F1 Evaluation Over 10 Epochs

