

## 4 רשת - CNN - עצמית

### A רשת

הרשת תהיה בעלת 2 שכבות פילטריות ו-2 שכבות fully connected.  
 הפונקציה  $\text{ReLU}$  תהיה הפונקציה הפעילה.  
 $\eta = 0.001$  !  
 batch size = 64  
 Adam optimizer  
 הפונקציה  $\text{L2}$  תהיה הפונקציה הפעילה.  
 $\eta = 0.0001$   
 הפונקציה  $\text{L2}$  תהיה הפונקציה הפעילה.

Train Set ACC = 90%

Train Set avg loss = 0.2662

• validation-set ACC = 88%

validation-set avg loss = 0.3248

• Test Set ACC = 86%

Test set avg loss = 0.3683

Word Embedding  
 מילה  
 , פונקציה

~~מחשבת~~

### B רשת

הרשת תהיה בעלת 2 שכבות פילטריות (multi-class) ו-2 שכבות fully connected.  
 הפונקציה  $\text{ReLU}$  תהיה הפונקציה הפעילה.  
 $\eta = 0.001$   
 batch size = 64  
 Adam optimizer  
 הפונקציה  $\text{L2}$  תהיה הפונקציה הפעילה.  
 הפונקציה  $\text{L2}$  תהיה הפונקציה הפעילה.

⏏

: B  $\sqrt{311}$   $\sqrt{16311}$

• Test Acc = 90%.

Test avg loss = 0.2432

• Validation Acc = 89%

val Avg loss = 0.3001

• Test acc = 87%

Test avg loss = 0.3455

— — —  
| 87% |  
| 89% |  
| 90% |  
— — —

C  $\sqrt{311}$

( $\sqrt{311}$  in) NN ~~layer~~, input in 2100 in 2  
activation  $\rightarrow$  Relu  $\rightarrow$  hidden layers 21 in 2100  
net time in A  $\sqrt{311}$  in 2100  
 $\rightarrow$  some in, some to batch-norm  
Internal covariate shift in 2100  $\sqrt{311}$  in 2100

$\sqrt{16311}$

Train Acc = 90%

avg loss = 0.2376

Validation Acc = 88%

Avg loss = 0.3415

Test Acc = 87%

loss = 0.3575

— — —  
| 87% 89% |  
| 90% |  
— — —