Training error vs Test error

Network characteristics
Number of Hidden Layers: 1
Units per Layer: 32
Activation function: Leaky ReLU
Batch-size: 32
Loss function: MEE
Learning Rate: 0.001
Regularization: elastic
α: 0.25
λ: 0.0001
Optimizer: NAG
Momentum: 0.9

Training Error
Test Error

MEE Error

Epochs