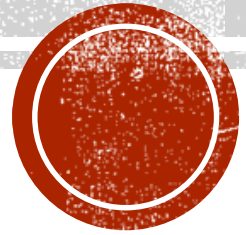
IU ST

NATURAL LANGUAGE PROCESSING WITH DEEP LEARNING



Dr.Minaei, IUST, Fall 2020

Presenting by Sara Rajaee

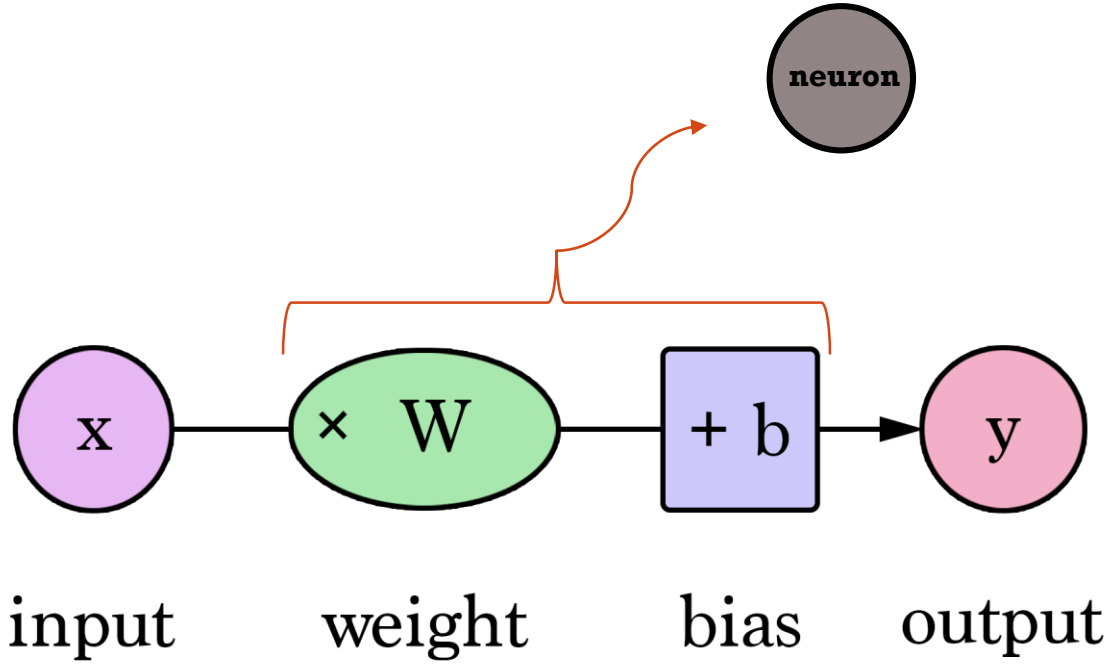
Most materials of this mini-course are provided by "Natural Language Processing" course, Mohammad Taher Pilehvar, IUST, Fall 2019 and "Natural Language Processing with Deep Learning" course, Stanford, Winter 2020

REVIEW – WORD2VEC

- A two-layer neural network learns a vector for every word in the vocabulary.
- The target task on which word2vec's model is trained is Language modeling.
- Word2Vec has two architecture to learn word representations, CBOW and Skip-gram.
- Both of them consider a fixed-size window of neighbors.

REVIEW - NEURAL NETWORKS

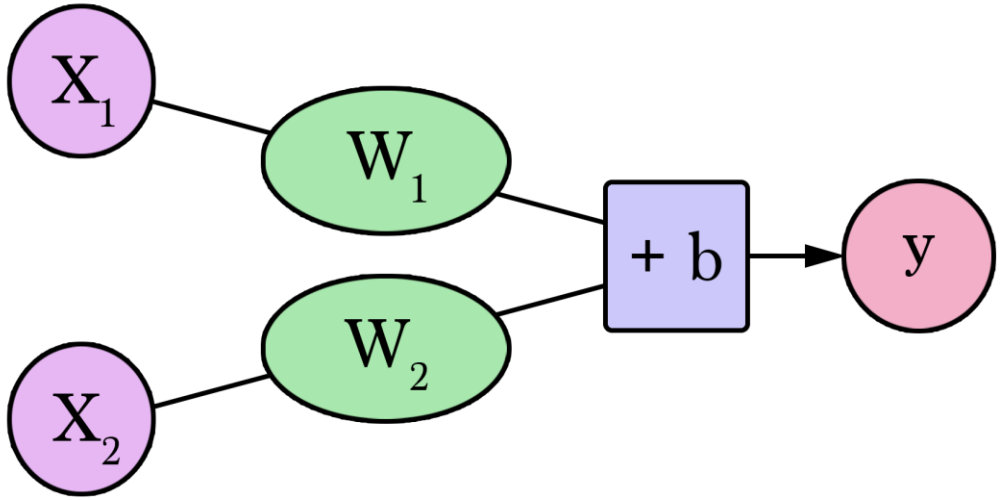
A NEURON



$$y = Wx + b$$

REVIEW - NEURAL NETWORKS

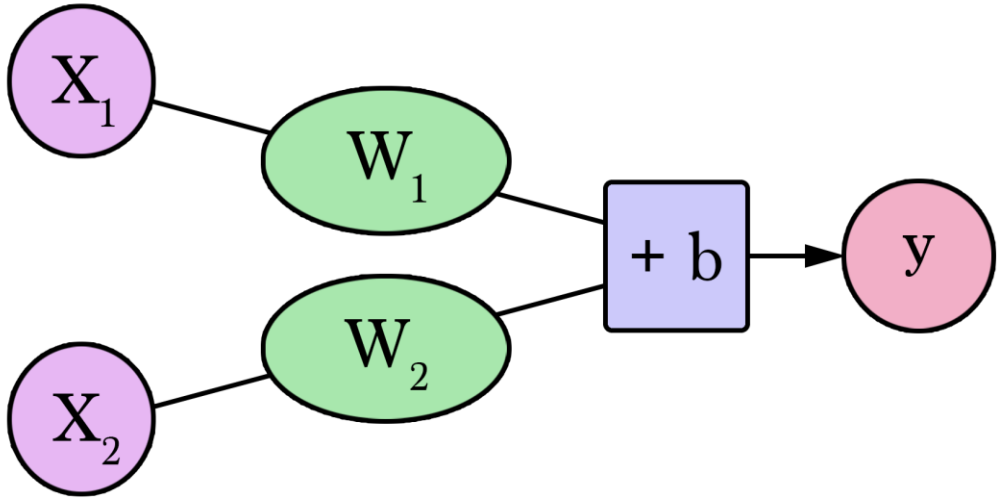
A NEURON



$$y = W_1 x_1 + W_2 x_2 + b$$

REVIEW - NEURAL NETWORKS

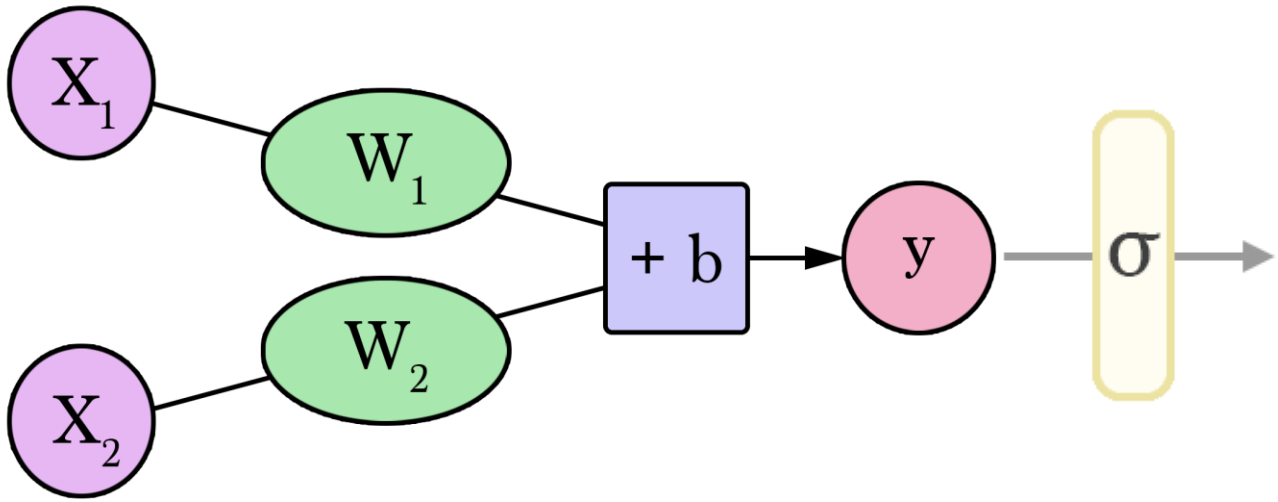
A NEURON



$$y = W_1 x_1 + W_2 x_2 + b$$

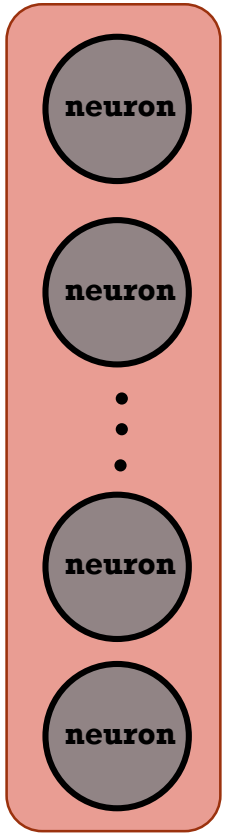
REVIEW - NEURAL NETWORKS

A NEURON



REVIEW - NEURAL NETWORKS
A DENSE LAYER

- A basic layer in neural networks.
- Add nonlinearity to neural nets.
- Feeds all input features to all neurons, and each neuron provides an output for the next layer in the model.



MEMORY

- Dense has no memory!
  - It processes the input shown to it independently (with no state in between them)
- This is called *feedforward networks*

- However, we read sentences word by word, keeping a memory of what came before

MEMORY

- Dense has no memory!
  - It processes the input shown to it independently (with no state in between them)
- This is called *feedforward networks*

- However, we read sentences word by word, keeping a memory of what came before

RECURRENT NEURAL NETWORKS

- In Dense, each input is processed independently, with no state kept in between inputs:
  - To process a sequence or a temporal series of data points, you have to show the entire sequence to the network at once: turn it into a single data point.
- In contrast, human processes a text word by word, while keeping memories of what came before.
  - Biological intelligence processes information incrementally while maintaining an internal model of what it's processing, built from past information and constantly updated as new information comes in.

RECURRENT NEURAL NETWORKS

- RNN adopts the same principle, albeit in an extremely simplified version:
  - It processes sequences by iterating through the sequence elements and maintaining a state containing information relative to what it has seen so far.
- The state of the RNN is reset between processing two different, independent sequences.
  - We can still consider one sequence a single data point: a single input to the network.

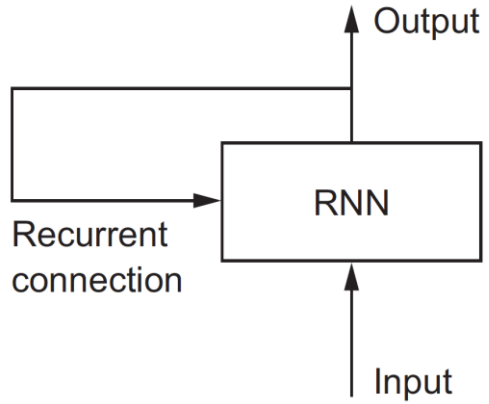


Figure 6.9 A recurrent network: a network with a loop

RECURRENT NEURAL NETWORKS

- To make it clear, let's implement the forward pass of a toy RNN in NumPy.

```
state_t = 0                          <---- The state at t
for input_t in input_sequence:       <---- Iterates over sequence elements
    output_t = f(input_t, state_t)
    state_t = output_t               <---- The previous output becomes the state for the next iteration.
```

REFERENCES AND FURTHER RESOURCES

Websites:

1. https://jalammar.github.io/visual-interactive-guide-basics-neural-networks/