Influence of the Choice of Dataset and Learning Rate (Training Set Size: 256) Mode: FUMBRELLA Mode: VANILLA 0.70 0.65 accuracy 09 0 0.55 Dataset: qnli, Learning Rate: 1e-05 Dataset: qnli, Learning Rate: 1e-05 0.50 Dataset: qnli, Learning Rate: 5e-05 Dataset: qnli, Learning Rate: 5e-05 1.8 1.6 1.4 1.0 8.0 0.6 Step Step 0.7 0.6 0.5 train_st2_loss 0 0 0 w + 0 0 0.2 0.1 0.0 200 600 700 100 700 0 100 300 400 500 800 200 300 400 500 600 800