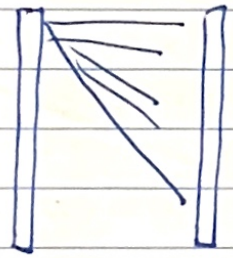
3-Step Framework

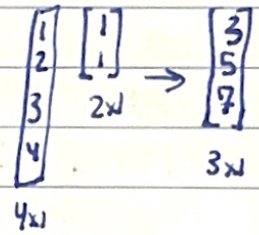
Model: Linear, Attention

Loss: Mean Squared, cross-entropy, contrastive

Optimizer: exact, gradient descent

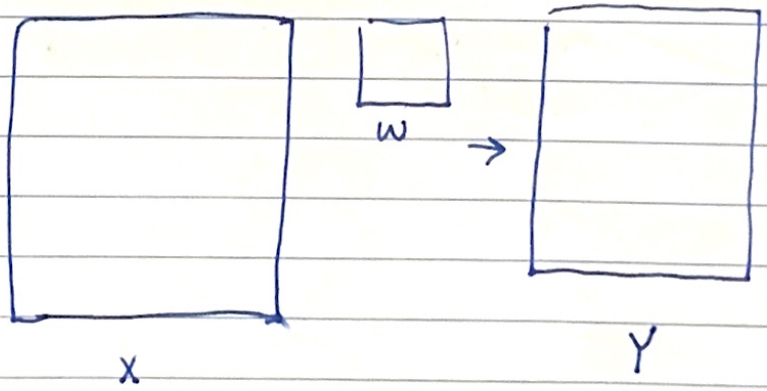
! Convolution is our next type & layer!





Complexity: $O(d^2)$ $O(d' \cdot \text{window-size})$

- Efficient
- shared weights
- Locality preserving



Activity: (in 2D)
↳ average
↳ detect vertical edge |
↳ can we do 3D?

Deeper = ___

Image Embeddings

Plan

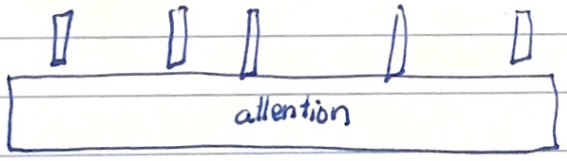
Review
Convolutions
Embeddings

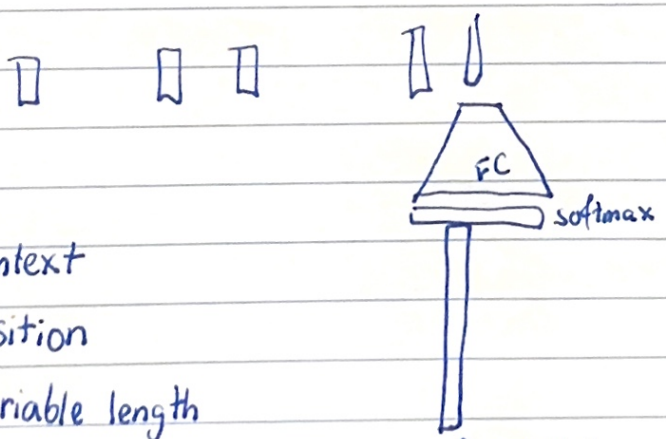
Logistics

Zoom :)
Game night!
Talk @ 1pm
PSet 2
Scribe Notes
Projects!

Review

"Roses are red violets are ____"





↑ distribution over next token

- context
- position
- variable length

Watermarking:
#1 Red/green
#2 Hash seed
#3 exponential minimum sampling

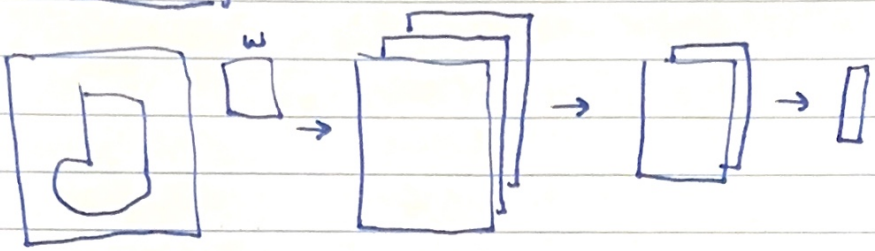
$i^* \sim p$

$\Leftrightarrow \Pr(i^* \text{ selected}) = p_i$

$\Leftrightarrow i^* = \arg\min_i \frac{-\log x_i}{p_i}$

for $x \sim \text{Unif}([0,1]^N)$

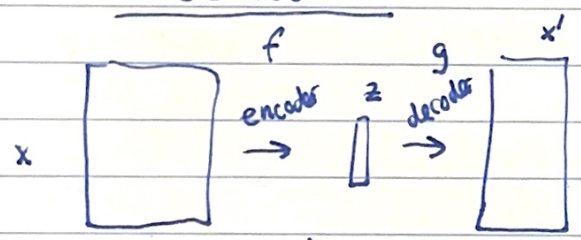
Embeddings



Contrastive learning: (x, x^+) close e.g., same clothing

(x, x^-) far e.g., different clothing

Autoencoders:

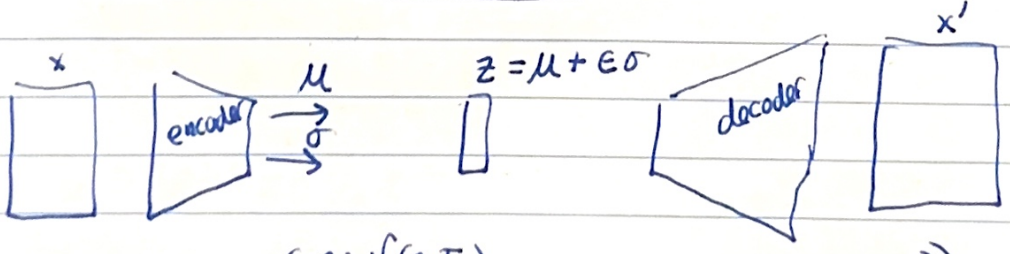


"condense information"

Reconstruction loss: $\|g(f(x)) - x\|_2^2$

Can we distribute nicely?

Variational Auto encoders



$\epsilon \sim N(0, I)$

Variational loss: $\underbrace{H(N(\mu,\sigma), N(0,I))}_{KL\ \text{divergence}} - H(N(0,I)) = -\log(\sigma) + \sigma^2 + \mu^2 - 1/2$