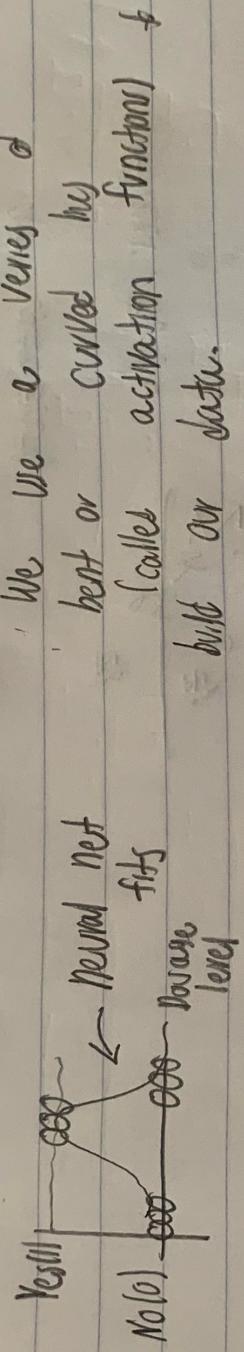


Imagine we have a dataset that estimates how effective a treatment of some medicine would be.



Many types of activation functions

Softplus

Cells (rectified linear unit)

sigmoid

$\ln(1 + e^x)$

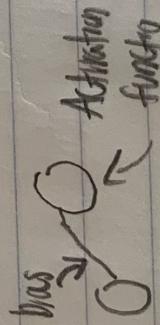
$\max(0, x)$

$\frac{1}{1 + e^{-x}}$

Many activation functions

Input Nodes  $\rightarrow$  Hidden Layer  $\rightarrow$  Output node

This is an estimation



Neural networks fitting gives us different parameters  
estimation on connection between nodes and that  
results in each node in the hidden layer to use  
different portions of activation function.