

**Mid Semester Examination**

**Course Name: Artificial Intelligence**  
**Full Marks-40**

**Code: CS 561**  
**Time: 2 hours**

*Answer ALL the questions*

*Make reasonable assumptions as and whenever necessary. You can answer the questions in any sequence. However, the answers to all the parts of any particular question should appear together.*

**Q1.** For document classification, occurrences of words are important'- verify this claim with the help of an appropriate naive Bayes classification model and necessary examples. State each step of the algorithm, compute the parameters and derive the training and test time complexities of the algorithm.

**12**

**Q2.** "Selecting attribute sequence in decision tree has significant effect on its induction and deduction processes. There is no universal way of splitting the data, rather it depends upon the problem at hand".

Discuss the above statement using proper examples, and any TWO of the three objective measures like "Information Gain", "Misclassification Error" and "Gini Index". Explain each step very clearly.

**8**

**Q3.** While training neural network with backpropagation you observe that the error does not decrease any further. Enumerate ALL possible reasons for this behavior and the solution for each such cause.

**7**

**Q4.** "If the constraint of pure feed-forward network is removed, then one can save in the number of neurons in the network". Justify this statement with an example. "Pure feedforward network" means, the connections can go from any layer to only the next upper layer towards the output layer and no other layer.

**6**

**Q5.** Assume that the error function in a feedforward-BP network is the KL-divergence given by

$$D_{KL}(Q \| T) = \sum_{i=1}^N Q(i) \log \frac{Q(i)}{T(i)}$$

Where  $T$  is the target distribution of neuron values on the output layer,  $Q$  is the observed distribution, and  $i$  varies from 1 to  $N$  (no. of output neurons). This means each of  $T(i)$  (and  $P(i)$ ) is a number from 0 to 1, and their sum is equal to 1. Assuming sigmoid neurons, give the weight change rules in the network both for output layer and hidden layer neurons.

**3+4= 7**