Weight-Decay λ =0.01, Mini-Batch Size B=500, Training Rate η = 0.001 1.0 Training Rate () Logistic Regression Validation Data Accuracy Logistic Regression Training Data Accuracy Linear Regression Validation Data Accuracy Linear Regression Training Data Accuracy 0.9 Training Loss (Cross-Entropy Loss) 8.0 0.7 0.6 100 700 0 200 300 400 500 600

of Epochs