**Experiments**

EX1

Batch Size = 64 for training

Epochs = 30

Num of samples

1238 total pairs

868 training pairs

185 validation pairs

185 test pairs

Embeded dimension = 128

Latent dimension = 512

Numbers heads = 6

learning\_rate=1e-4

Loss: 1.2315

Accuracy: 0.3569

Val\_loss: 1.2878

Val\_accuracy: 0.3618

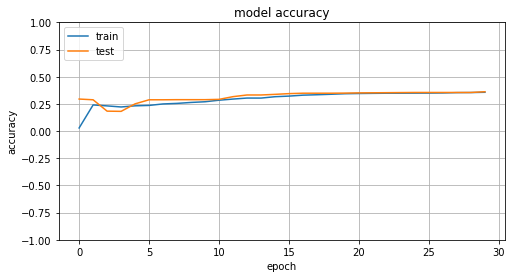
Bleu-4

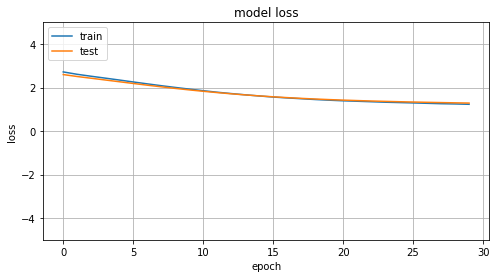
Scores

[0.1793958582523853, 0.325220509812945, 0.08977326109967479, 0.2846138428754854, 0.19374017157980686, 0.2092314418483734, 0.19374017157980686, 0.22596137858748616, 0.16611358241134278, 0.2092314418483734, 0.33194838105495905, 0.14242643658526624, 0.18059702173007178, 0.2092314418483734, 0.2092314418483734, 0.1793958582523853, 0.1221169849202633, 0.2154521273737615, 0.22596137858748616, 0.16611358241134278, 0.24402902432875526, 0.09695144123261011, 0.2570342453335712, 0.22596137858748616, 0.1793958582523853, 0.056585270274915354, 0.22596137858748616, 0.1221169849202633, 0.14242643658526624, 0.14242643658526624]

Average Bleu-4

0.19174615897319894





EX2

Batch Size = 64 for training

Epochs = 100

Num of samples

1238 total pairs

868 training pairs

185 validation pairs

185 test pairs

Embeded dimension = 128

Latent dimension = 512

Numbers heads = 6

learning\_rate=1e-4

Loss: 0.4671

Accuracy: 0.6947

Val\_loss: 1.2654

Val\_accuracy: 0.3868

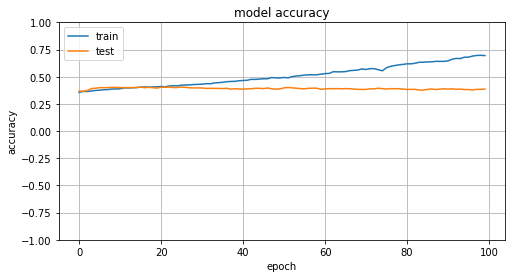
Bleu-4

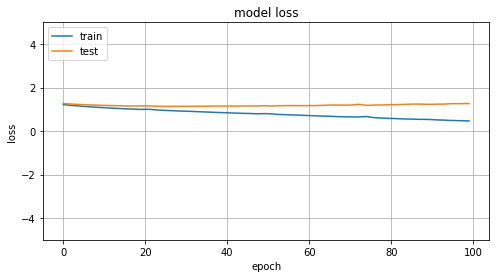
Scores

[0.5671276403520978, 0.5274742318366016, 0.537284965911771, 0.5274742318366016, 0.5946035575013605, 0.5671276403520978, 0.5327106629263971, 0.5946035575013605, 0.5577503997480662, 0.48832168314095387, 0.5885661912765424, 0.5623413251903491, 0.33125669191122536, 0.5721248424548514, 0.5533409598501607, 0.5946035575013605, 0.5410822690539396, 0.5450176720923848, 0.5336172773868456, 0.5410822690539396, 0.5623413251903491, 0.4760116549244004, 0.5919814398339353, 0.5946035575013605, 0.5577503997480662, 0.5721248424548514, 0.5533409598501607, 0.5229027085252375, 0.3945277164315128, 0.5631101643283382]

Average Bleu-4

0.541540213188904





EX3

Batch Size = 64 for training

Epochs = 30

Num of samples

9963 total pairs

6975 training pairs

1494 validation pairs

1494 test pairs

Embeded dimension = 128

Latent dimension = 512

Numbers heads = 6

learning\_rate=1e-4

Loss: 1.5741

Accuracy: 0.3531

Val\_loss: 1.9979

Val\_accuracy: 0.3007

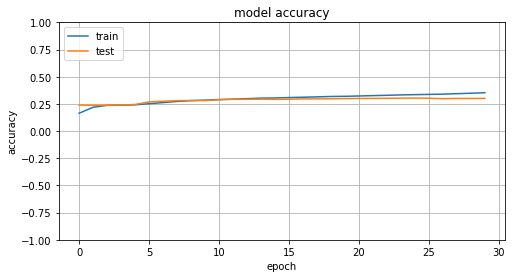
Bleu-4

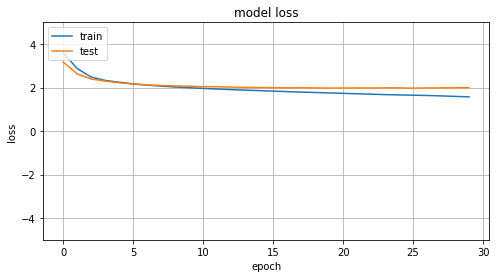
Scores

[0.5533409598501607, 0.5139409225754329, 0.5139409225754329, 0.5533409598501607, 0.5139409225754329, 0.5623413251903491, 0.5491004867761125, 0.5695988432761473, 0.5623413251903491, 0.5266403878479265, 0.5695988432761473, 0.5200970009529277, 0.5734090813814468, 0.5, 0.5292031904718658, 0.5300714512917181, 0.5336172773868456, 0.5336172773868456, 0.5533409598501607, 0.5169731539571706, 0.5555238068023582, 0.4832697830906221, 0.5233175696960528, 0.5139409225754329, 0.5336172773868456, 0.5336172773868456, 0.5345224838248488, 0.5623413251903491, 0.5814307369682193, 0.5109955811291784]

Average Bleu-4

0.5389244634411074





EX4

Batch Size = 64 for training

Epochs = 30

Num of samples

29271 total pairs

20491 training pairs

4390 validation pairs

4390 test pairs

Embeded dimension = 256

Latent dimension = 512

Numbers heads = 6

learning\_rate=1e-4

Loss: 1.9540

Accuracy: 0.4069

Val\_loss: 3.2032

Val\_accuracy: 0.2309

Bleu-4

Scores

[0.5623413251903491, 0.4653024295510498, 0.5026386140359251, 0.33652780691981066, 0.5200970009529277, 0.5336172773868456, 0.4215503093673728, 0.5623413251903491, 0.5247357977607321, 0.44811252183533623, 0.6447788067558897, 0.5338508734369847, 0.5109955811291784, 0.5750816584478015, 0.5734090813814468, 0.5154458901398173, 0.4685387070236361, 0.6116908849062525, 0.5659119256652702, 0.6039072332841391, 0.5588805988556161, 0.2572851870647722, 0.463502386414385, 0.4518010018049224, 0.5151431317771185, 0.4728708045015879, 0.5806030754582867, 0.4653024295510498, 0.4502457257096904, 0.5019724248795793]

Average Bleu-4

0.5066160605459374

