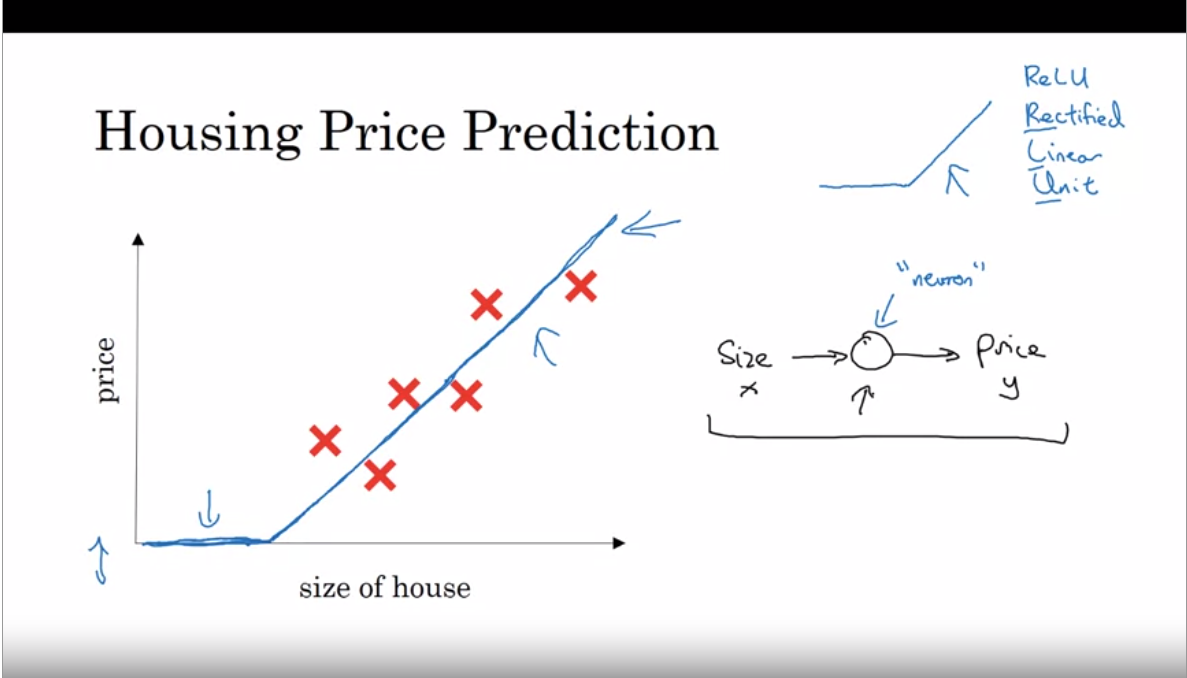
Q. What is deep learning and how is it powering the future?

It is a ML based on artificial neural networks using multi-processing layers used to extract features from data. Refers to training neural networks.

Eg. Self-driving cars

Just how electricity transformed our society AI is going to do that too and deep learning is a part of that.

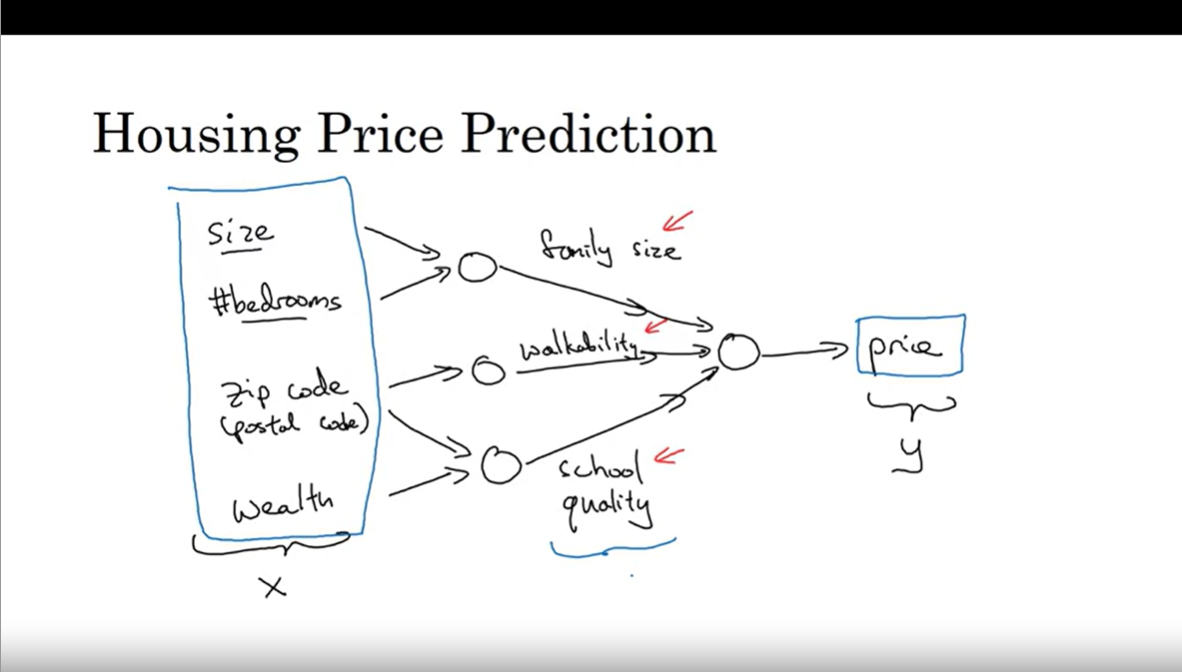
Linear regression is a statistical method used for modeling the relationship between a dependent variable (usually denoted as "Y") and one or more independent variables (usually denoted as "X"). Its primary goal is to find a linear equation that best represents this relationship.



Neuron--Inputs size computes size, takes max as zero and outputs price

ReLu—Rectified Linear Unit

Many neurons🡪Large Neural Network



One must train the neural network giving input x and output y

Hidden layers take all inputs and compute results.

*In supervised learning, the algorithm learns to make predictions or decisions based on labeled training data, where the correct answers are provided.*

Based on user data and searches ads are predicted to the user. Neural networks show the most relevant ads based on user preference and show you an ad that you are most likely to click on.

Computer vison uses neural networks for photo tagging. (CNN)

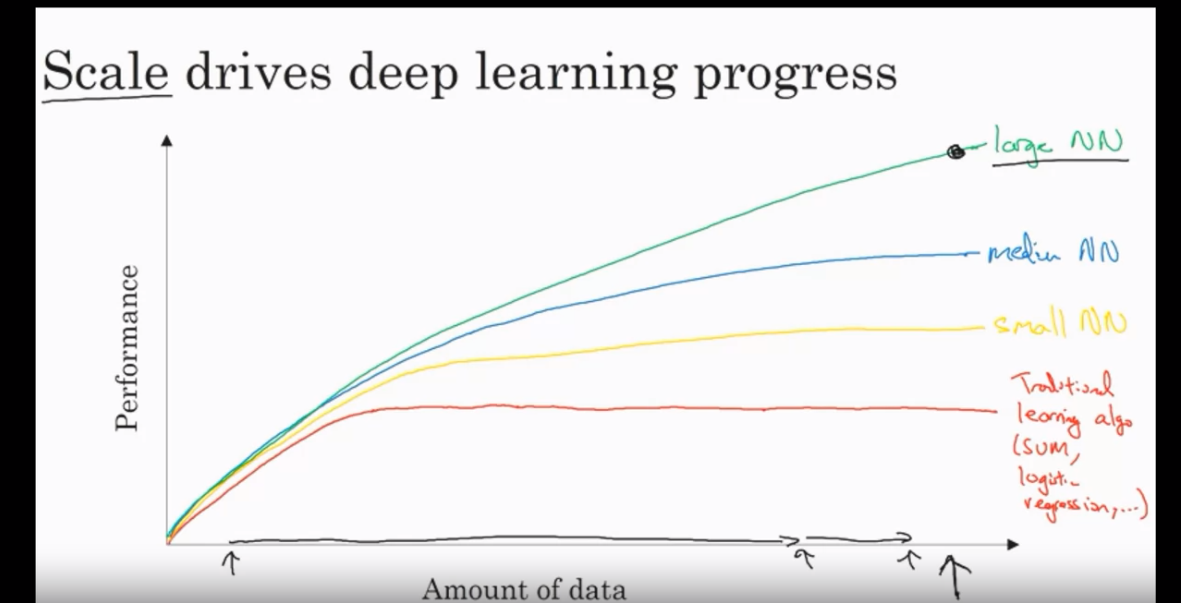
Speech recognition—input speech to text conversion. (RNN)(for sequence data)

Language translation is another application. (RNN’s and its complex versions)

Autonomous driving—helps predict position of other cars on road using image fed and radar info. (Customized neural networks)

Structured Data— Tabular form—Features have well defined meaning---used by companies to process giant database

Unstructured Data--- audio, image



The first curve plateaus out because of availability of small amount of data only.

Due to digitization of society we now have a large amount of data because of humans spending time on computers, websites, apps and because of sensors.

Performance keeps getting better as the size of the net increases because of training large amount of data.

Scale refers to—

* Scale of data
* Size of neural network with a lot of parameters, hidden units

m—number of training examples

sigmoid---slope 0

learning slow

parameters change slowly

ReLU—

gradient=1, it outputs the input directly if it's positive, and zero otherwise.

Gradient descent works well

Avoiding Vanishing Gradients: The sigmoid function squashes input values into the range (0, 1), causing gradients to become very small (close to zero) for inputs far from zero. This leads to the vanishing gradient problem, making it challenging for deep networks to train. ReLU, on the other hand, doesn't suffer from this problem because its gradient is either zero (for negative inputs) or one (for positive inputs), making it less likely to encounter vanishing gradients.

Helped computation(faster)

Idea—neural network

Code-implement and experiment

Make changes