280722_1053 - 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: 0.999
)
Training Data: 43863
Validation Data: 10570
_____

---------------------- Epoch 1 / 5 ----------------------
Train Loss: 1.728

Dev Set Scores --   EM: 31.64%   F1 (mean): 50.80%   F1 (max): 100.00%

---------------------- Epoch 2 / 5 ----------------------
Train Loss: 0.994

Dev Set Scores --   EM: 32.21%   F1 (mean): 51.85%   F1 (max): 100.00%

---------------------- Epoch 3 / 5 ----------------------
Train Loss: 0.744

Dev Set Scores --   EM: 31.70%   F1 (mean): 50.84%   F1 (max): 100.00%

---------------------- Epoch 4 / 5 ----------------------
Train Loss: 0.606

Dev Set Scores --   EM: 31.01%   F1 (mean): 50.80%   F1 (max): 100.00%

---------------------- Epoch 5 / 5 ----------------------
Train Loss: 0.539

Dev Set Scores --   EM: 29.81%   F1 (mean): 48.60%   F1 (max): 100.00%