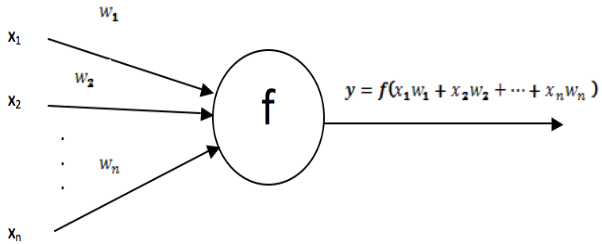
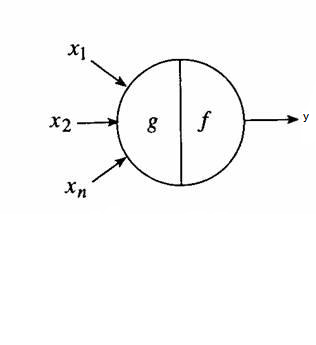
**CSC 417 Unit 2 Day 1 Outline**

1. Deep Learning
   1. Mind vs. Machine
      1. Search in AI doesn’t seem to work as the human mind does
         1. Machines perform extensive calculations, humans perform minimal calculations
         2. Humans often compare new information to stored knowledge
         3. Humans make associations between pieces of information
      2. Neural Networks are inspired by the human brain
         1. Copy observed functionality of neurons
            1. Strength of synapse connections can strengthen or weaken over time
   2. **Artificial Neurons
      1. Components
         1. Input vector *x*
         2. Weight vector *w*
         3. Function *f*
         4. Output *y*
            1. *f* applied to dot product (matrix multiplication) of *w* and *x*
   3. Artificial Neural Network
      1. Definition
         1. Collection of artificial neurons (processing units) arranged in some topology
            1. Input layer, some number of hidden layers, and an output layer
      2. Function
         1. An input vector *x* should produce a certain output *y*… but how?
         2. Adjust weights of connections between neurons
   4. Development of ANNs
      1. McCulloch-Pitts Network (1943)
         1. First artificial neuron
         2. Input and output are binary
            1. If *g*(*x*) < θ, y = 0
            2. Else y - 0
         3. No learning is possible due to lack of weights
            1. MC-P is not adaptive
      2. Perceptron Learning Rule
         1. Weights are adjusted based on output from network
            1. Replace threshold θ with an activation function
            2. If output is incorrect, adjust weights
            3. Calculate new weight as wnew = wold + η(t – o)xi

η = learning rate

* + - 1. Network converges (learns correct output if)
         1. Training data is linearly separable
         2. Learning rate is sufficiently small

If learning rate is too small, network will take too long to converge

* + - * 1. Perceptrons are adaptable
      1. Classic shortcoming – single layer perceptron cannot learn XOR (not linearly seperable)
    1. Backpropagation
       1. Method for adjusting weights in a multilayer network
          1. Attempts to determine “share” of error to assign to each weight
       2. Requires large training sets and many epochs (passes through training data)
       3. Activation function must be continuous and differentiable
       4. Procedure
          1. Feed forward computation (input -> output)
          2. Backpropagation of error to output layer
          3. Backpropagation of error to hidden layer
          4. Update all weights
    2. Feature Extraction
       1. Deep neural network = network with two or more *hidden* layers
       2. Diagram

          Description automatically generatedDeep networks *automatically* perform feature extraction (identification of “best” variables in data) – separate extraction step is not needed