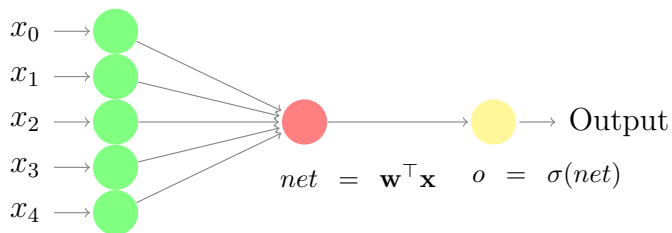


2.2 An Example of a Multilayer Neural Network

- Linear Unit
- Perceptron Unit
- Sigmoid Unit
 - Smooth, differentiable threshold function

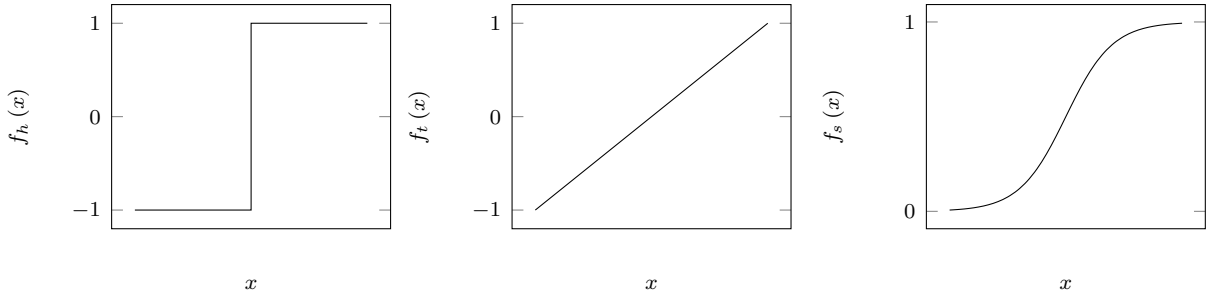
$$\sigma(net) = \frac{1}{1 + e^{-net}}$$

- Non-linear output



As mentioned earlier, the perceptron unit cannot be used as it is not differentiable. The linear unit is differentiable but only learns linear discriminating surfaces. So to learn non-linear surfaces, we need to use a non-linear unit such as the sigmoid.

2.3 Properties of Sigmoid Function

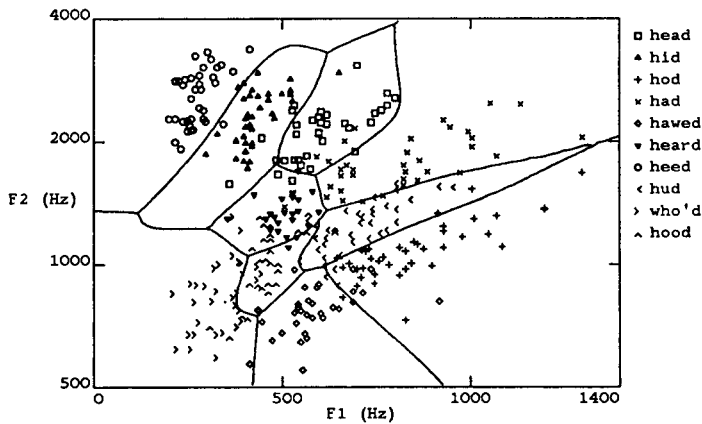


The threshold output in the case of the sigmoid unit is continuous and smooth, as opposed to a perceptron unit or a linear unit. A useful property of sigmoid is that its derivative can be easily expressed as:

$$\frac{d\sigma(y)}{dy} = \sigma(y)(1 - \sigma(y))$$

One can also use e^{-ky} instead of e^{-y} , where k controls the “steepness” of the threshold curve.

2.4 Motivation for Using Non-linear Surfaces



The learning problem is to recognize 10 different vowel sounds from the audio input. The raw sound signal is compressed into two features using spectral analysis.