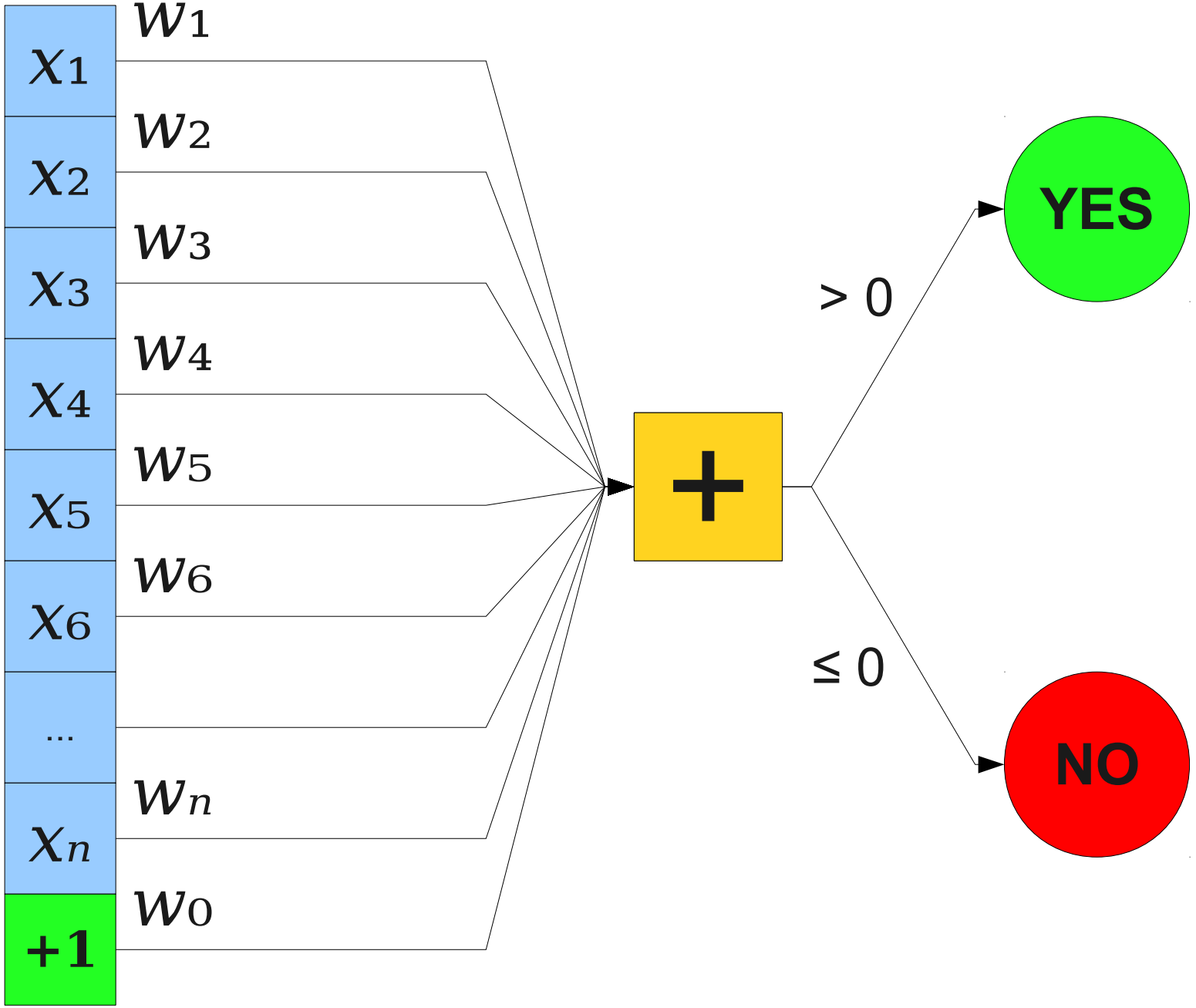
Machine Learning

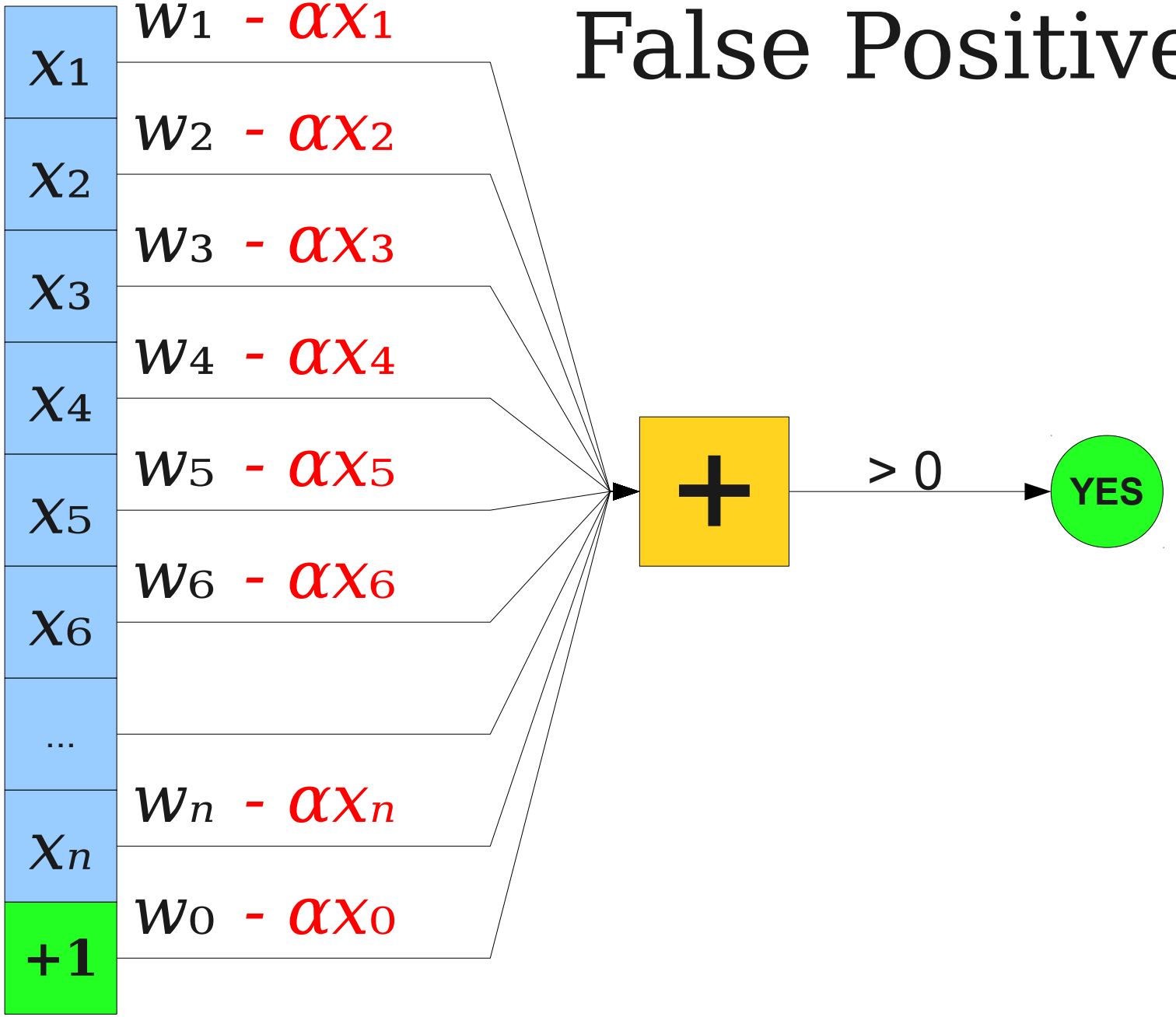
Perceptron Learning

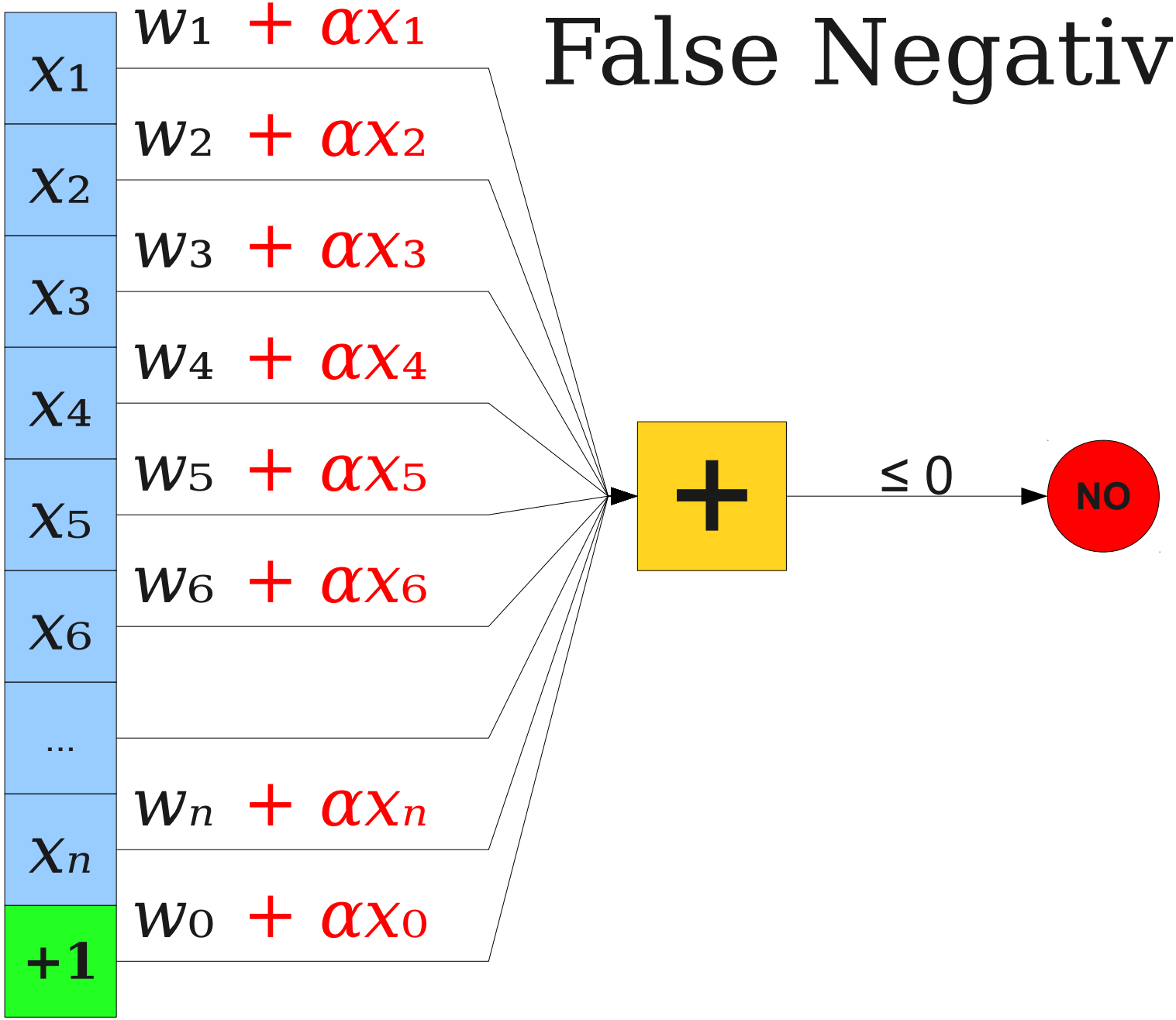

x_1
x_2
x_3
x_4
x_5
x_6
...
x_n
+1
w_1
w_2
w_3
w_4
w_5
w_6
w_n
w_0
+
> 0
≤ 0
YES
NO

How do we choose good values for w_0 ... w_n?

One Approach

- **Train** the perceptron on valid data.
- For each data point:
  - Ask the perceptron what it thinks.
  - If correct, do nothing.
  - Otherwise, nudge w_0 … w_n in the right direction.
- Repeat until number of errors is “small enough.”
- Question: What kind of mistakes can we make?


False Positive
x1
x2
x3
x4
x5
x6
...
xn
+1
w1 - αx1
w2 - αx2
w3 - αx3
w4 - αx4
w5 - αx5
w6 - αx6
wn - αxn
w0 - αx0
+
> 0
YES


False Negative
x_1
x_2
x_3
x_4
x_5
x_6
...
x_n
+1
$w_1 + \alpha x_1$
$w_2 + \alpha x_2$
$w_3 + \alpha x_3$
$w_4 + \alpha x_4$
$w_5 + \alpha x_5$
$w_6 + \alpha x_6$
$w_n + \alpha x_n$
$w_0 + \alpha x_0$
+
≤ 0
NO

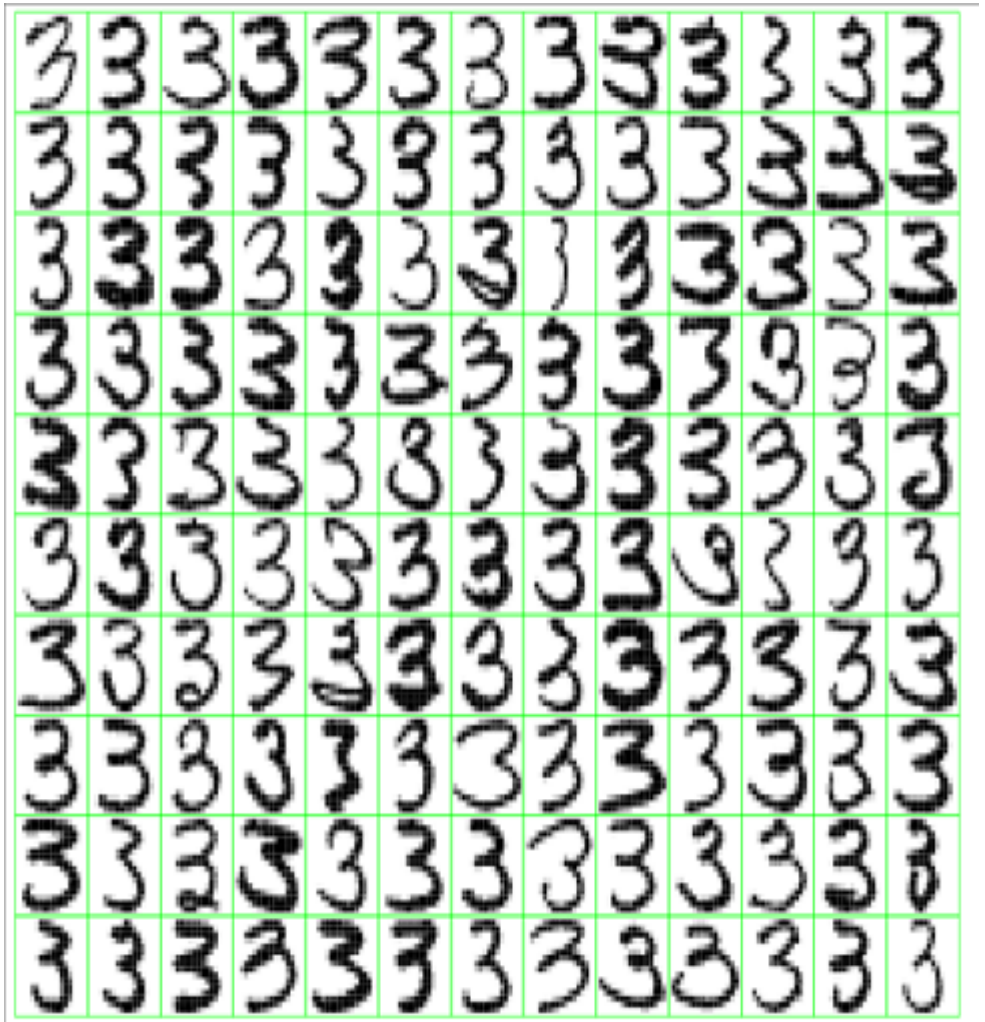
A Cute Math Trick

- For false positives, set $w_k = w_k - \alpha x_k$.
- For false negatives, set $w_k = w_k + \alpha x_k$.
- For correct answers, set $w_k = w_k$.
- Let "YES" be 1 and "NO" be 0.
- Consider the difference between **actual answer** and **perceptron guess**:
  - False positive: Actually NO, we say YES. Difference is -1.
  - False negative: Actually YES, we say NO. Difference is +1.
  - Correct answer: Both YES or both NO. Difference is 0.
- General update rule: $\mathbf{w_k = w_k + \alpha(real - guess)x_k}$.

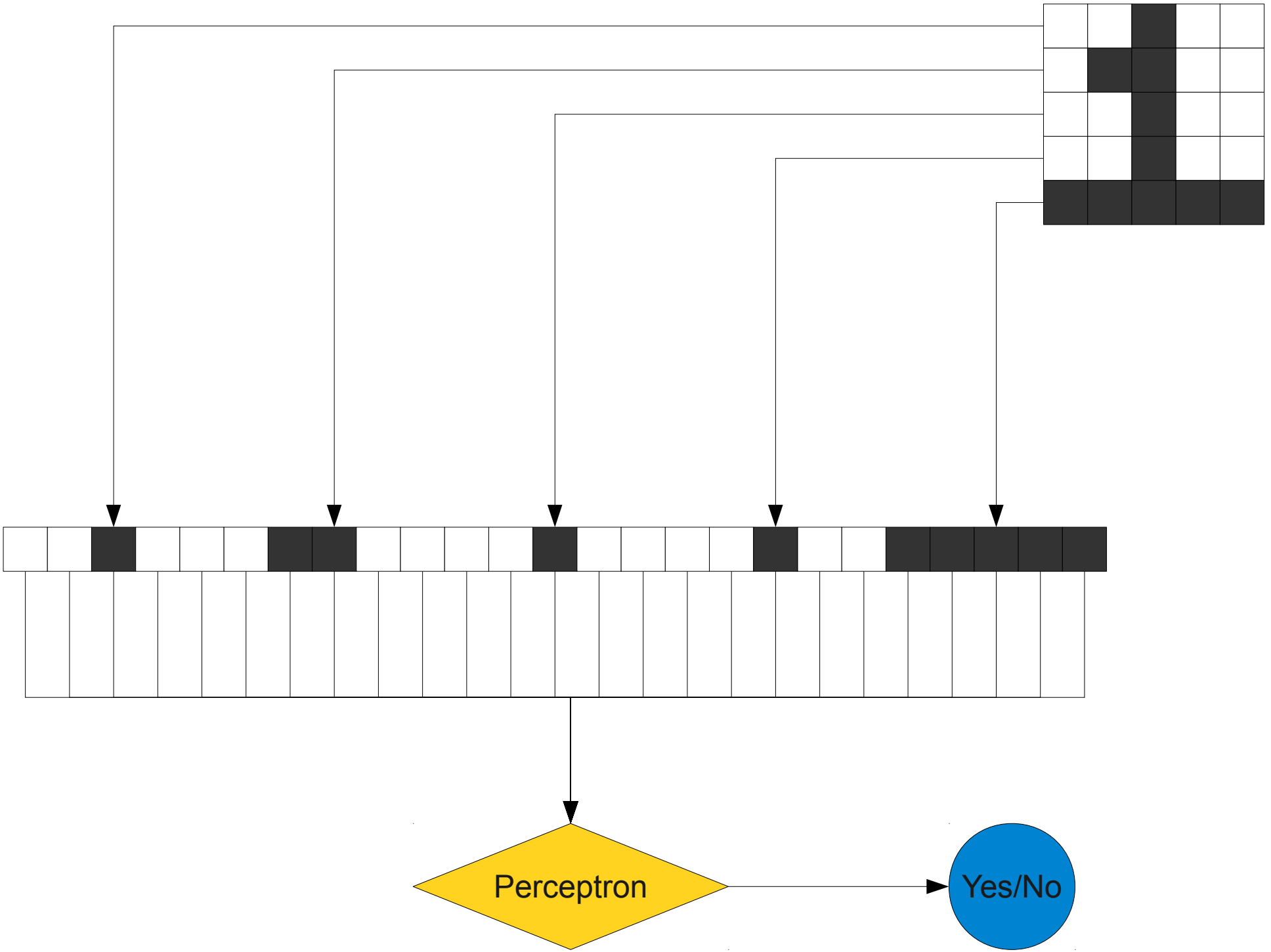
Perceptron Learning Algorithm

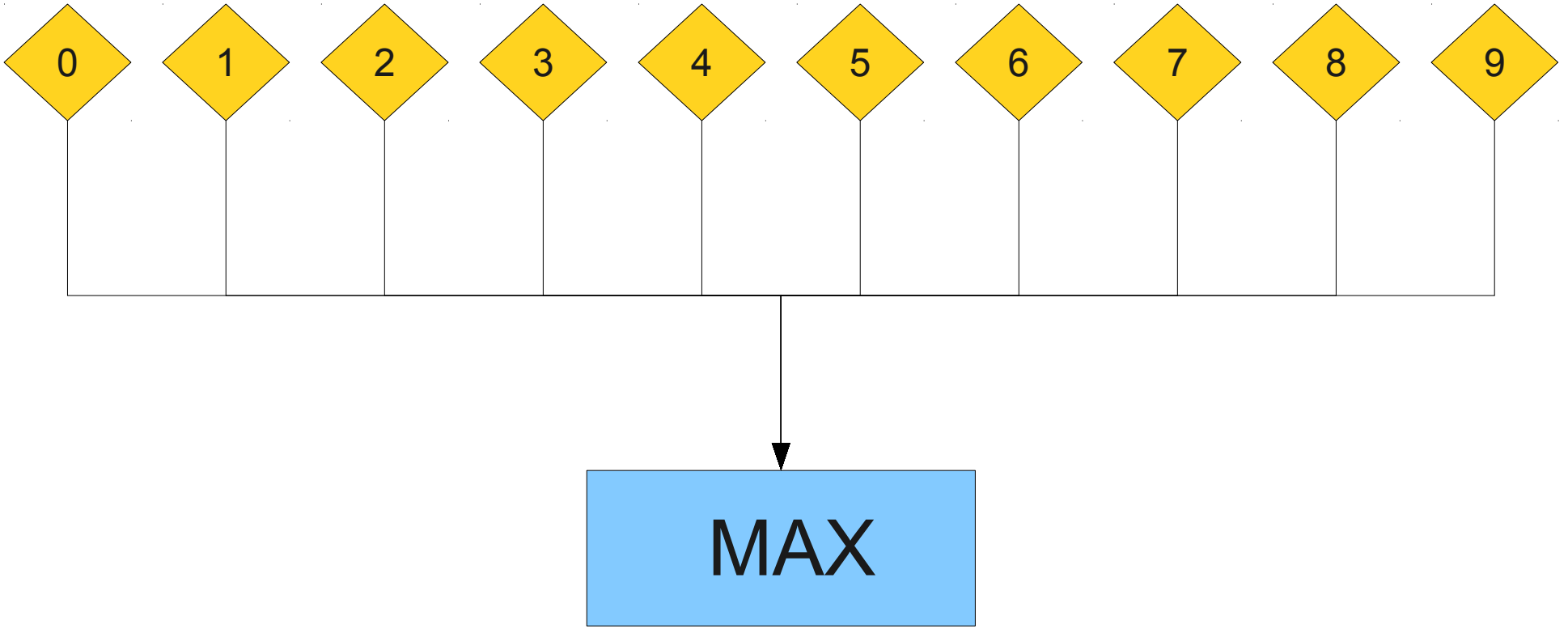
- Start with a random guess of each w_k.
- Repeat until perceptron is sufficiently accurate:
  - Choose a training example (x_0, x_1, ..., x_n).
  - Let *real* be the real answer, *guess* be the perceptron's guess.
  - For each k, set $\boldsymbol{w_k = w_k + \alpha(real - guess)x_k}$
- Note: Use **batching** in practice.
  - Update everything all at once.

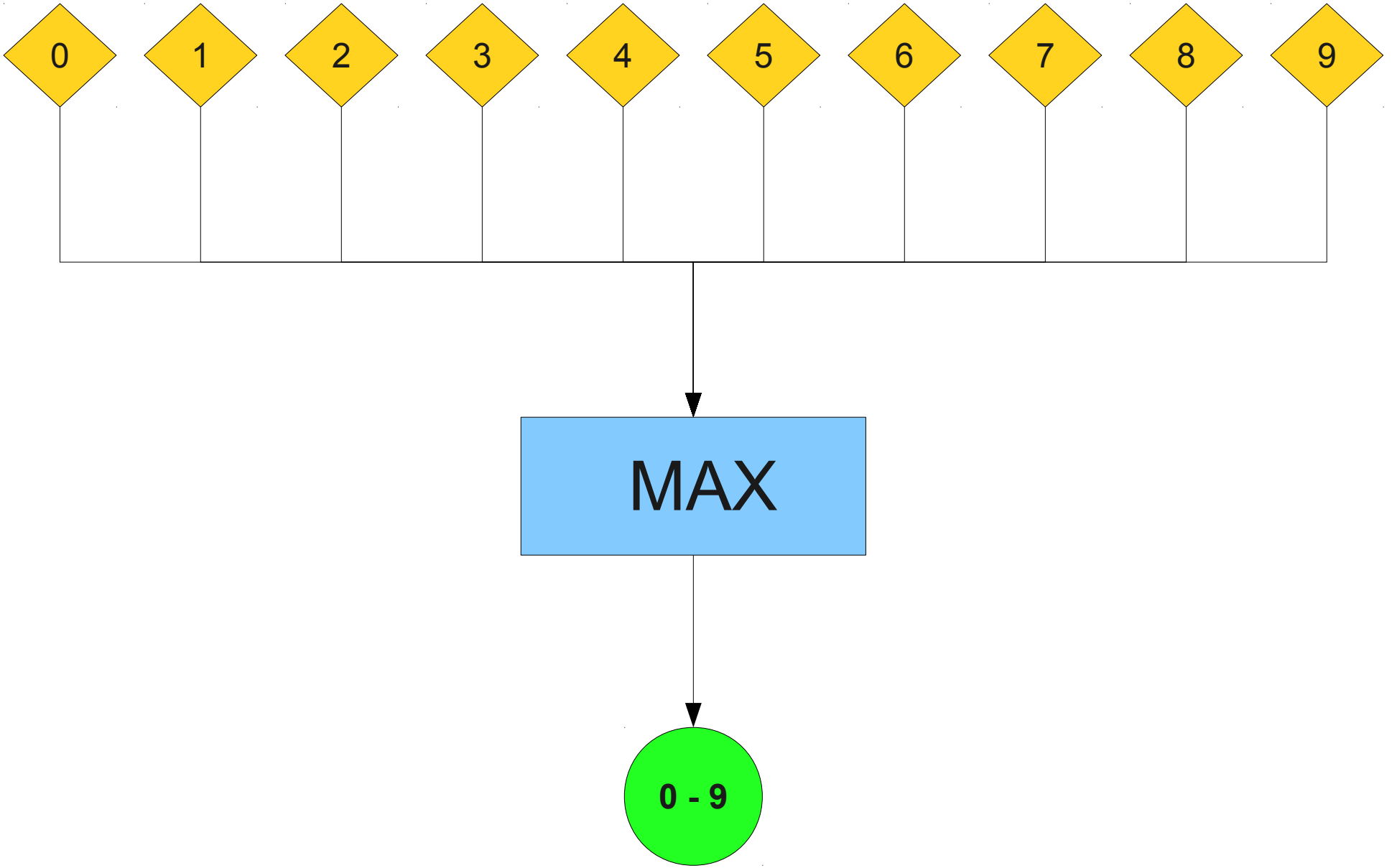
Application: Handwriting Analysis



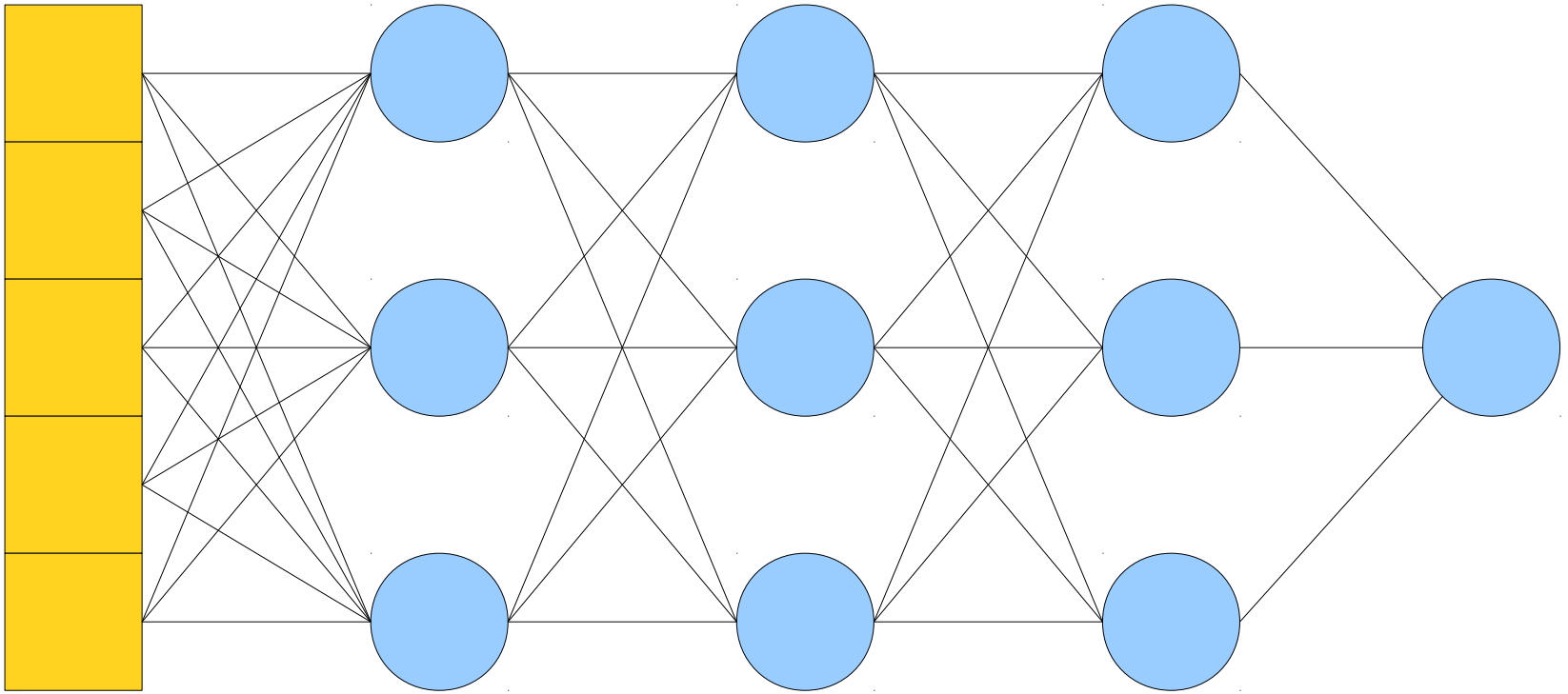
- Train a computer to recognize handwritten numbers 0 – 9.
- Large training and test set available (MNIST Handwritten Digit Database)


Perceptron
Yes/No


0
1
2
3
4
5
6
7
8
9
MAX


0
1
2
3
4
5
6
7
8
9
MAX
0 - 9

Combining Perceptrons



This is called a neural network.

Machine Learning

- Interesting in machine learning? Take CS109 and CS229!
- Many beautiful algorithms:
  - Naive Bayes classifiers (used in spam filtering).
  - Decision trees (used in hospitals for diagnostics).
  - Bayesian networks (used in cancer research to learn what causes tumors).
  - Autoencoders (cutting-edge research; can detect patterns in just about anything)