

---

meetings:  
wednesdays 14:00  
Thursd. -11-


14 weeks submit report: 5.1.

presentation max. 2-3 weeks after subm.:

- grading scheme:

- challenges taken by himself
- quality of docum.
- style of work

||

- motivation
  - creativity
  - enthusiasm
- 
- 

1) • experiments from paper: 2-layer FCNN, CNN  
on Gaussian/CIFAR-10 inputs + ● 1. week

• experiments on real data:

2-layer + MNIST: FCNN vs CNN

≈ 10<sup>2</sup> layers + CIFAR10: — // —

start with FCNN,  
use all training samples.  
What is test loss  $\epsilon$ ?  
→ how many tr. samples  
to

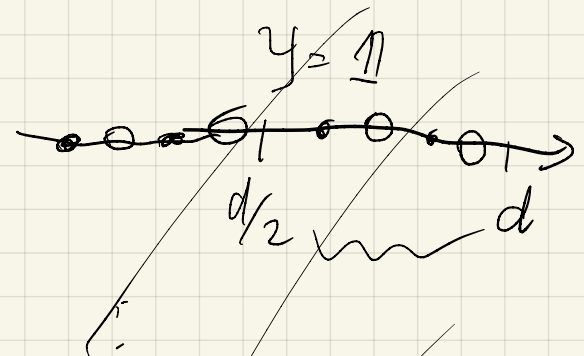
• experiments on synthetic data:

1D: labeling fct.:  $h(x) = \text{sgn}\left(\sum_{i=1}^{d/2} x_i - \sum_{i=d/2+1}^d x_i\right)$  (linear)

$h(x) = \text{sgn}\left(\sum_{i=1}^{d/2} x_i^2 - \sum_{i=d/2+1}^d x_i^2\right)$  (quadratic)

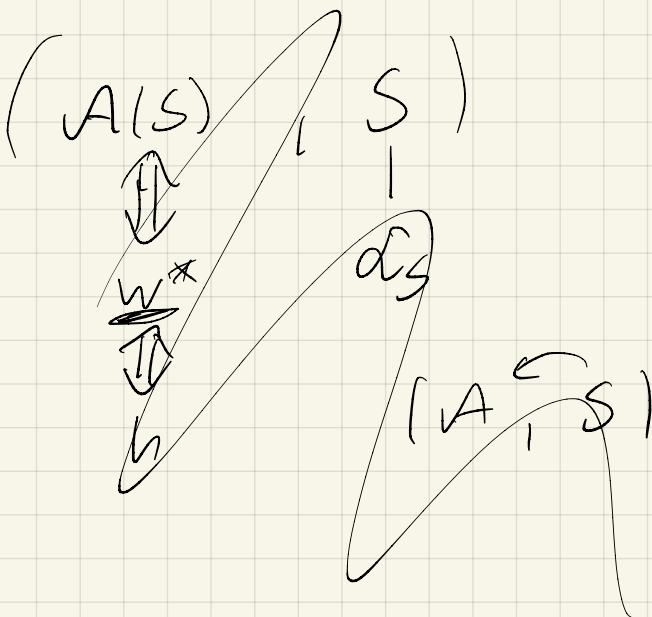
paper

$x = [0, 1]$



$L_S + C_{\text{sharp}}$

$\leq \underline{10}$



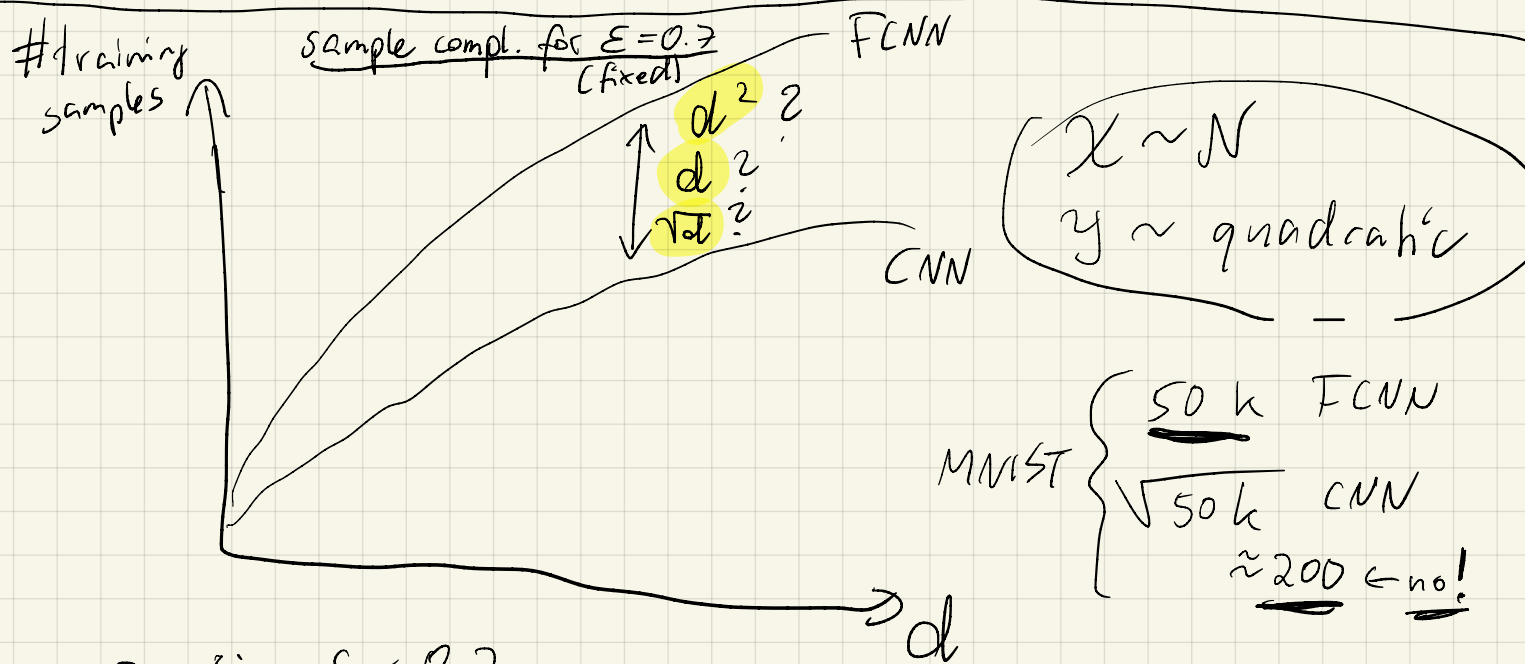
Paper: FCNN Lower bd.:  $\Omega(d^2)$   $d^{20}$   $e^d$   $\epsilon, \delta$

CNN upper bd.:  $\underbrace{O(1)}$  perhaps:  $O(d^{1.5})$ ? then separation:  $\sqrt{d}$

highly suspect... 50k tr. 10k test

MNIST FCNN 50k tr.  $\rightarrow$  70%  
CNN ? to samples  $\rightarrow$  70%

$d \hookrightarrow$  MNIST



0. fix  $\epsilon = 0.7$

iterate over  $d = 2 \dots 100$

1. create 30k samples { 25k training set, 5k test set
2. train FCNN on 15k (?)
3. get test loss  $\epsilon \approx 0.7$  FCNN
4. train CNN's till convergence with line search:
  - $\hookrightarrow$   $\begin{cases} 15k \rightarrow \epsilon_{CNN} \approx 0.9 \\ 5k \rightarrow \epsilon_{CNN} \approx 0.5 \end{cases}$ 
    - $\begin{cases} 14k \rightarrow 0.8 \\ 12k \rightarrow 0.7 \\ \epsilon_{CNN} = \epsilon_{FCNN} \end{cases}$

$\rightarrow$  pt. in above graph