Learning Curve: Batch Size 5, Dropout 0.5 Max_norm(2)