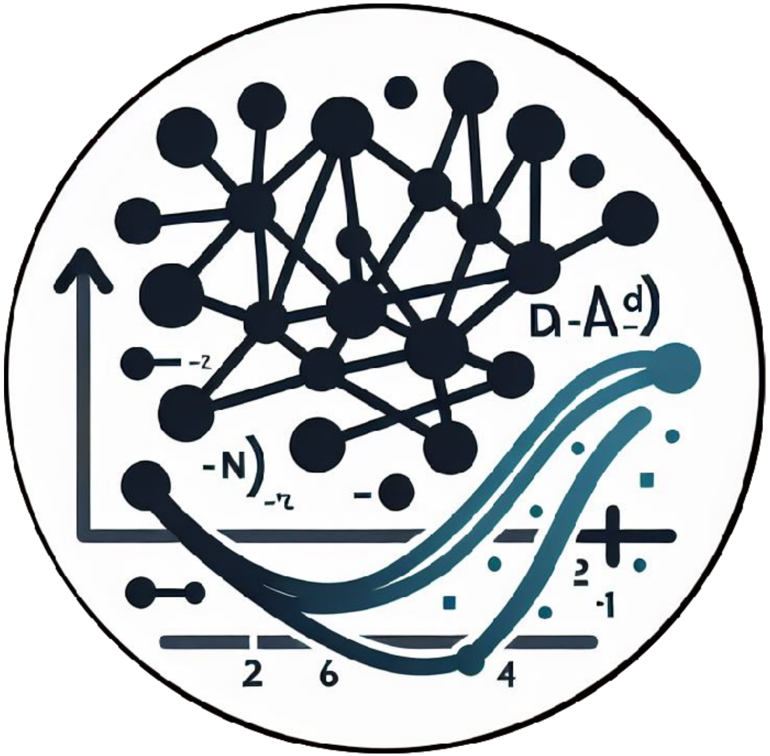
Deep Learning Basics



Deep Learning for Engineers
Andrew Ning
aning@byu.edu

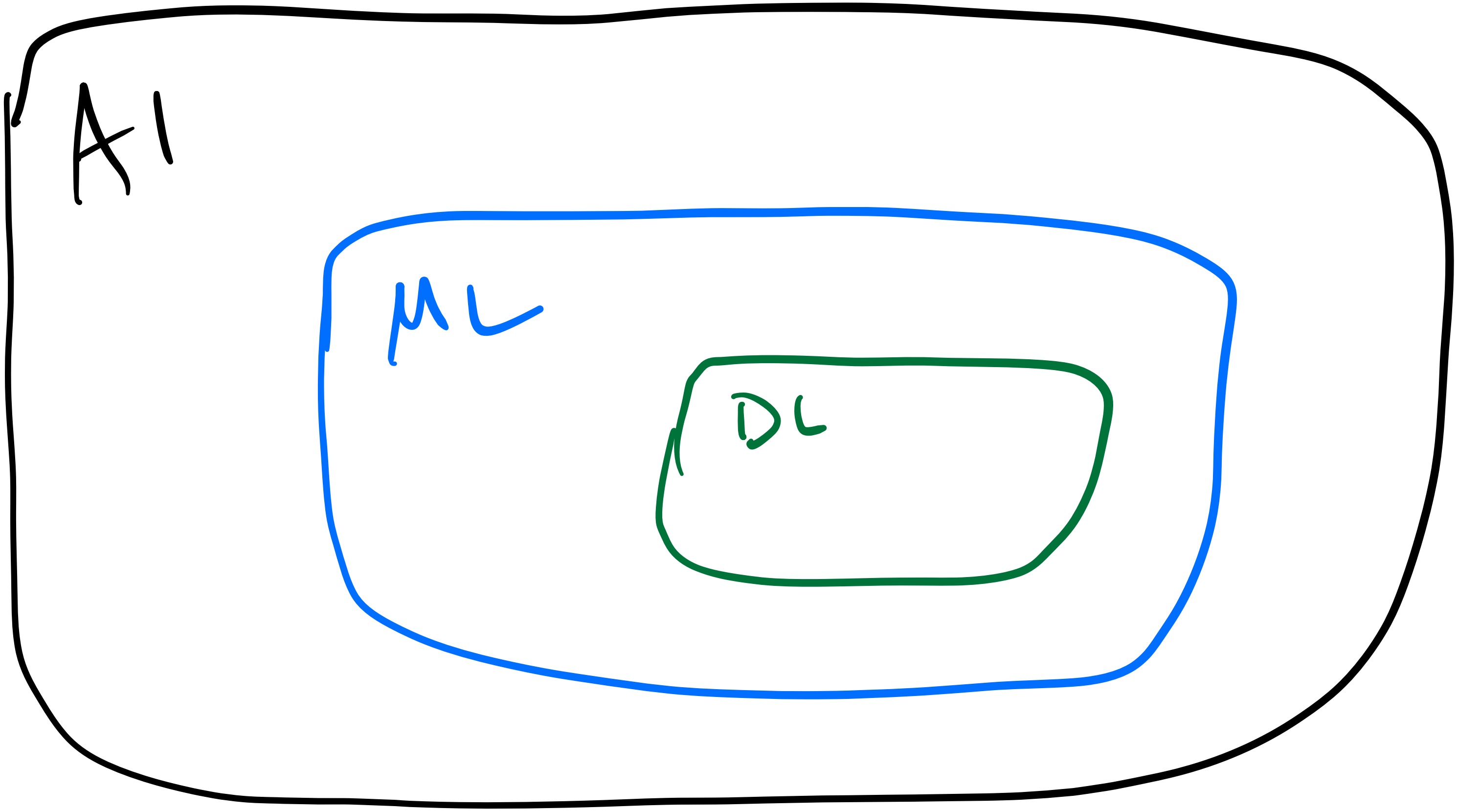
2006, BS Applied Physics

2008 MS, 2011 PhD, Aeronautics & Astronautics

2011-2014, Wind Energy

2014-present, Mechanical Engineering

What is Deep Learning?



Machine Learning

supervised - has labels.

regression
classification

unsupervised - no labels.

reinforcement. - learn by interacting w/ environment.

Intersection of Deep Learning with Engineering

2017: physics informed neural networks

2017: deep Koopman

2018: neural ODEs

2020: DL with symbolic regression

2021: neural operators

2024: Kolmogorov-Arnold Networks

Syllabus

https://flow.byu.edu/me595r/

Undergrad alternatives

EC EN 471: Machine Learning: Foundations and Applications

CH EN 426: Machine Learning for Engineers

CS 474: Introduction to Deep Learning

Universal Function Approximator

$$x_i \longrightarrow \boxed{\hat{f}(x)} \longrightarrow \hat{y}_i$$

Start with Linear Regression

$$y = \theta_1 x_1^2 + \theta_2 x_1 x_2 + \theta_3 \exp(x_2) + \theta_4 x_1 + \theta_5$$

$$y = \theta^T x \qquad x = [x_1^2,\ x_1 x_2,\ e^{x_2},\ \ldots]$$

$$y = w^T x + b$$

weights   bias

$$\underset{n_y}{y} = \underset{n_y \times n_x}{W} \underset{n_x}{x} + \underset{n_y}{b}$$

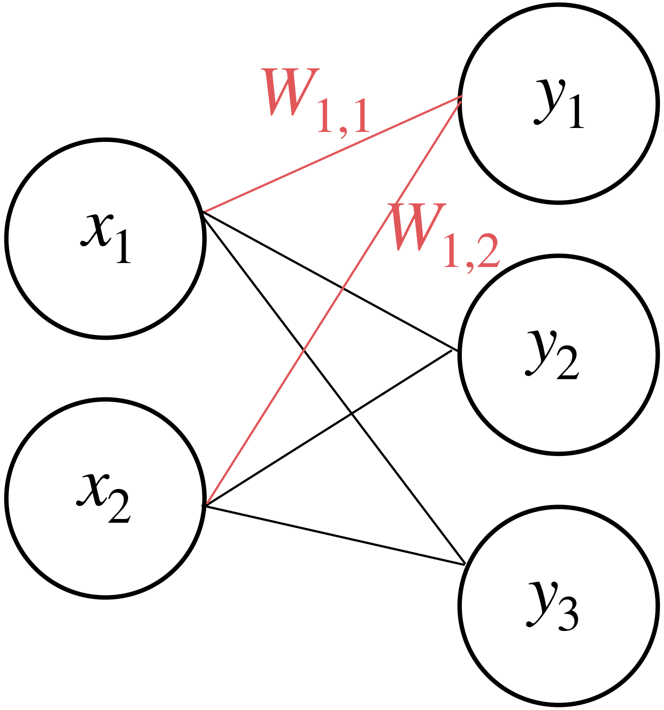
multiple outputs:

$$y_1 = w_1^T x + b_1$$
$$y_2 = w_2^T x + b_2$$

where

$$W = \begin{bmatrix} w_1^T \\ w_2^T \end{bmatrix}$$

Start with Linear Regression



$$y_1 = W_{1,1}x_1 + W_{1,2}x_2 + b_1$$

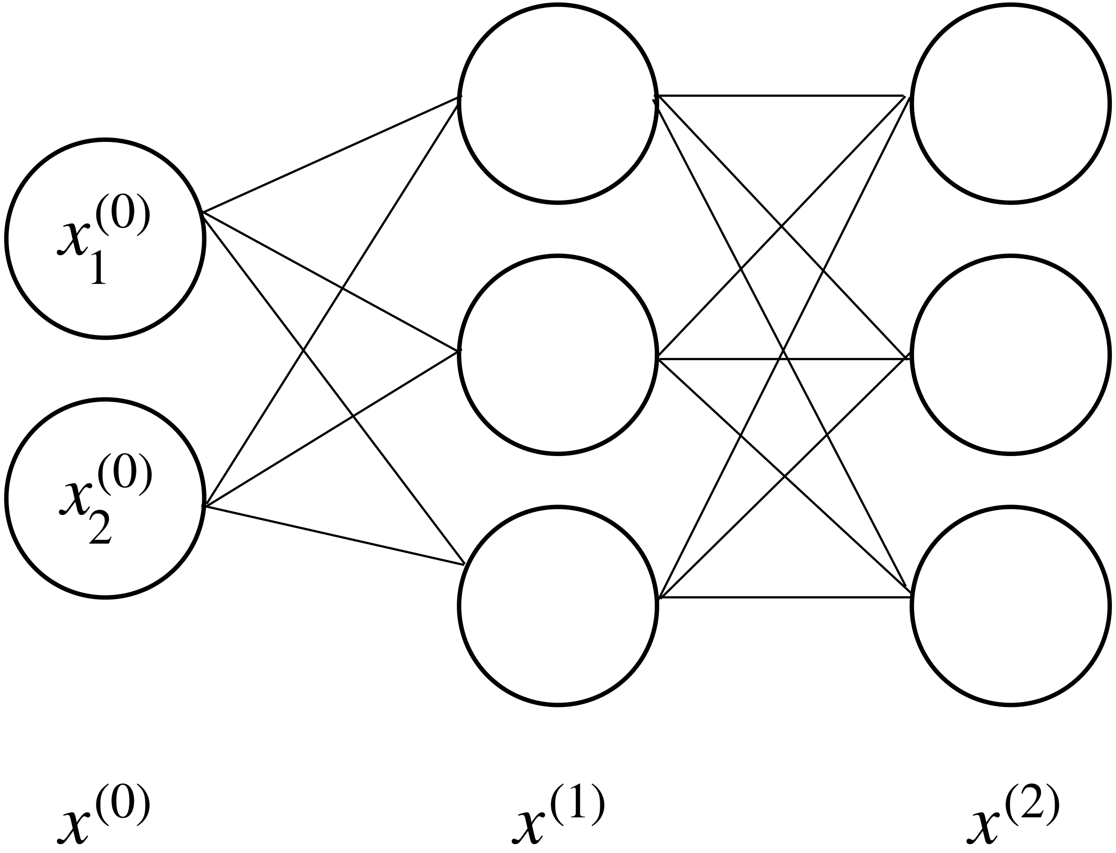
$$y_2 = W_{2,1}x_1 + W_{2,2}x_2 + b_2$$

$$y_3 = W_{3,1}x_1 + W_{3,2}x_2 + b_3$$

$$y = Wx + b$$

why can't we add more layers?

Linear Regression: why can't we add more layers?



$$x^{(1)} = W^{(1)} x^{(0)} + b^{(1)}$$

$$x^{(2)} = W^{(2)} x^{(1)} + b^{(2)}$$

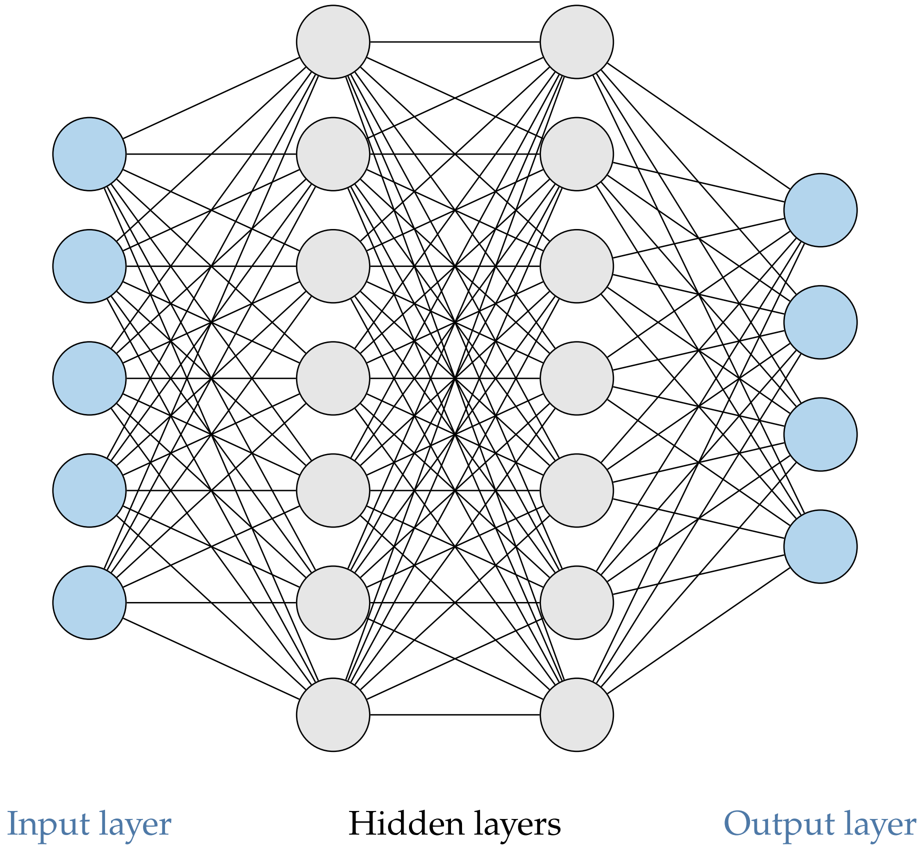
$$x^{(2)} = \underbrace{W^{(2)} W^{(1)}}_{W} x^{(0)} + \underbrace{W^{(2)} b^{(1)} + b^{(2)}}_{b}$$

Function Composition

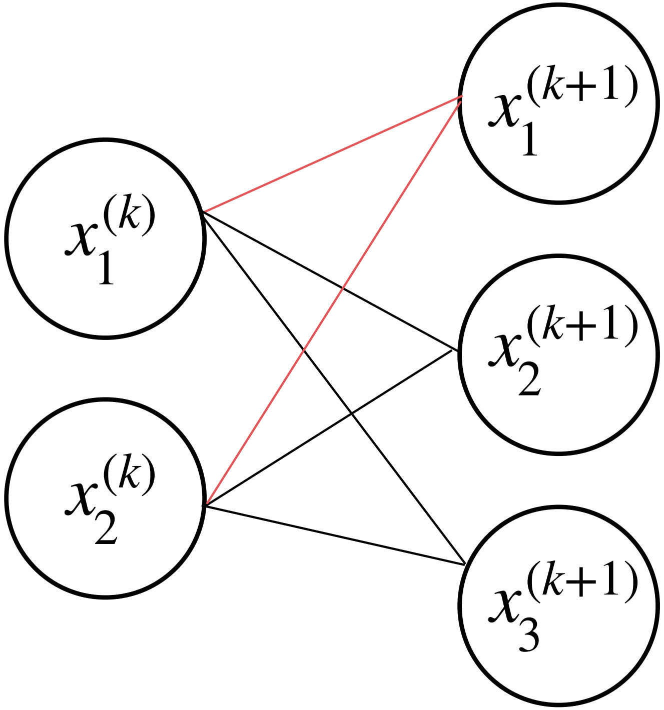
$$\hat{f}(x) = f^{(3)}(f^{(2)}(f^{(1)}(x)))$$

$$\hat{f}(x) = (f^{(3)} \circ f^{(2)} \circ f^{(1)})(x)$$

Neural Network



Dense, Linear, or Fully Connected Layer



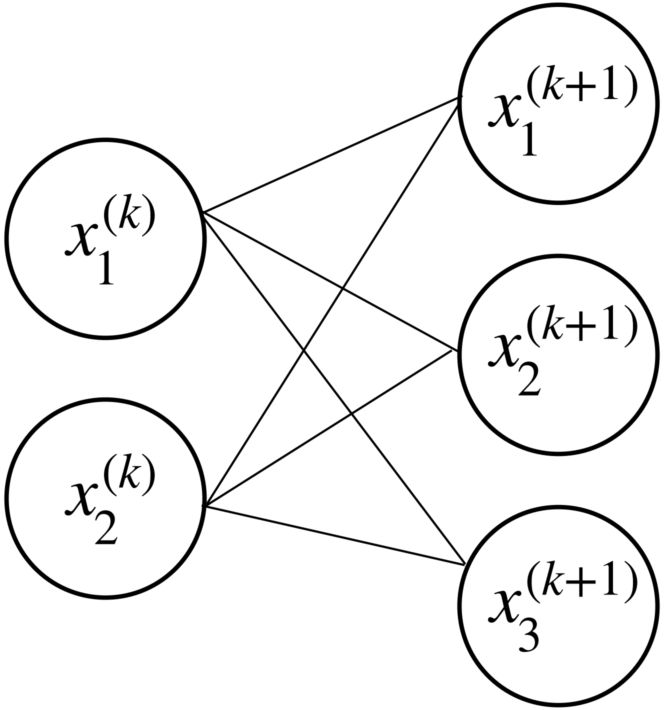
$$z_1^{(k+1)} = w_{1,1} x_1 + w_{1,2} x_2 + b_1$$

$$z_2^{(k+1)} = w_{2,1} x_1 + w_{2,2} x_2 + b_2$$

$$z_3 = \cdots$$

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Dense, Linear, or Fully Connected Layer



$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} W_{1,1} & W_{1,2} \\ W_{2,1} & W_{2,2} \\ W_{3,1} & W_{3,2} \end{bmatrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$z^{(k+1)} = W^{(k+1)} x^{(k)} + b^{(k+1)}$$

Activation Function (nonlinearity)

$$z^{(k+1)} = W^{(k+1)} x^{(k)} + b^{(k+1)}$$

$$x^{(k+1)} = a(z^{(k+1)})$$

activation function