Week 1 - Introduction to Deep Learning

Introduction

- AI is the new Electricity
- Electricity had once transformed countless industries and so is AI nowadays

Course Structure:

1. Neural Networks and Deep Learning
2. Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization
3. Structuring your Machine Learning Project
4. Convolutional Neural Networks (CNN)
5. Natural Language Processing: Building sequence models (RNN, LSTM, GRU)

What is a neural network?

House pricing prediction:

- predict the price of house based on its size. Single neuron. ReLU
- predict on multiple features given an input **x**
- given *enough* data a neural network is good enough to map the input x to a given output y

Supervised Learning with Neural Networks

In supervised learning you have a given input (x) and a desired given output (y) labeled.

| Input (x)	Output (y)	Application
Home Features	Price	Real State
Ad, user info	Click on ad? (0/1)	Online Advertising
Image	Object (1, …, 1000)	Photo Tagging
Audio	Text Transcript	Speech Recognition
English text	Chinese text	Machine Translation
Image, Radar Info	Position of other cars	Autonomous driving

Image applications usually uses CNN, sequence data (audio and text) usually uses RNN. Complex models usually mix many strategies. Simple models usually are simply fully connected layers.

Structured Data: data in a tabular format with a very well defined meaning **Unstructured Data**: data in no particular format but with underling meaning as images, text or audio

Why is deep learning taking off?

The basic ideas of deep learning have been around for decades but why the sudden growth on its popularity?

Scale drives deep learning progress
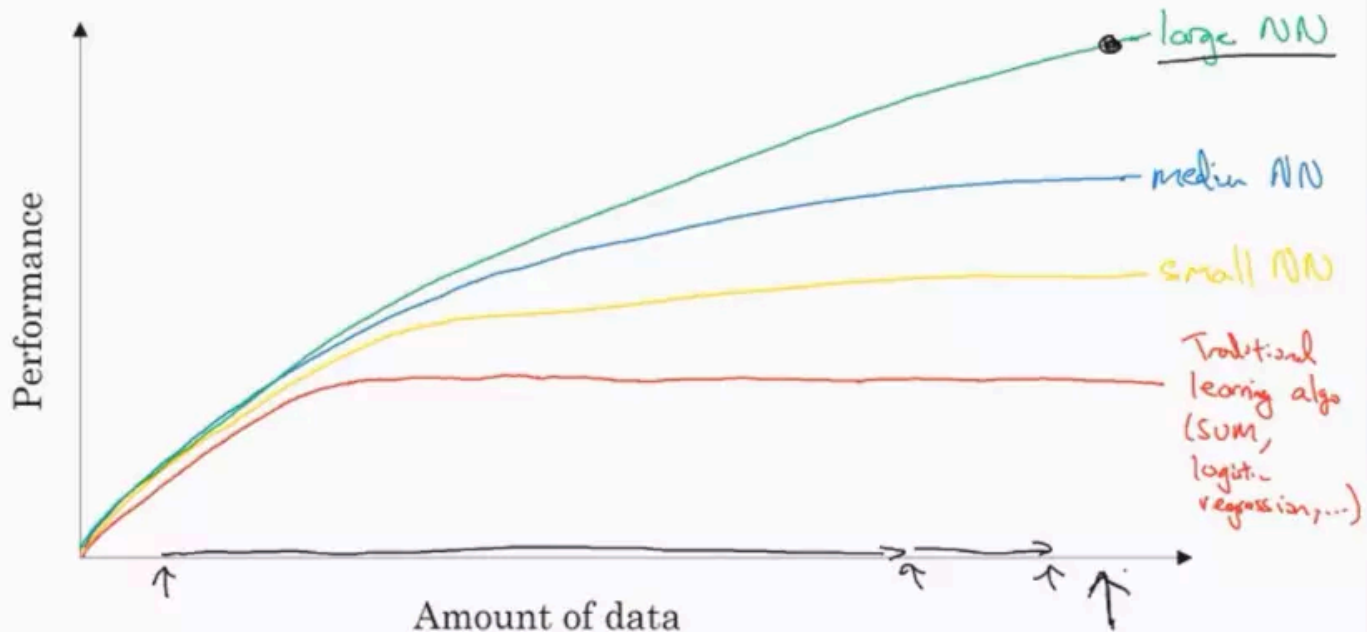
Performance x Amount of Data on different models

- Increase of data availability and collected through all means and available devices
- Performance plateaus in old algorithms
- NN increases performances with data
- GPU and computational power availability (either local and cloud sources)
- Algorithms changes. Activation functions (Sigmoid vs ReLU), Optimizers (SGD vs Adam)
- Machine Learning a iterative cycle. Idea -> Code -> Experiment. As the training got much faster the cycle could be iterated much faster.

Quiz

1. Question 1: What does the analogy "AI is the new electricity" refer to? (1 point)
   - ☐ Through the "smart grid", AI is delivering a new wave of electricity.
   - ☐ AI is powering personal devices in our homes and offices, similar to electricity.
   - ☑ Similar to electricity starting about 100 years ago, AI is transforming multiple industries.
   - ☐ AI runs on computers and is thus powered by electricity, but it is letting computers do things not possible before.
2. Question 2: Which of these are reasons for Deep Learning recently taking off? (Check the three options that apply.)
   - ☐ Neural Networks are a brand new field.
   - ☑ We have access to a lot more data.
   - ☑ We have access to a lot more computational power.
   - ☑ Deep learning has resulted in significant improvements in important applications such as online advertising, speech recognition, and image recognition.
3. Question 3: Recall this diagram of iterating over different ML ideas. Which of the statements below are true? (Check all that apply.) IDEA->CODE->EXPERIMENT
   - ☑ Being able to try out ideas quickly allows deep learning engineers to iterate more quickly.
   - ☑ Faster computation can help speed up how long a team takes to iterate to a good idea.
   - ☐ It is faster to train on a big dataset than a small dataset.

- ☑ Recent progress in deep learning algorithms has allowed us to train good models faster (even without changing the CPU/GPU hardware).

4. Question 4: When an experienced deep learning engineer works on a new problem, they can usually use insight from previous problems to train a good model on the first try, without needing to iterate multiple times through different models. True/False?
   - ☐ True
   - ☑ False

5. Question 5: Which one of these plots represents a ReLU activation function?
   - ☐ Figure 1 (Sigmoid)
   - ☐ Figure 2 (Sigmoid + y)
   - ☑ Figure 3 (ReLU)
   - ☐ Figure 4 (Leaked ReLU)

6. Question 6: Images for cat recognition is an example of "structured" data, because it is represented as a structured array in a computer. True/False?
   - ☐ True
   - ☑ False

7. Question 7: A demographic dataset with statistics on different cities' population, GDP per capita, economic growth is an example of "unstructured" data because it contains data coming from different sources. True/False?
   - ☐ True
   - ☑ False

8. Question 8: Why is an RNN (Recurrent Neural Network) used for machine translation, say translating English to French? (Check all that apply.)
   - ☑ It can be trained as a supervised learning problem.
   - ☐ It is strictly more powerful than a Convolutional Neural Network (CNN).
   - ☑ It is applicable when the input/output is a sequence (e.g., a sequence of words).
   - ☐ RNNs represent the recurrent process of Idea->Code->Experiment->Idea->….

9. Question 9: In this diagram which we hand-drew in lecture, what do the horizontal axis (x-axis) and vertical axis (y-axis) represent?
   - ☐ x-axis is the amount of data and y-axis is the size of the model you train.
   - ☐ x-axis is the input to the algorithm and y-axis is outputs.
   - ☐ x-axis is the performance of the algorithm and y-axis (vertical axis) is the amount of data.
   - ☑ x-axis is the amount of data and y-axis (vertical axis) is the performance of the algorithm.

10. Question 10: Assuming the trends described in the previous question's figure are accurate (and hoping you got the axis labels right), which of the following are true? (Check all that apply.)
    - ☑ Increasing the size of a neural network generally does not hurt an algorithm's performance, and it may help significantly.
    - ☐ Decreasing the size of a neural network generally does not hurt an algorithm's performance, and it may help significantly.
    - ☑ Increasing the training set size generally does not hurt an algorithm's performance, and it may help significantly.
    - ☐ Decreasing the training set size generally does not hurt an algorithm's performance, and it may help significantly.