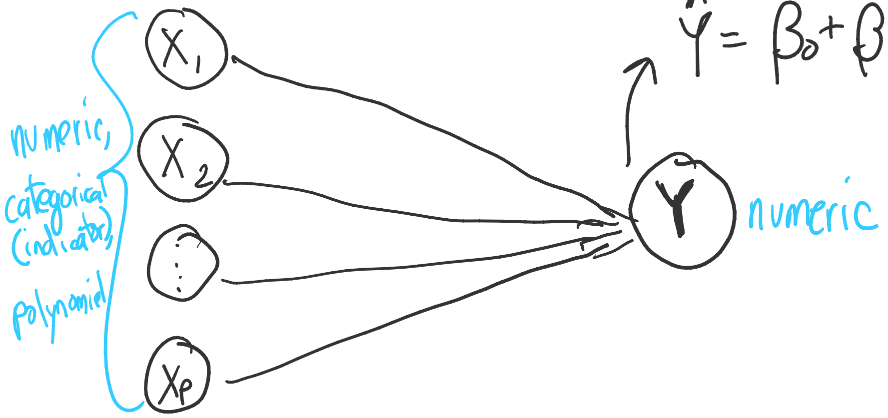
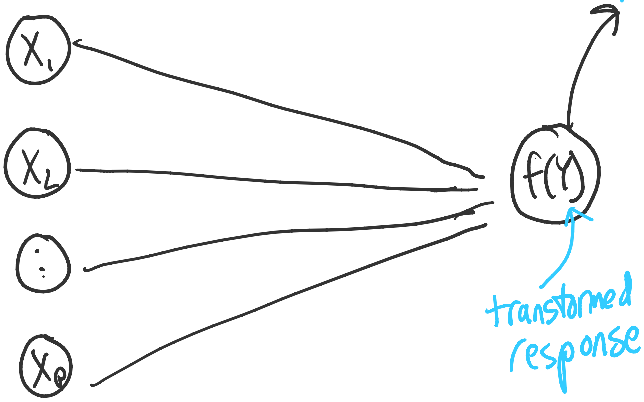
Wednesday, April 12, 2023 2:42 PM



$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_p X_p$$



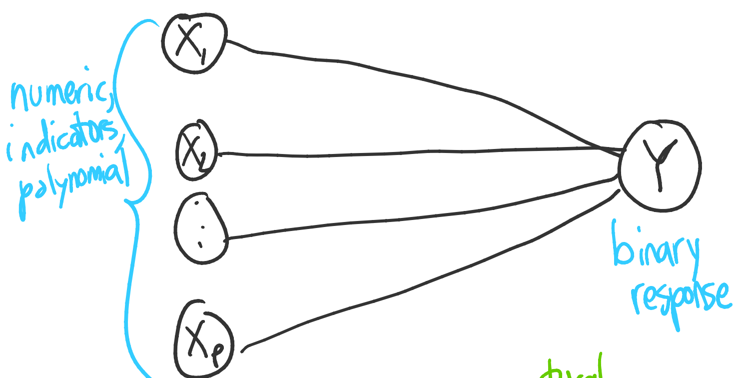
$$f^{-1}(f(\hat{Y})) = f^{-1}(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_p X_p)$$

$$\hat{Y} = f^{-1}(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_p X_p)$$

Machine Learning terms: activation function (f^{-1}), bias ($\hat{\beta}_0$), weights ($\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p$)

Machine Learning for binary responses

- The choice of activation function depends on the range of possible values of the predicted response



$$\hat{Y} = f^{-1}(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_p X_p)$$

What is our activation function?

$$\hat{Y} = P(Y=1) \qquad 0 \leq \hat{Y} \leq 1$$

We need an activation f'n (f^{-1}) that can take a real number and give us a number between 0 & 1.

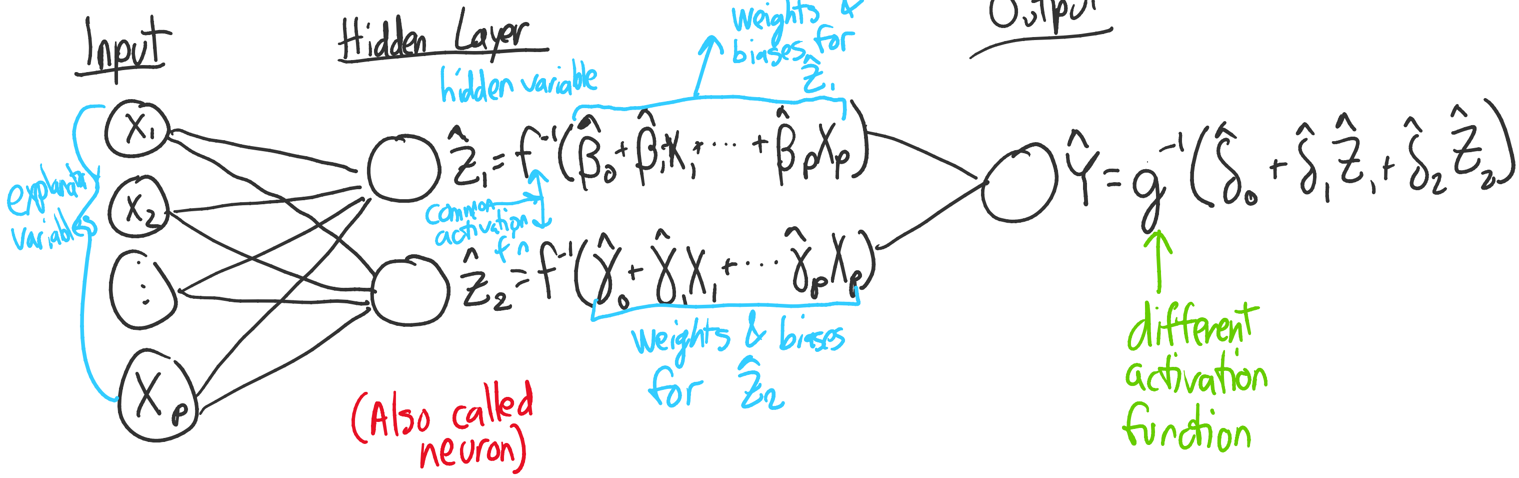
$$f(\hat{Y}) = \text{logit}(\hat{Y}) = \log\left(\frac{\hat{Y}}{1-\hat{Y}}\right)$$

log: natural logarithm; $\log\left(\frac{\hat{Y}}{1-\hat{Y}}\right)$: log odds ratio

$$f^{-1}(\;) = \frac{\exp(x)}{1+\exp(x)} \quad \text{inverse logit function} \implies \hat{Y} = \frac{\exp\{\hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_p X_p\}}{1 + \exp\{\hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_p X_p\}}$$

log odds ratio

$$f^{-1}(x) = \frac{\exp(x)}{1+\exp(x)}$$ } inverse logit function $$\Rightarrow \hat{Y} = \frac{\exp\{\hat{\beta}_0+\hat{\beta}_1X_1+\cdots+\hat{\beta}_pX_p\}}{1+\exp\{\hat{\beta}_0+\hat{\beta}_1X_1+\cdots+\hat{\beta}_pX_p\}}$$



This type of algorithm is called a **neural network**

↳ this algorithm tries to mimic the functionality of the human brain!

Notes about neural networks

- You will NOT be tested on complex neural networks
- You can have as many hidden layers & as many neurons in each layer as you want (but don't want too many)
- Logistic Regression and Linear regression are special types of neural networks with no hidden layers

Major Drawback of Neural Networks

- You lose interpretability of model parameters for perhaps only moderately improved predictions