AI Society Meeting Spring 2020

Basics of Deep Learning

Curriculum

- Week 7 -› Basics of Deep Learning
- Week 8 -› Basics of Natural Language Processing
- Week 9 -› Long Short Term Memory/Recurrent Neural Networks

Outline

1. Artificial Neural Networks (ANNs) and Deep Learning
2. Layers
3. Important Terms
4. Different Neural Network Architectures
5. Applications
6. Exercises

Artificial Neural Networks (ANNs) and Deep Learning

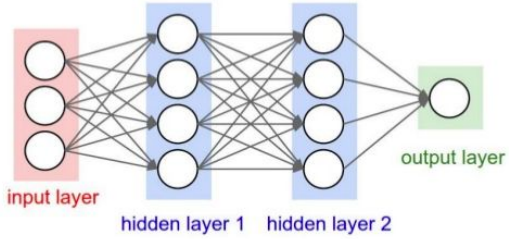
ANNs

- ANNs have units of calculation called **neurons (**connected by weighted values).
- What is a neuron? Consists of an activation function + weighted connections
- Long story short:
    - neuron will perform some sort of a calculation, possibly through an activation function, and then the result of this calculation will be multiplied by a weight as it passes through.
    - The weighted result can sometimes be the output of your neural network or you can have more neurons configured in **layers,** which is the basic concept to **deep learning.**
- When we ask the system a large number of questions - that's how the neurons learn how to carry out any specific calculations (recall supervised learning) and provide them with model/expected answers.
- Keep in mind, **Deep Learning** always has **more than 1 hidden layer** (you can jot down the dots once you reach the slides on Layers).
- **Deep learning** is a branch of Machine Learning which uses different types of neural network architectures. These algorithms mimic the way our brain functions.

Layers

Layers

- Layers are a set of neurons which help us determine the number of neurons we can use.
- The 3 different kinds:
  - Input: Data provided to the neural network
  - Hidden: Depends on your architecture and processes **activation functions**
  - Output: Provide us with the results/finished computations
- This hierarchy helps the output layer with predictions.
- Each layer extracts higher features from the input layer until the final layer makes a decision about what the input shows.
- More layers = higher level features can be learned



Important Terms

Important terms

- Linear Algebra behind DL
- Activation function
  - Neurons detect features and transform them through these non-linear activation functions before passing them onto the output layer; needed to learn complex/non-linear representations of data. Example: Sigmoid function.
- Neural networks learn weights and biases using the gradient descent algorithm -- how are the gradients of the cost functions computed or really, how does the neural network learn? **Backpropagation**! More on math behind backpropagation here.

Important terms (continued)

- More on backpropagation:
  - Used to find the weights/biases that lead to the most accurate output for the NN, where the weights are initialized randomly at most times.
  - These weights and biases keep adjusting depending on what the model learns.
  - Gradient descent is then used to come up with a minimized cost function (talked about in our previous sessions) using these weights.
  - The “training” of the model comprises of working with more and more data to give you a neural network with a high accuracy rate.
    - Output from the neural network -› calculate cost -› backpropagate to tune weights
  - The most important ingredient? Data. If you don’t have clean data, your model (no matter how optimal it is) can only help you so much.

Different Neural Network Architectures

Different Neural Network Architectures

- Recurrent Neural Networks
  - Help in understanding context based on previous behavior (used in Natural Language Processing) and in retaining information (example: LSTM) by going back through the process and building on top of it
- Convolutional Neural Networks
  - Help in reducing the number of parameters that need to be optimized
  - Notes symmetry in how the neurons are connected; helps in recognizing patterns
  - Commonly used in working with images by recognizing patterns around pixels
  - More on CNNs
- Feel free to look more into these
  - Feedforward Neural Network
  - Radial Basis Function Neural Network
  - Multilayer perceptron
  - Modular Neural Networks

Applications

- Speech Recognition (NLP)
- Image Recognition (computer vision, try ECS 174 to learn more)
- Ads, web search
- Notable people who have contributed significantly: Geoffrey Hinton, Andrew Ng, Yann LeCun, Ian Goodfellow (one of the authors of the deep learning book), Jeff Dean

Exercises

- Tinker with a Neural Network here
- Try this Neural Network exercise here
- Try this Sentiment Analysis exercise by following this article

Thank you!
Next week: Basics of Natural Language Processing