NAME : MD.ASIF HOSSEN

ROLL: 201831052061

SUBJECT: PATTERN RECOGNITION

Answer the question no:1

DECISION TREE:Decision Trees (DTs) are a non-parametric supervised learning method used for classification and regression. Decision trees learn from data to approximate a sine curve with a set of if-then-else decision rules. The deeper the tree, the more complex the decision rules and the fitter the model.Decision tree builds classification or regression models in the form of a tree structure. It breaks down a data set into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes. A decision node has two or more branches. Leaf node represents a classification or decision. The topmost decision node in a tree which corresponds to the best predictor called root node. Decision trees can handle both categorical and numerical data.

BAYESIAN:Bayesian inference is a method of statistical inference in which Bayes' theorem is used to update the probability for a hypothesis as more evidence or information becomes available. Bayesian inference is an important technique in statistics, and especially in mathematical statistics.

NEURAL NETWORK: A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature.

VERSUS LAZY CLASSIFICATION:Eager classification is seen to be much faster than the lazy classification. ... Lazy classification on the other hand involves a better hypothesis level this improves accuracy of classification . in this classification trading tuples have to be stored increasing costs and involved high end indexing structures. I wish to better understand the difference between lazy and eager learning. I am having difficulty conceptualising what the "abstraction" refers to between the two.

According to the text book I am reading it says, "The distinction between easy learners and lazy learners is based on when the algorithm abstracts from the data."A lazy learner delays abstracting from the data until it is asked to make a prediction while an eager learner abstracts away from the data during training and uses this abstraction to make predictions rather than directly compare queries with instances in the dataset.I understand that KNN algorithm loads all the data into memory so depending on the size of the dataset, the computing requirements could be enormous as well as the fact it's a non-parametric algorithm so there is no training to find parameters, the classification is done on a per query requirement.

ANSWER THE QUESTION NO 2:

Genetic algorithm: Genetic algorithms have been proved to be quite effective in solving certain optimization and artificial intelligence (AI) problems. They have been used in many application areas, including pattern recognition. However, the applications of genetic algorithms in pattern recognition have concentrated primarily on training neural networks for pattern recognition (Montana 1989, Whitley 1992, Kitano 1994). The research in this paper is aimed at using a genetic algorithm to perform pattern matching directly. The basic idea is to use a genetic algorithm to find the best match between nodes of the two patterns. An objective function can be defined in terms of the total difference in the magnitudes of angles between the corresponding edges of the two patterns. Experiments designed to evaluate the algorithm have shown very promising results with high accuracy in classifying the input patterns.

Rough setes algorithm: This article provides an overview of recent literature on some tasks of pattern recognition using rough sets and its hybridization with other soft computing paradigms. Rough set theory is an established tool for dealing with imprecision, noise, and uncertainty in data. In this article we will focus on two recent applications using rough sets; *viz.*, feature selection in high dimensional gene expression data, and collaborative clustering. The experimental results demonstrate that the incorporation of rough set improves the performance of the system.

Keywords

Feature Selection Discernibility Matrix Cluster Prototype Indiscernibility Relation Redundancy Reduction.

*These keywords were added by machine and not by the authors. This process is experimental and the keywords may be updated as the learning algorithm improves.*

*FuZZY SETS:*

Fuzzy Set Theory is concerned with non-dichotomous structures, *i.e*., with situations which are not of the either-or-structure, which cannot be characterized by black-or- white, by true-or-false, by certain-or-impossible. These situations are generally considered as containing uncertainty, even though the term “uncertainty” is not defined unequivocally. Often uncertainty is considered to refer to the occurrence of events while vagueness refers to the description of events. Let us consider uncertainty with respect to occurrence first: The statement “The probability of hitting the target is.6” is certainly probabilistic in nature and could well be modelled using classical probability theory. “The chances of winning are good” will already pose some problems, because “good chances” are not well (*i.e*., crisply) defined. In a phrase like “It is likely that we will make a good profit” the event itself nor its occurrence are well defined. In these situations classical probabilistic models will not even be appropriate for the expression of uncertain occurrences.

Keywords

Membership Function Fuzzy Cluster Fuzzy Relation Cluster Validity Fuzzy Cluster Algorithm

*These keywords were added by machine and not by the authors. This process is experimental and the keywords may be updated as the learning algorithm improves.*