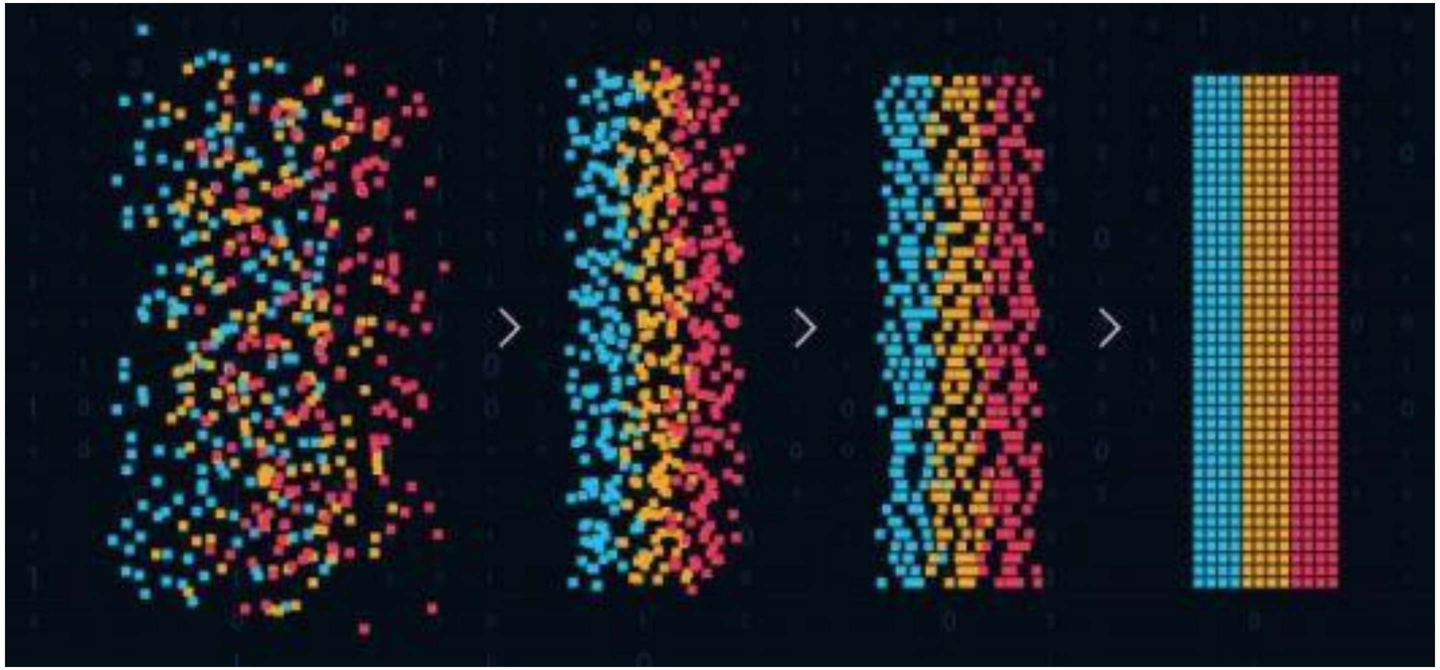
Pattern and Anomaly Detection



Source: Edureka

B. Tech., CSE + AI/ML

Dr Gopal Singh Phartiyal

11/11/2021

Neural Networks

- Output of form $$y(\mathbf{x}, \mathbf{w}) = f\left(\sum_{j=1}^{M} w_j \phi_j(\mathbf{x})\right)$$
- In linear models of classification,
  - fixed number of basis functions
  - f is non-linear function (decision function)
- In linear models for regression
  - Fixed number of basis functions
  - f is identity function
- Neural networks
  - Basis functions are parameter based and these parameters are learned during training f is non-linear function (decision.
  - Many ways to form parametric basis functions

Neural Networks

- M different linear combinations for D dimensional input.
- ‘a’ is activation and w is weights

$$a_j = \sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)}$$

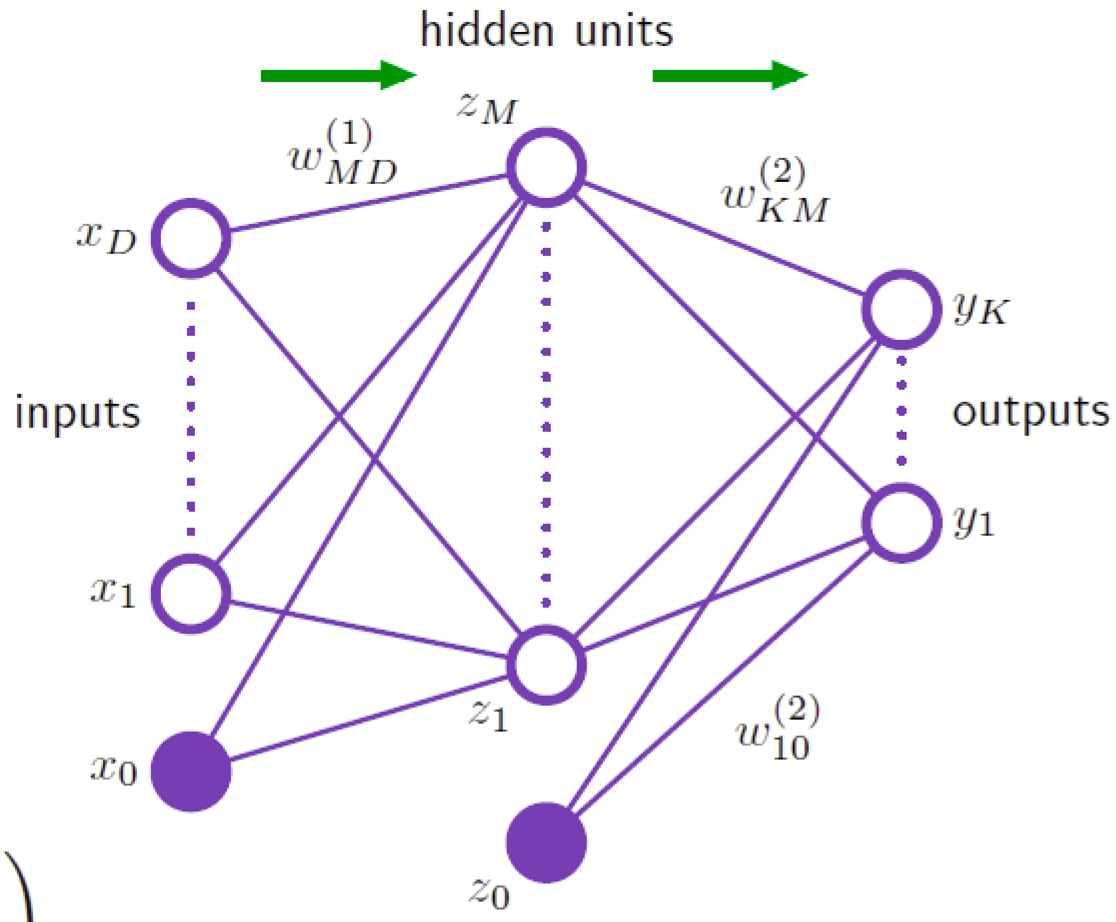
- ‘h’ is activation function.
- z is the output after activation function (hidden unit)

$$z_j = h(a_j).$$

- Repeat

$$a_k = \sum_{j=1}^{M} w_{kj}^{(2)} z_j + w_{k0}^{(2)}$$

Neural Networks

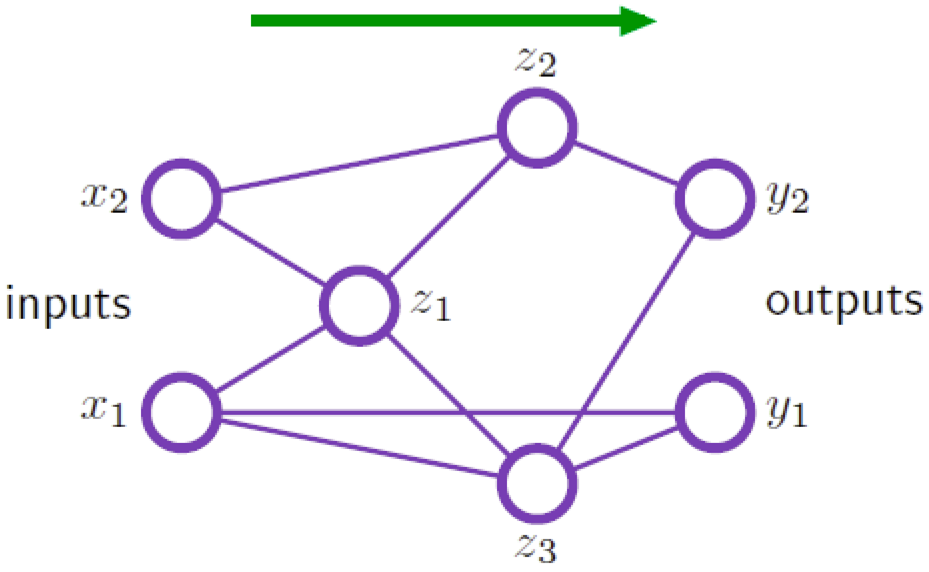


$$y_k(\mathbf{x}, \mathbf{w}) = \sigma\left(\sum_{j=1}^{M} w_{kj}^{(2)} h\left(\sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)}\right) + w_{k0}^{(2)}\right)$$

Neural Networks

- MLP: continuous sigmoidal non-linear function
- Perceptron: step-function non-linear function
- NNs can have other activation functions
- General architecture

$$z_k = h\left(\sum_j w_{kj} z_j\right)$$



Neural Networks: Network Training

- Find w?????
- Intuitive approach: Sum-of-squares error and its minimization with least squares.
- Alternative and more general approach: Probabilistic interpretation
  - Consider Gaussian
  - Create maximum likelihood
  - Either maximize it or minimize the negative of it.
  - We also know relationship between maximum likelihood and sum-of-squares error
  - Use this relationship to create error functions

Neural Networks: Network Training

- Error Functions in NN based models
- 1. Regression: When the output layer of the NN has linear or identity activation function

$$\frac{\partial E}{\partial a_k} = y_k - t_k$$

- 2. Classification
  - Binary: Two class

$$E(\mathbf{w}) = -\sum_{n=1}^{N} \{t_n \ln y_n + (1 - t_n) \ln(1 - y_n)\}$$

  - Multiple two-class

$$E(\mathbf{w}) = -\sum_{n=1}^{N} \sum_{k=1}^{K} \{t_{nk} \ln y_{nk} + (1 - t_{nk}) \ln(1 - y_{nk})\}$$

Neural Networks: Network Training: Errors

- Error Functions in NN based models
- 1. Regression: When the output layer of the NN has linear or identity activation function

$$\frac{\partial E}{\partial a_k} = y_k - t_k$$

- 2. Classification
  - Binary: Two class

$$E(\mathbf{w}) = -\sum_{n=1}^{N} \{t_n \ln y_n + (1 - t_n) \ln(1 - y_n)\}$$

  - Multiple two-class

$$E(\mathbf{w}) = -\sum_{n=1}^{N} \sum_{k=1}^{K} \{t_{nk} \ln y_{nk} + (1 - t_{nk}) \ln(1 - y_{nk})\}$$

  - Multi-class

$$E(\mathbf{w}) = -\sum_{n=1}^{N} \sum_{k=1}^{K} t_{kn} \ln y_k(\mathbf{x}_n, \mathbf{w}).$$

Next time: Neural Networks

Thank You

UPES
UNIVERSITY WITH A PURPOSE