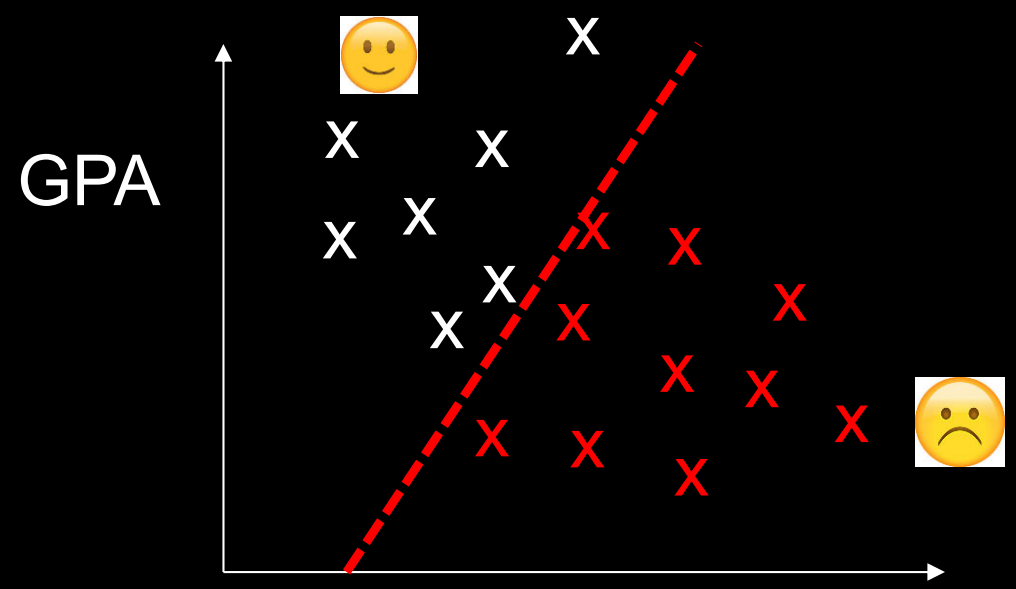
Problem: How to classify stimuli?

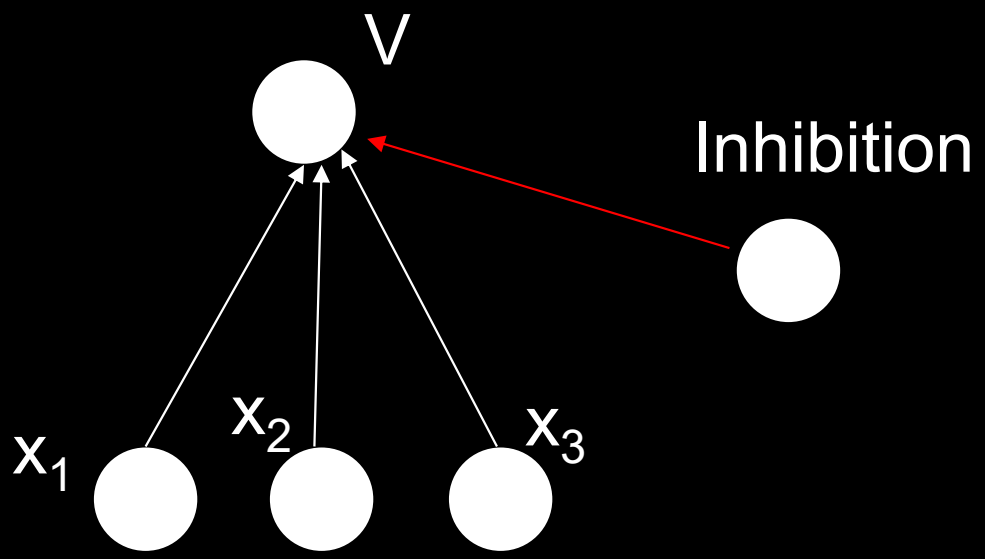
- Suppose you are trying to hire someone for an entry level job. Ideally you want someone who is smart and who doesn't drink too much and miss work. You only know their GPA out of college and how much their friends say they drink and miss work. Which ones would you call for an interview?

- You might draw a boundary line to decide who is "above threshold"



History of neural networks – McCulloch-Pitts

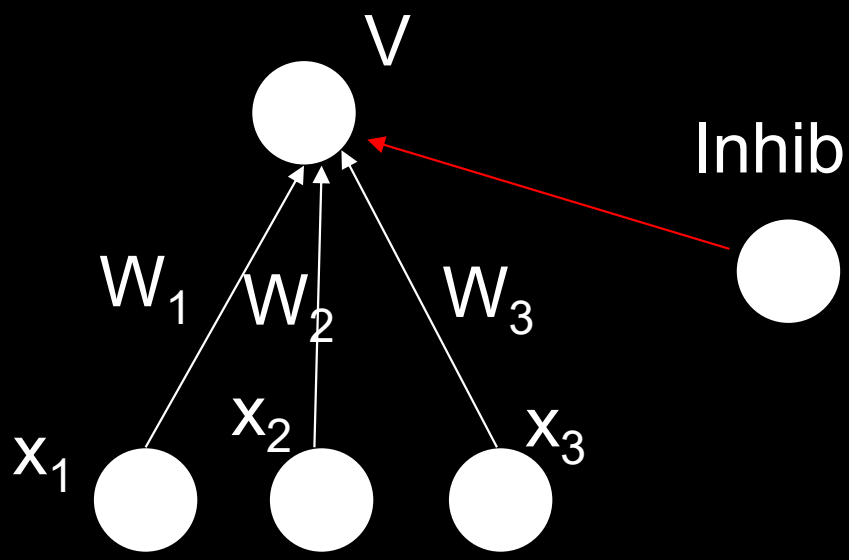
- McCulloch & Pitts (1943) developed first computation model of neurons.
- V is the firing rate of the “output” cell. V > 0 if the number of active input cells exceeds a threshold (e.g., 2 cells), and if no “inhibitory” cells to V are active
- Note: no learning in this model

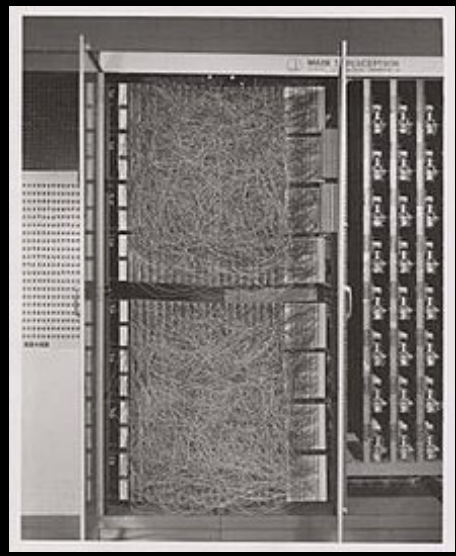


Perceptrons

- Rosenblatt (1962) introduced the Perceptron model
- Like McCulloch & Pitts, but with graded inputs rather than all or none. Also, the signals to the output unit are WEIGHTED, so that some are more influential than others on the output
- The weights W_i indicate how strong the synapse is.
- $V = bias + x_1W_1 + x_2W_2 + x_3W_3$

Generally $V = \sum_{i=1}^{N} W_i x_i + bias$ (such that V >0)



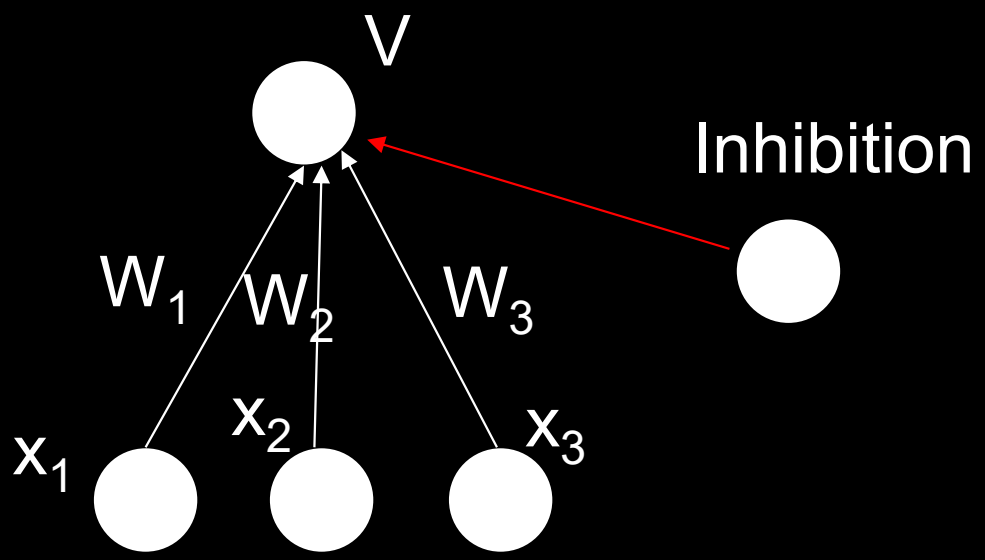


Perceptrons

- Let's look at the equation more closely:

$$V = \sum_{i=1}^{N} W_i x_i + bias$$

- Notice that this is just matrix multiplication:
  - V=Wx, where
  - $W = [W_1\ W_2\ W_3]$
  - $X = [x_1\ x_2\ x_3]^T$

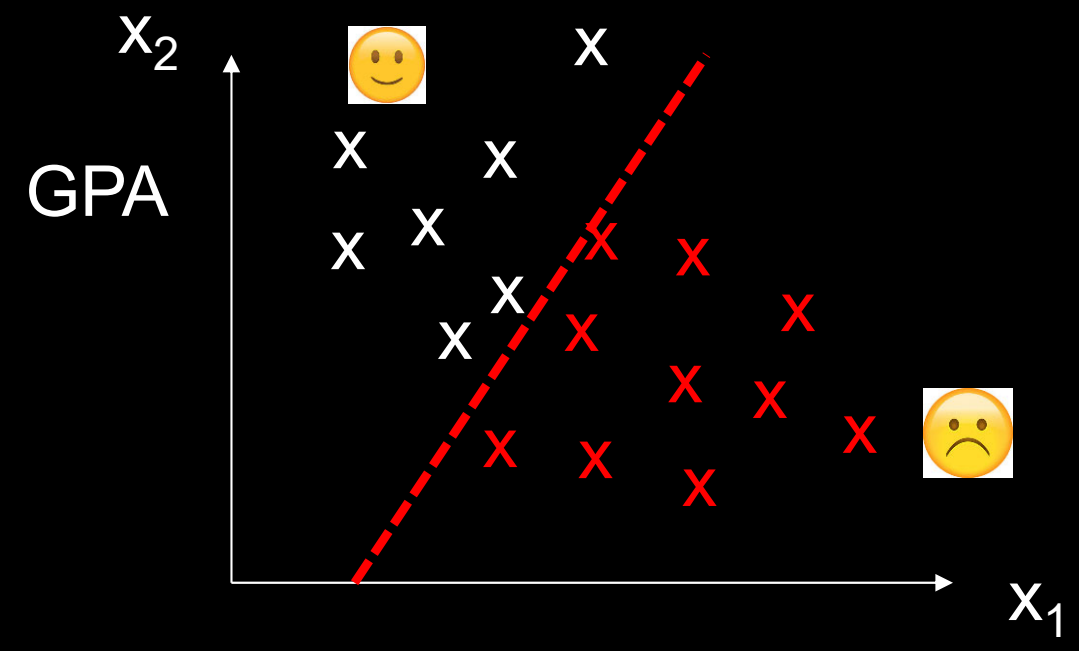


Problem: How to classify stimuli?

- Now we can classify as "above threshold" if
- V=Wx (matrix multiplication) is greater than some threshold b
- $V = W_1 * x_1 + W_2 * x_2 - b$
- i.e. $V = W_1 * x_1 + W_2 * x_2 - b > 0$
- → solve $W_1 * x_1 + W_2 * x_2 - b > 0$
- $X_2 > \frac{-W_1 x_1 + b}{W_2}$ is boundary line (flip "<" if $W_2 > 0$)

Equation defines this boundary line

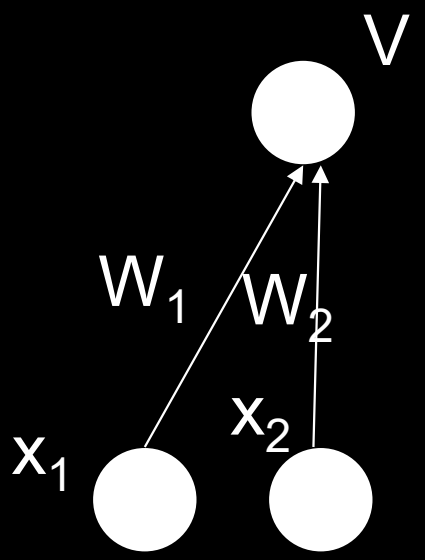
Score penalty for drinking more



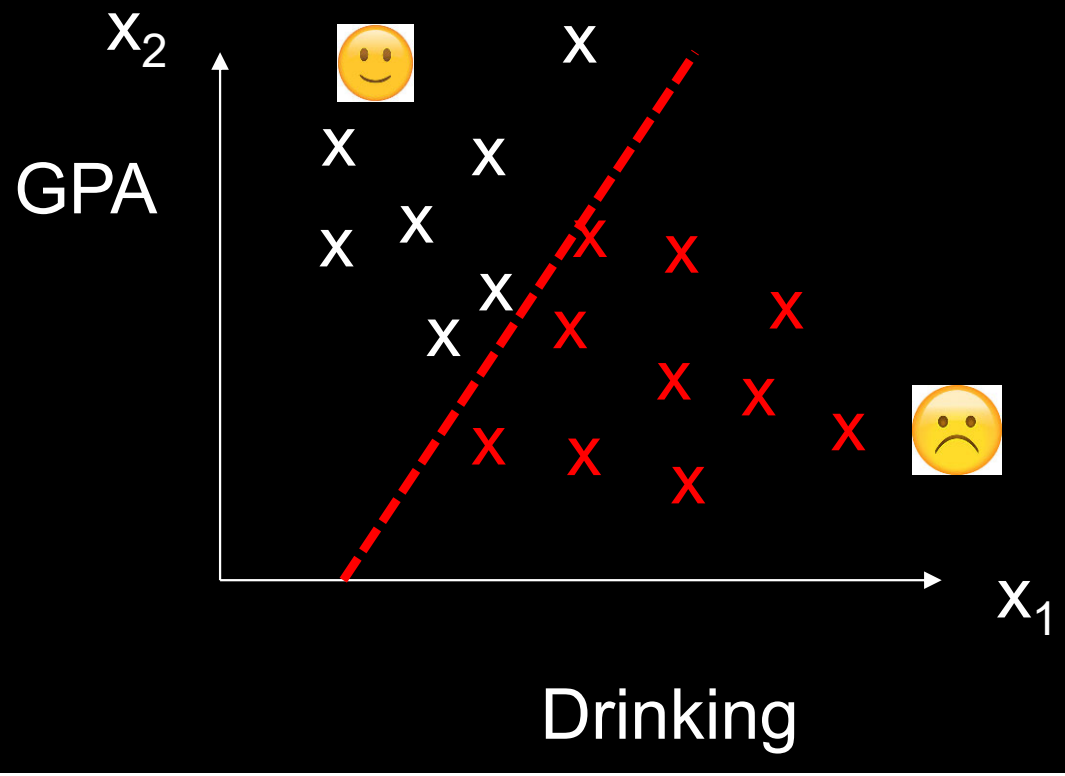
Perceptrons

- The decision boundary is exactly what the perceptron gives us!
- If V = $\sum_{i=1}^{N} W_i x_i + bias$ > 0, then classify the job candidate as "above threshold" for an interview

Perceptron

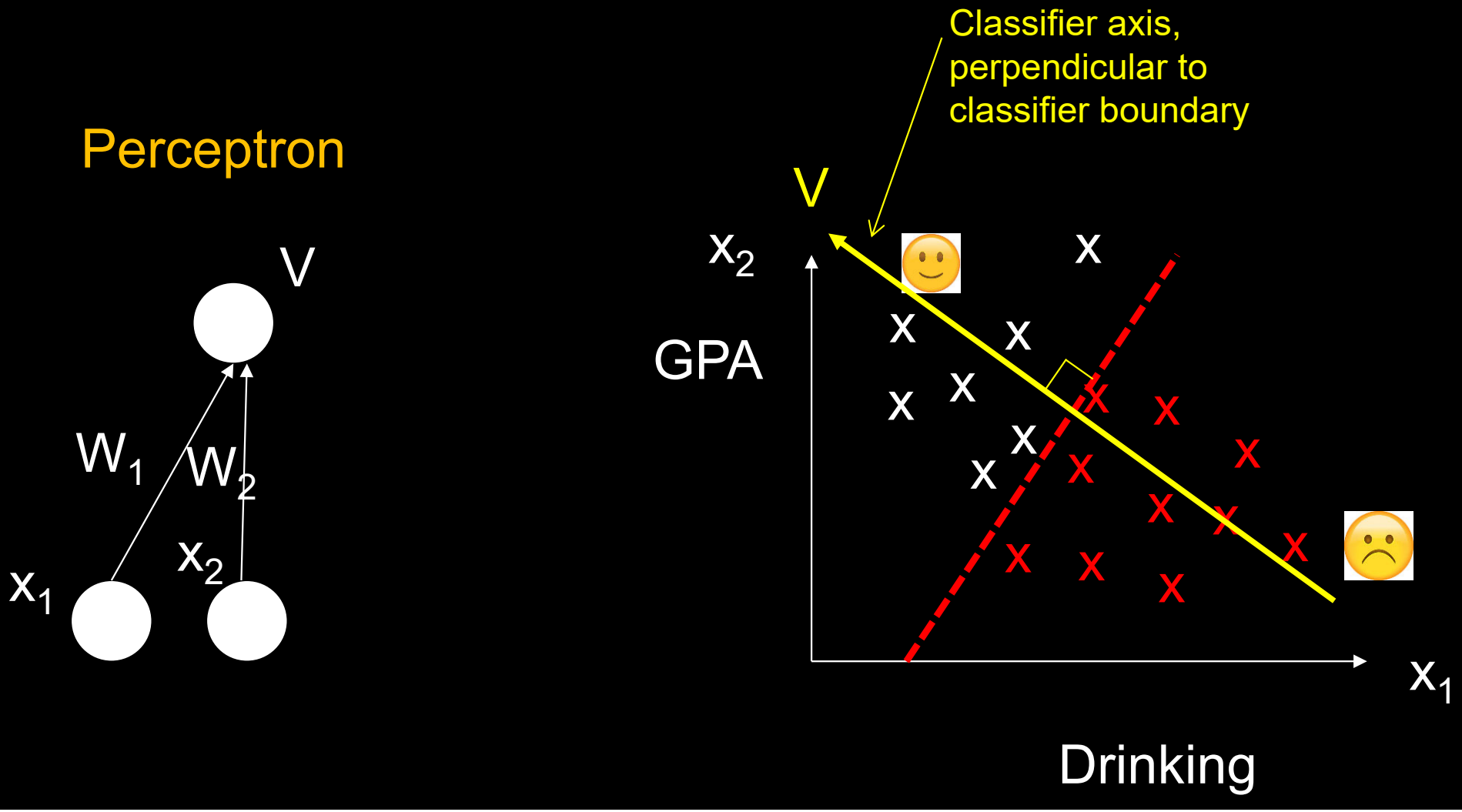


Classification



Perceptrons

- Another way to think about it is as a dot product
- Dot product V=Wx tells us where on the classifier axis the data point falls: above threshold, or below?



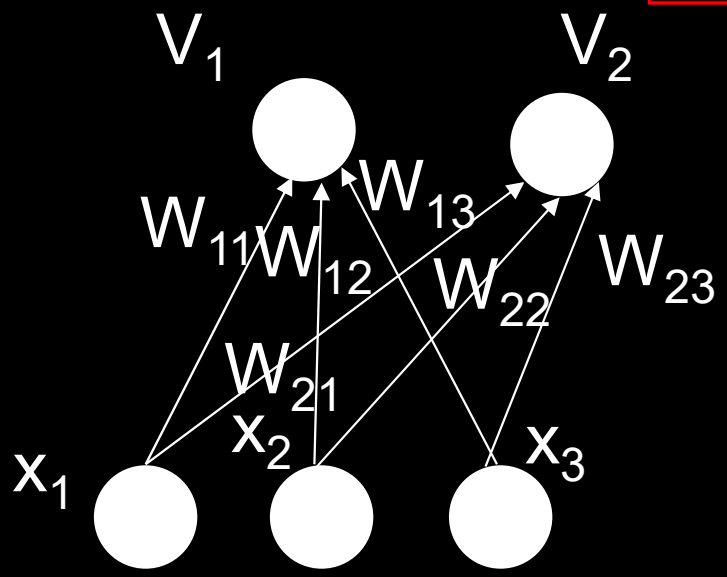
Perceptrons

- We can generalize this further with multiple outputs:
  - V=Wx, where
  - $W = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \end{bmatrix}$
  - $X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$

So V=Wx

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} W_{11}X_1 & W_{12}X_2 & W_{13}X_3 \\ W_{21}X_1 & W_{22}X_2 & W_{23}X_3 \end{bmatrix}$$



$$V_j = \sum_{i=1}^{N} W_{ij} x_i$$

Perceptrons – learning?

- Problem: How do we know what W should be for a good decision???
- Answer: next time…

Perceptron