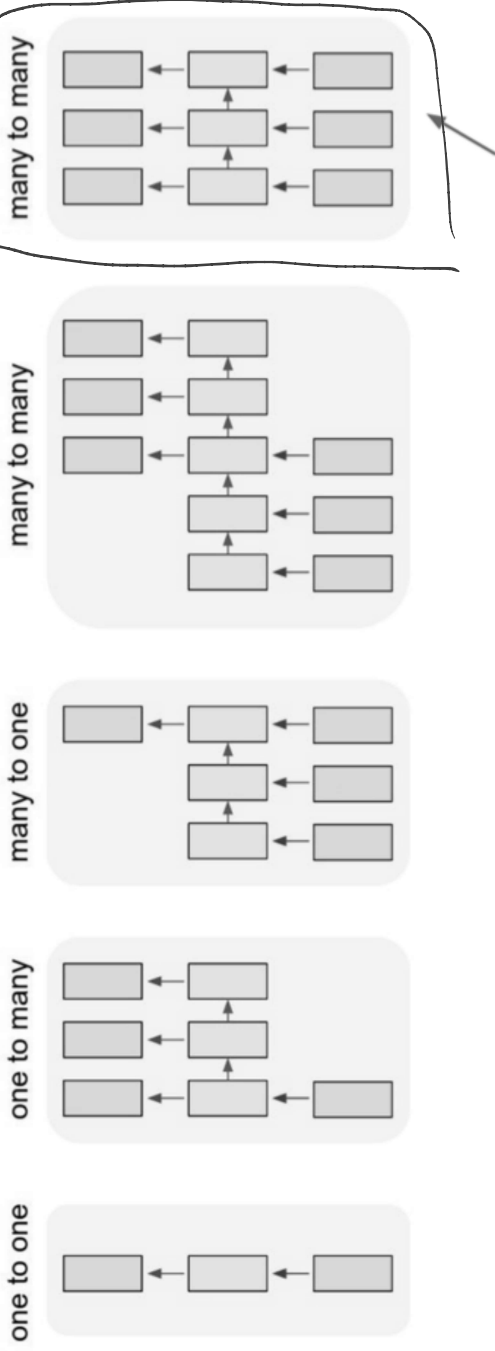
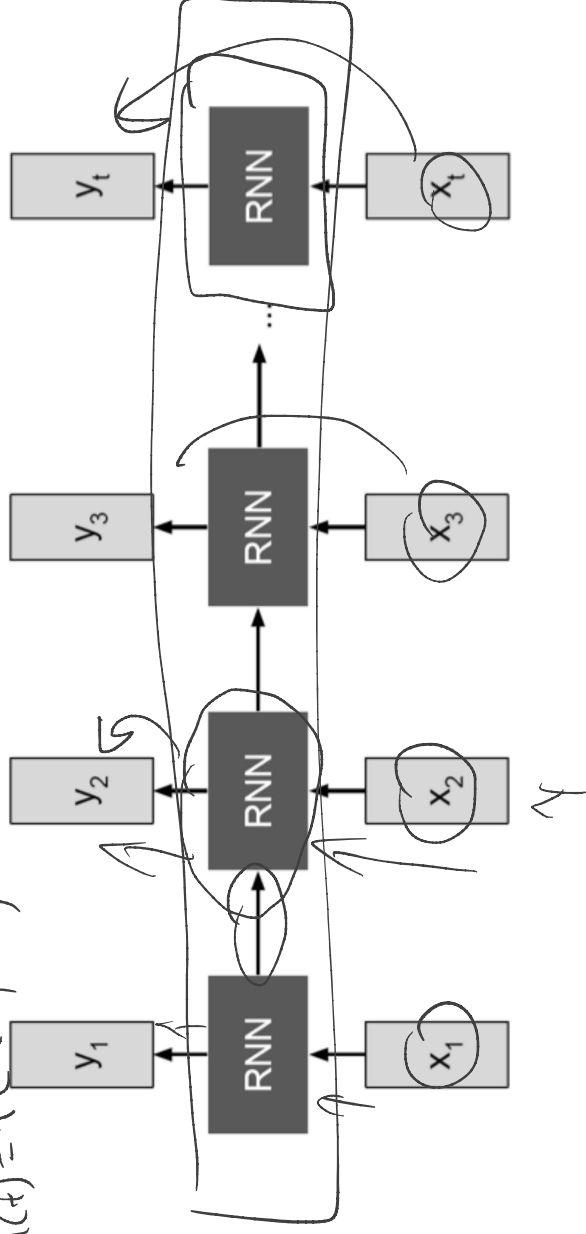


# Recurrent Neural Networks



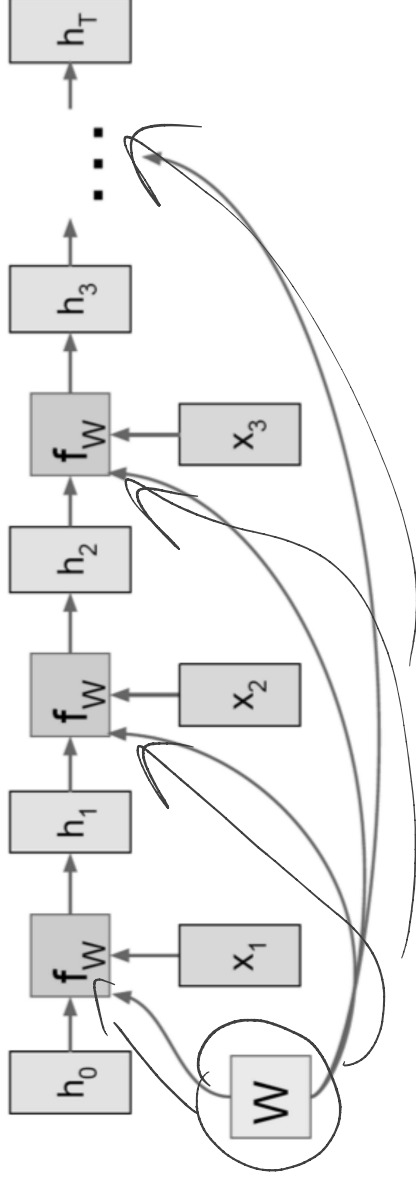
# Recurrent Neural Networks

$$h_t = f(h_{t-1}, x_t)$$

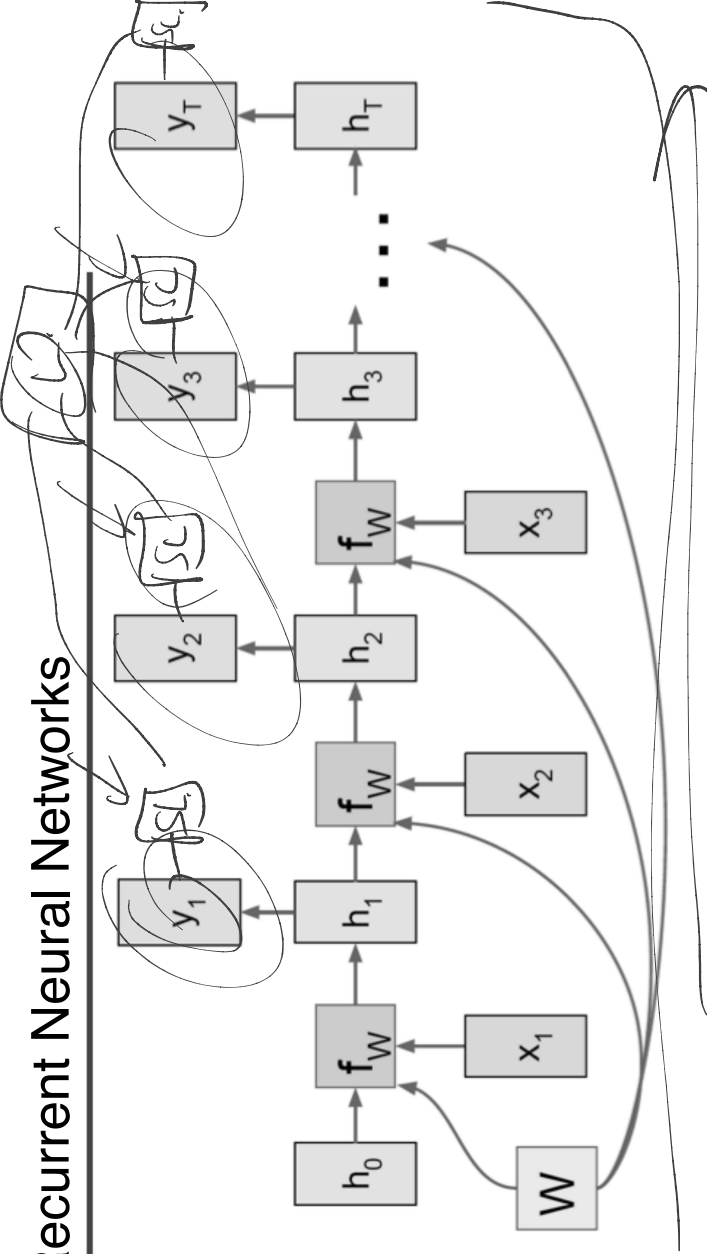


# Recurrent Neural Networks

Re-use the same weight matrix at every time-step

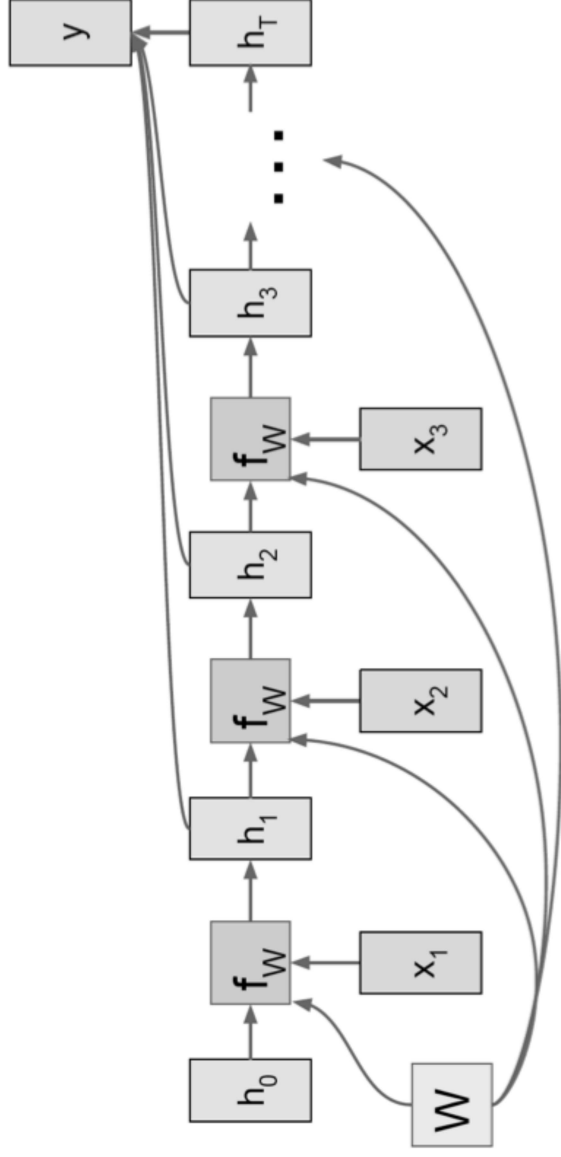


# Recurrent Neural Networks

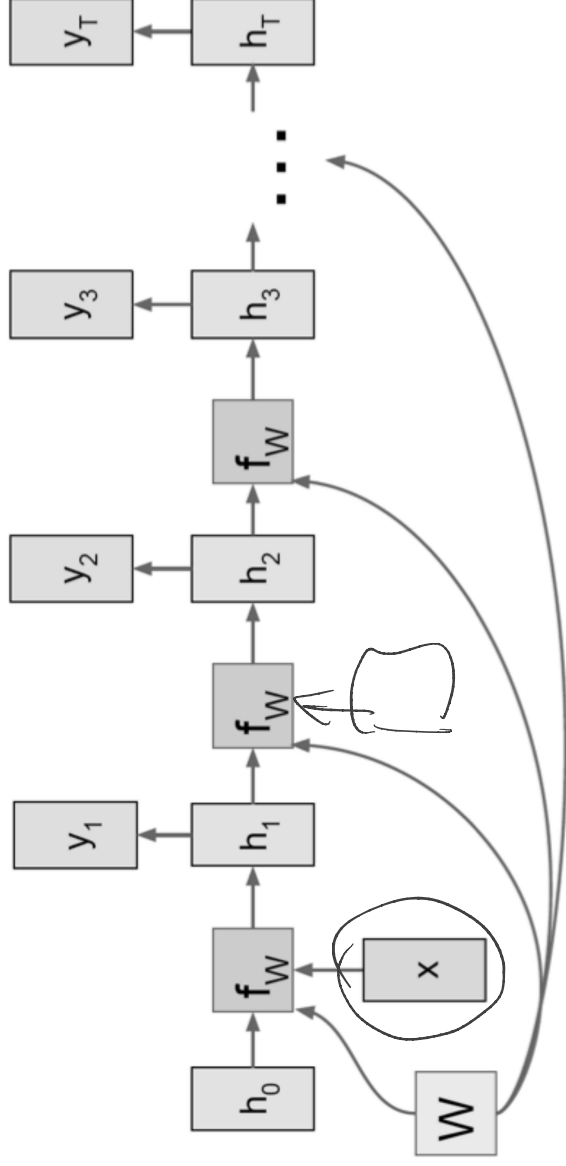


# Recurrent Neural Networks

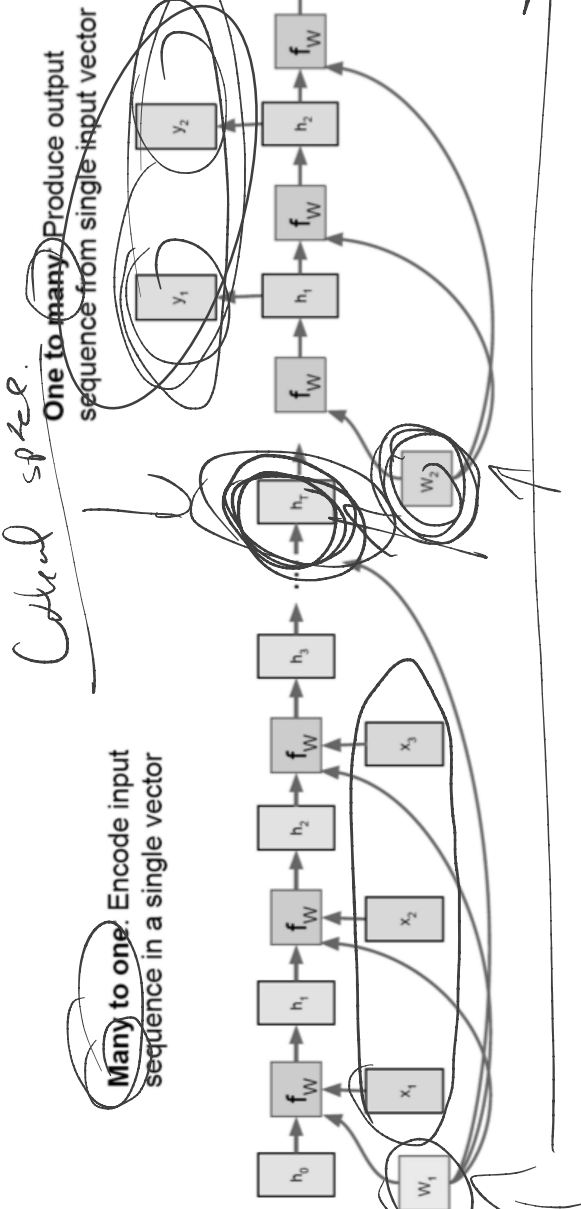
---



# Recurrent Neural Networks



## Recurrent Neural Networks

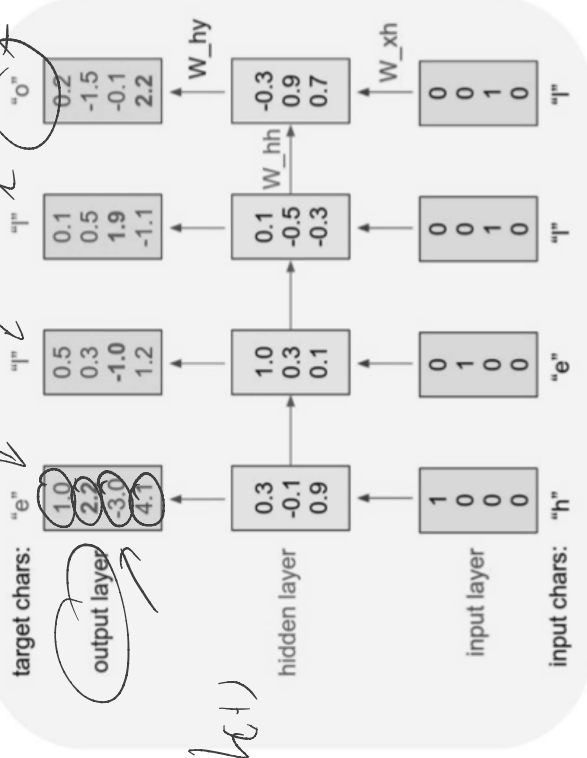


# Recurrent Neural Networks

## Example: Character-level Language Model

Vocabulary:  
[h,e,l,o]

Example training  
sequence:  
"hello"

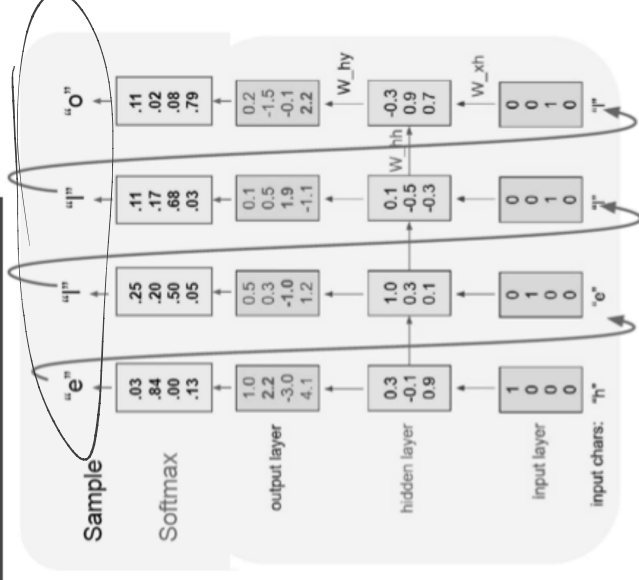




## Example: Character-level Language Model Sampling

Vocabulary:  
[h,e,l,o]

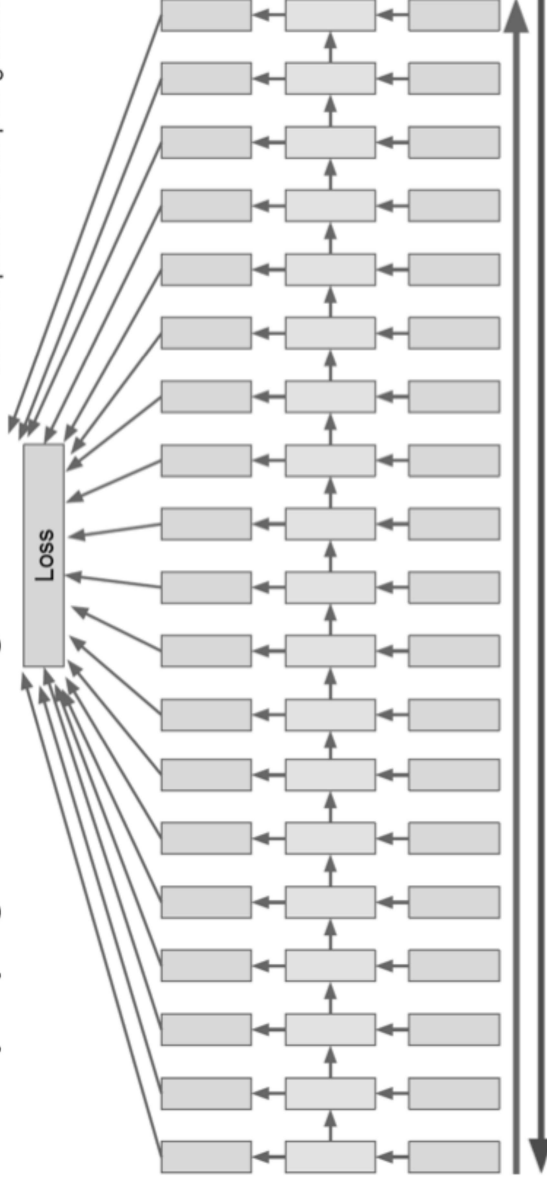
At test-time sample  
characters one at a time,  
feed back to model



# Recurrent Neural Networks

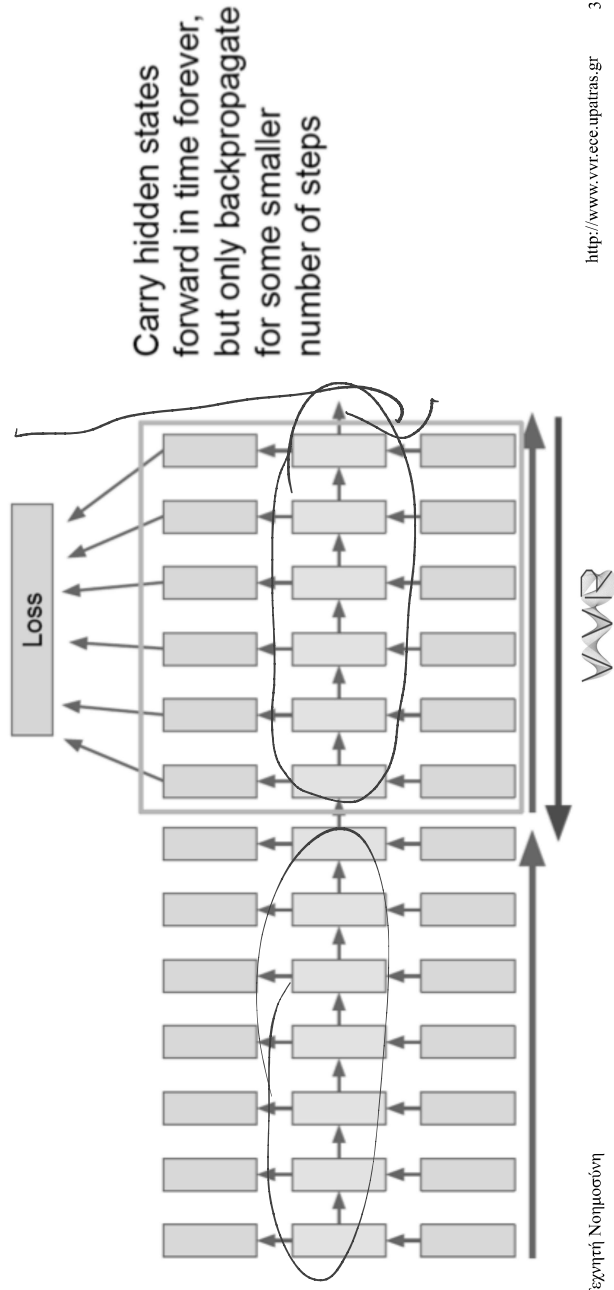
Forward through entire sequence to compute loss, then backward through entire sequence to compute gradient

Backpropagation through time



# Recurrent Neural Networks

## Truncated Backpropagation through time



# Recurrent Neural Networks

---

## RNN Advantages:

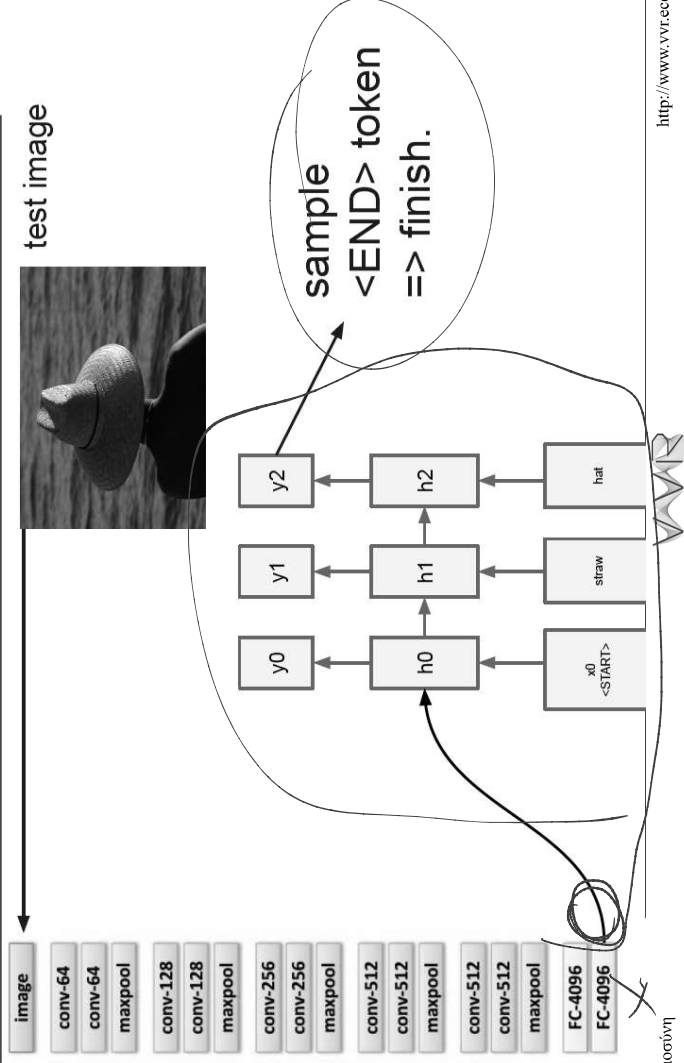
- Can process any length input
- Computation for step  $t$  can (in theory) use information from many steps back
- Model size doesn't increase for longer input
- Same weights applied on every timestep, so there is symmetry in how inputs are processed.

## RNN Disadvantages:

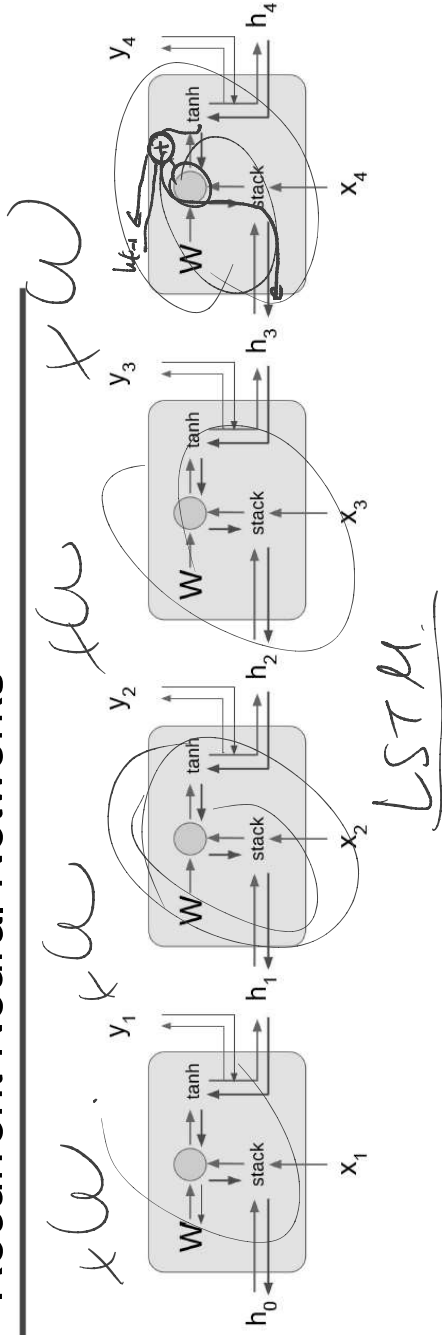
- Recurrent computation is slow
- In practice, difficult to access information from many steps back



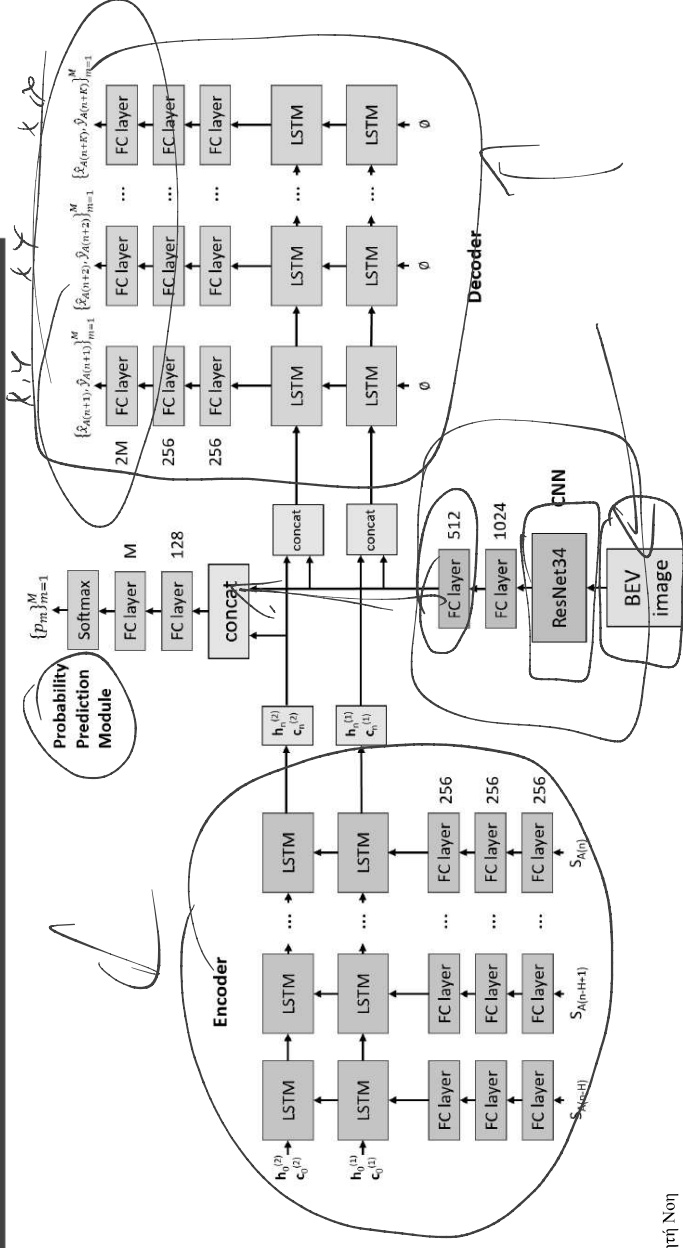
# Recurrent Neural Networks



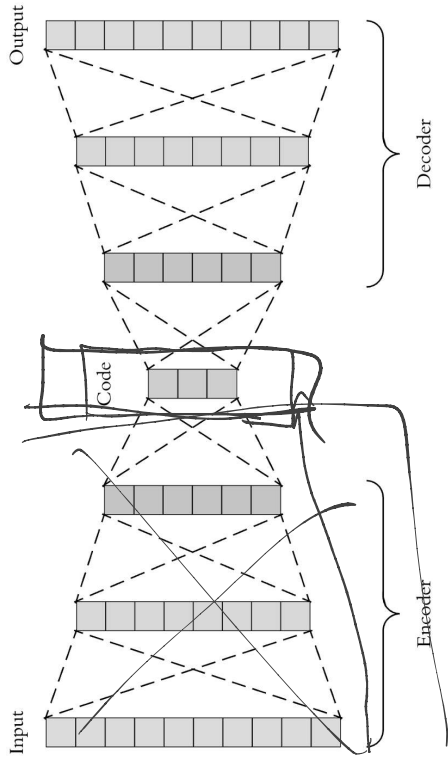
# Recurrent Neural Networks



# Examples – Motion prediction?

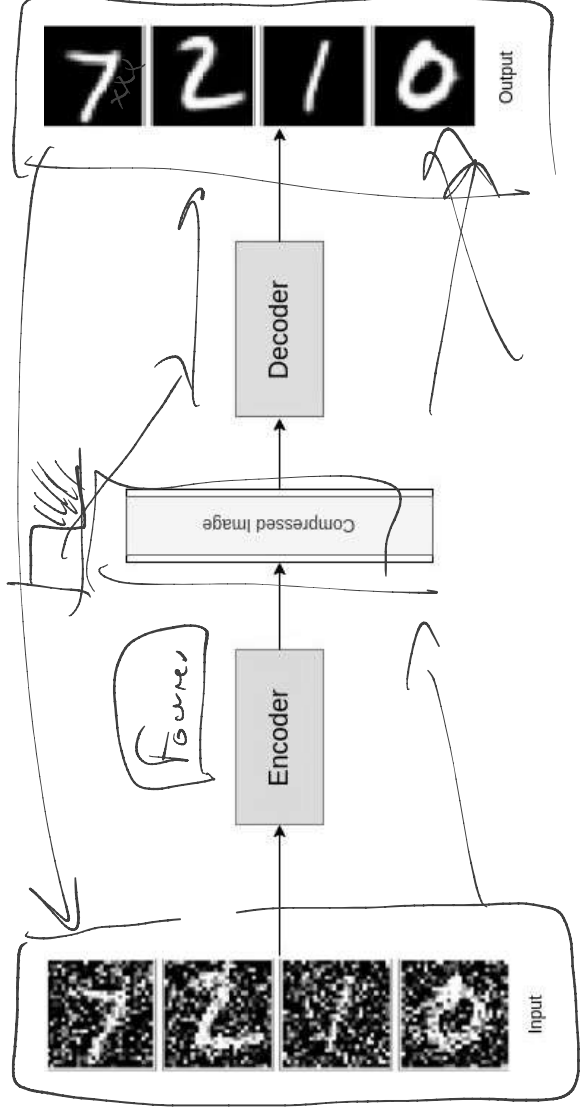


# Examples – Autoencoders





# Examples – Autoencoders & Denoising



# Examples – Autoencoders & Superresolution

