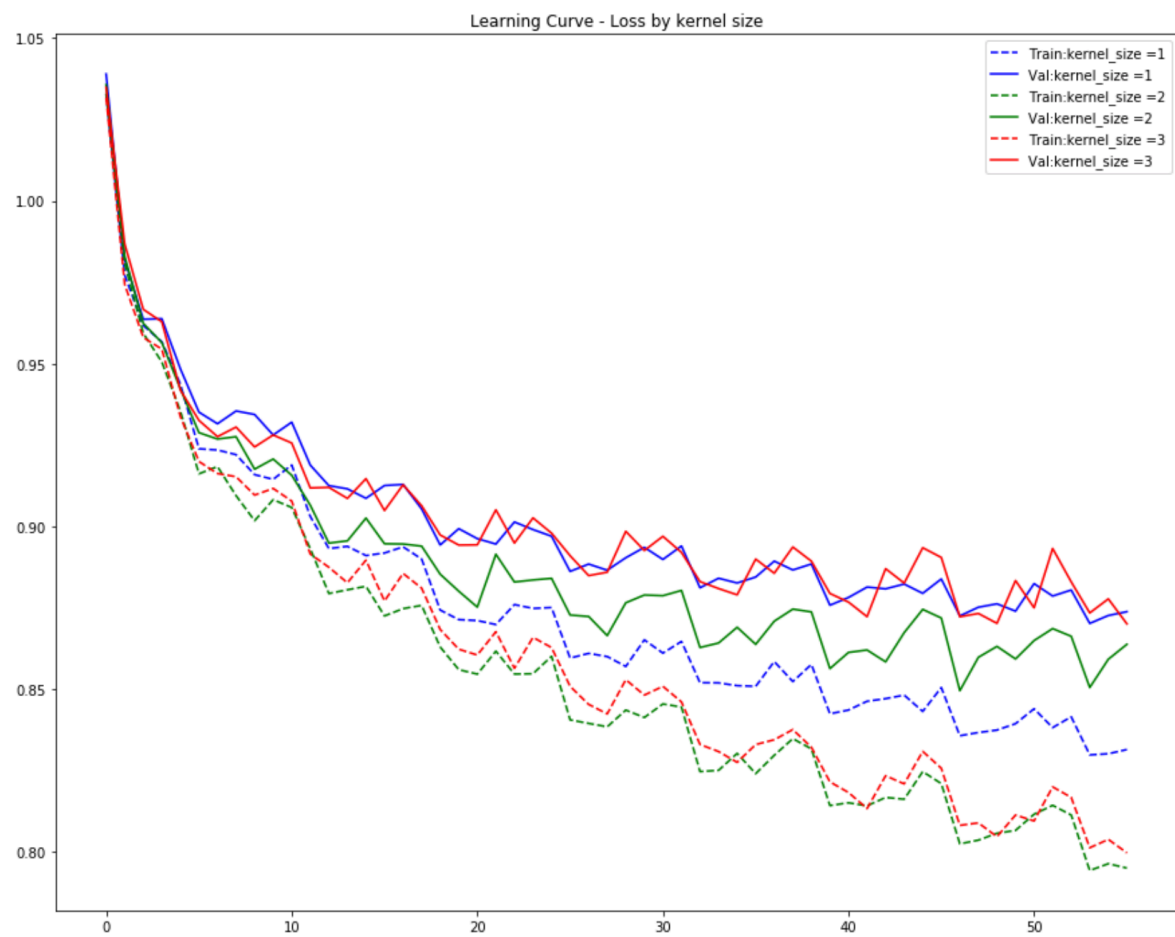
```
CNN(  
  (embedding): Embedding(100002, 300, padding_idx=0)  
  (conv1): Conv2d(1, 300, kernel_size=(5, 300), stride=(1, 1))  
  (conv2): Conv2d(300, 100, kernel_size=(5, 1), stride=(1, 1))  
  (max1): MaxPool2d(kernel_size=(10, 1), stride=(10, 1), padding=0, dilation=1, ceil_mode=False)  
  (linear1): Linear(in_features=200, out_features=500, bias=True)  
  (linear2): Linear(in_features=500, out_features=3, bias=True)  
  (dpl): Dropout(p=0.8)  
)
```



2. RNN models

- single-layer bi-directional GRU model
- training process
- parameter-tuning(hidden size, dropout rate),
- evaluation on MNLI by genres
- 3 correct/incorrect samples
- Bonus: Fine-tuning on MNLI by genres





3. LSTM _ Glove model (final model)

Layer (type)	Output Shape	Param #	Connected to
input_5 (InputLayer)	(None, 42)	0	
input_6 (InputLayer)	(None, 42)	0	
embedding_1 (Embedding)	(None, 42, 300)	12717300	input_5[0][0] input_6[0][0]
time_distributed_2 (TimeDistrib	(None, 42, 300)	90300	embedding_1[2][0] embedding_1[3][0]
lstm_1 (LSTM)	(None, 300)	721200	time_distributed_2[2][0] time_distributed_2[3][0]
batch_normalization_1 (BatchNor	(None, 300)	1200	lstm_1[0][0]
batch_normalization_2 (BatchNor	(None, 300)	1200	lstm_1[1][0]
concatenate_1 (Concatenate)	(None, 600)	0	batch_normalization_1[0][0] batch_normalization_2[0][0]
dropout_1 (Dropout)	(None, 600)	0	concatenate_1[0][0]
dense_3 (Dense)	(None, 600)	360600	dropout_1[0][0]
dropout_2 (Dropout)	(None, 600)	0	dense_3[0][0]
dense_4 (Dense)	(None, 600)	360600	dropout_2[0][0]
dropout_3 (Dropout)	(None, 600)	0	dense_4[0][0]
dense_5 (Dense)	(None, 600)	360600	dropout_3[0][0]
dropout_4 (Dropout)	(None, 600)	0	dense_5[0][0]
dense_6 (Dense)	(None, 3)	1803	dropout_4[0][0]
Total params: 14,614,803			
Trainable params: 1,896,303			
Non-trainable params: 12,718,500			

Got a text accuracy of 80%

Google drive link for model : https://drive.google.com/file/d/1A_TqHy6PMPQPKVci3nAIMWES9sWcgSIZ/view?usp=sharing

