Data Mining II: Advanced Methods and Techniques

Assignment #1: Artificial Neural Networks

Please provide sufficient support so that your models could be reproduced and provide enough discussion to illustrate your understanding of the methods and concepts

1. What is the difference between artificial and biological neuron?
2. What is a perceptron?
3. What is the basic structure of a biological neuron and how does it relate/differ to/from the McCulloch-Pitts neuron? What is the main difference between McCullogh-Pitts neuron and general artificial neuron?
4. Describe how does Gradient Descent works
5. Describe how does Backpropagation works
6. What is the purpose of learning rate and momentum in back-propagation learning? Briefly explain how this purpose is achieved.
7. Use either the Letter dataset that can be found under Resources section of the Blackboard. Build the most accurate neural network for predicting the class value using Weka's neural network classifier. Experiment with at least three different ANN configurations (# of hidden layers and nodes), varying the learning rate and number of training epochs of your choice. Describe all of your experiments in enough detail that they can be reproduced.
8. Built the most accurate neural network from a dataset of your choice available via UCI KDD archive: http://kdd.ics.uci.edu/summary.data.alphabetical.html or Kaggle: https://www.kaggle.com/datasets or UCI Machine Learning Repository: http://archive.ics.uci.edu/ml/. Preprocess the data. Conduct different neural network experiments by modifying the neural network topology, the learning rate, the number of epochs to train, etc. Describe how you came to the model you decided to keep. Report your findings in enough detail that they can be reproduced.