IIT ROORKEE

NPTEL ONLINE CERTIFICATION COURSE

ARTIFICIAL NEURAL NETWORK PART-2

LECTURE 54

DR. GAURAV DIXIT

DEPARTMENT OF MANAGEMENT STUDIES

ARTIFICIAL NEURAL NETWORKS

- Computing output values at nodes of each layer type
  - Hidden layer nodes
    - θ_j and w_{ij} are typically initialized to small random values in the range 0.0 ± 0.05
    - Network updates these values after learning from data during each iteration or round of training
  - Output layer nodes
    - Steps are same as for hidden layer nodes, except the fact that input values are received from last hidden layer
    - Output values produced by nodes are used as
      - Predictions in a prediction task
      - Scores to be used to classify a record in a classification task

ARTIFICIAL NEURAL NETWORKS

- Open RStudio
- Neural Network training process
  - Steps to compute neural network output values are repeated for all the records in the training partition
  - Prediction errors are used for learning after each iteration
- Linear and Logistic regression as special cases
  - A neural network with single output node and no hidden layers would approximate the linear and logistic regression models

ARTIFICIAL NEURAL NETWORKS

- Linear and Logistic regression as special cases
  - If a linear transfer function (g(x) = bx) is used, output would be

$$y = \theta + \sum_{i=1}^{p} w_i x_i$$

    - A formulation equivalent to multiple linear regression equation
    - However, estimation method (least squares) is different from neural network (back propagation)

ARTIFICIAL NEURAL NETWORKS

- Linear and Logistic regression as special cases
  - If a logistic transfer function ($g(x) = 1/1+e^{-bx}$) is used, output would be

$$P(y = 1) = \frac{1}{1 + e^{\theta + \sum_{i=1}^{p} w_i x_i}}$$

    - A formulation equivalent to logistic regression equation
    - However, estimation method (maximum-likelihood method) is different from neural network (back propagation)

ARTIFICIAL NEURAL NETWORKS

- Normalization
  - Scale of [0,1] is typically recommended for neural network models for performance purposes
  - For numeric variables,

$$V_{norm} = \frac{V - \min(V)}{\max(V) - \min(V)}$$

ARTIFICIAL NEURAL NETWORKS

- Normalization
  - Binary variables (categorical variables with two classes)
    - Create dummy variables: set of values {0, 1}
  - Nominal variables with m (>2) classes
    - Create m-1 dummy variables: set of values {0, 1}
  - Ordinal variables with m (>2) classes
    - Map the values to the set {0, 1/(m-1), 2/(m-1), …, (m-2)/(m-1), 1}

Key References

- Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data by EMC Education Services (2015)
- Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner by Shmueli, G., Patel, N. R., & Bruce, P. C. (2010)

Thanks…