

270722_2257 - DistilBertSQuADv1

Max Sentence length: 256

Batch size: 32

Optimizer: AdamW (

Parameter Group 0

amsgrad: False

betas: (0.9, 0.999)

eps: 1e-08

lr: 5e-05

maximize: False

weight_decay: 0.9

)

Training Data: 21931

Validation Data: 10570

----- Epoch 1 / 5 -----

Train Loss: 2.098

Dev Set Scores --	EM: 28.09%	F1 (mean): 47.20%	F1 (max): 100.00%
-------------------	------------	-------------------	-------------------

----- Epoch 2 / 5 -----

Train Loss: 1.074

Dev Set Scores --	EM: 31.33%	F1 (mean): 50.77%	F1 (max): 100.00%
-------------------	------------	-------------------	-------------------

----- Epoch 3 / 5 -----

Train Loss: 0.688

Dev Set Scores --	EM: 31.49%	F1 (mean): 51.06%	F1 (max): 100.00%
-------------------	------------	-------------------	-------------------

----- Epoch 4 / 5 -----

Train Loss: 0.457

Dev Set Scores --	EM: 29.75%	F1 (mean): 48.81%	F1 (max): 100.00%
-------------------	------------	-------------------	-------------------

----- Epoch 5 / 5 -----

Train Loss: 0.331

Dev Set Scores --	EM: 29.66%	F1 (mean): 48.53%	F1 (max): 100.00%
-------------------	------------	-------------------	-------------------