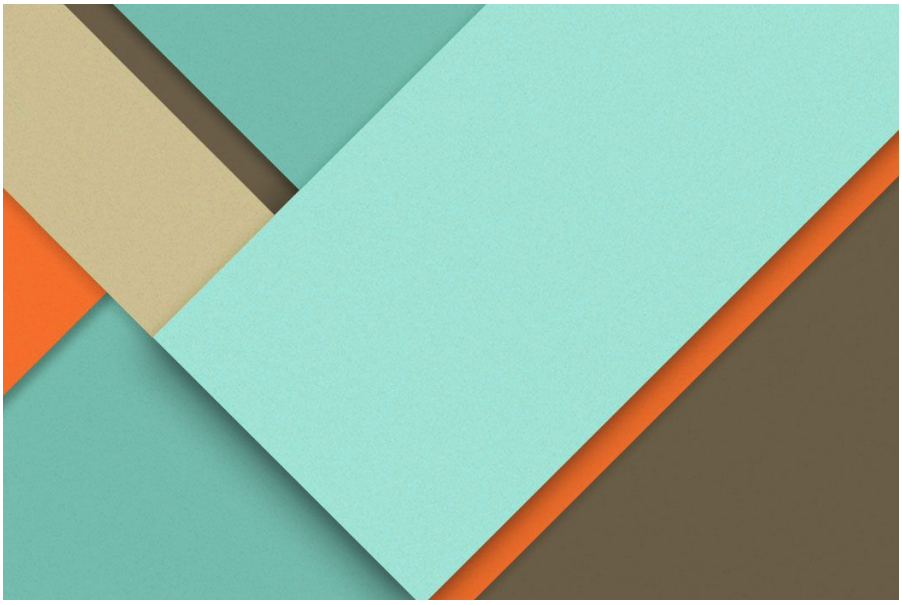

AML T2

10.10.2018

- Sanketh Rangreji 01FB15ECS267
- Shahid Ikram 01FB15ECS274
- Sondhi Nishant 01FB15ECS298
- Sumanth S. Rao 01FB15ECS314

Objectives

Developing model that can determine semantic similarity between 2 sequences.

Using distributed word representations and building models using Keras embedding layer.

Using Keras for sequence processing.

Model Architecture

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	(None, 791)	0	
input_2 (InputLayer)	(None, 791)	0	
embedding_1 (Embedding)	(None, 791, 100)	500000	input_1[0][0] input_2[0][0]
lstm_1 (LSTM)	(None, 50)	30200	embedding_1[0][0] embedding_1[1][0]
concatenate_1 (Concatenate)	(None, 100)	0	lstm_1[0][0] lstm_1[1][0]
dropout_1 (Dropout)	(None, 100)	0	concatenate_1[0][0]
batch_normalization_1 (BatchNor	(None, 100)	400	dropout_1[0][0]
dense_1 (Dense)	(None, 25)	2525	batch_normalization_1[0][0]
dropout_2 (Dropout)	(None, 25)	0	dense_1[0][0]
batch_normalization_2 (BatchNor	(None, 25)	100	dropout_2[0][0]
dense_2 (Dense)	(None, 1)	26	batch_normalization_2[0][0]

Total params: 533,251
Trainable params: 33,001
Non-trainable params: 500,250

Specifications

No of records fetched : **40,000**

No. of records used for training : **22,400**

No. of records used for cross-validation : **5,600**

No. of records used for testing : **12,000**

No. of epochs : **5**

Goals Achieved

We were able to successfully build the specified architecture and obtained results for the same.

Training Accuracy : **71.12%**

Cross Validation Accuracy : **71.29%**

Testing Accuracy : **72.71%**

Loss : **0.1898**