

Classification:
MNIST



Regression

NNs with a single Hidden Layer

100 hidden
neurons

+ L2 Regul.

50 hidden
neurons

→ more train.
Data = 10.000

Human Error?

Around 0%

↕ Avoidable Bias

Error on Training Set

2 %

1 %

3 %

1 %

0.1 %

↕ Variance

*Problem? 100 hidden neurons
= $(28 \times 28 + 1) \times 100 + (100 + 1) \times 10$ weights
that have to be learned!*

Error on Dev/Test Set

17 %

14 %

13 %

14 %

4 %