Week 10: Daily Morning Challenge

Day 2: Thursday 12th March 2020

Question 1: Describe the concept of connectionist computing approach as regards biological neural networks

Connectionism is a movement in cognitive science that hopes to explain intellectual abilities using artificial neural networks (also known as "neural networks" or "neural nets"). Neural networks are simplified models of the brain composed of large numbers of units (the analogs of neurons) together with weights that measure the strength of connections between the units. These weights model the effects of the synapses that link one neuron to another. Experiments on models of this kind have demonstrated an ability to learn such skills as face recognition, reading, and the detection of simple grammatical structure.

Connectionist networks are made up of interconnected processing units which can take on a range of numerical activation levels (for example, a value ranging from 0 – 1). A given unit may have incoming connections from, or outgoing connections to, many other units. The excitatory or inhibitory strength (or weight) of each connection is determined by its positive or negative numerical value. One common sort of connectionist system is the two-layer feed-forward network. In these networks, units are segregated into discrete input and output layers such that connections run only from the former to the latter. Often, every input unit will be connected to every output unit, so that a network with 100 units, for instance, in each layer will possess 10,000 inter-unit connections. Let us suppose that in a network of this very sort each input unit is randomly assigned an activation level of 0 or 1 and each weight is randomly set to a level between -0.01 to 0.01.

Question 2: Illustrate with a table the benefits and limitations of the following data mining techniques

Linear Regression

Benefits	Limitations
It is very easy and intuitive to use and understand	The algorithm assumes data is normally distributed in real they are not.
Being a parametric algorithm, Linear Regression can work well when the data available is very limited	Prone to outliers and over fitting
Advanced techniques of regularisation can be easily used for feature selection, which improves model interpretability drastically	Linear regression only models relationships between dependent and independent variables that are linear. It assumes there is a straight-line relationship between them which is

	incorrect sometimes.

Multivariate Regression

Benefits	Limitations
The ability to determine the relative influence of one or more predictor variables to the criterion value.	Collinearity or multicollinearity.
The ability to identify outliers, or anomalies	
Multiple regression model allows us to examine the causal relationship between a response and multiple predictors	

Concept learning (List-then-Eliminate)

Benefits	Limitations
A good concept learning algorithm should be able to backtrack the choice of hypothesis found so that the resulting hypothesis can be improved over time.	Inconsistent sets of training examples can mislead the finds algorithm as it ignores negative data samples, so an algorithm that can detect inconsistency of training data would be better to use.
many concept learning algorithms organize the search through the hypothesis space by relying on the general-to-specific ordering	No way to determine if the only final hypothesis (found by Find-S) is consistent with data or there are more hypothesis that is consistent with data.
description of the set of all hypotheses consistent with D without explicity enumerating them	performs poorly with noisy data

Decision tree

Benefits	Limitations
Are simple to understand and interpret. People are able to understand decision tree models after a brief explanation.	They are unstable, meaning that a small change in the data can lead to a large change in the structure of the optimal decision tree
Help determine worst, best and expected values for different scenarios.	Calculations can get very complex, particularly if many values are uncertain and/or if many outcomes are linked.
Can be combined with other decision techniques.	They are often relatively inaccurate. Many other predictors perform better with similar data

Naive Bayes

Benefits	Limitations

. When assumption of independent predictors holds true, a Naive Bayes classifier performs better as compared to other models.	Main imitation of Naive Bayes is the assumption of independent predictors. Naive Bayes implicitly assumes that all the attributes are mutually independent. In real life, it is almost impossible that we get a set of predictors which are completely independent
Naive Bayes requires a small amount of training data to estimate the test data. So, the training period is less.	If categorical variable has a category in test data set, which was not observed in training data set, then model will assign a 0 (zero) probability and will be unable to make a prediction. This is often known as Zero Frequency
Naive Bayes is also easy to implement.	

Artificial Neural Networks

Benefits	Limitations
The ability to work with inadequate knowledge: After ANN training, the data may produce output even with incomplete information. The lack of performance here depends on the importance of the missing information.	Hardware dependence: Artificial neural networks require processors with parallel processing power, by their structure. For this reason, the realization of the equipment is dependent.
It has fault tolerance: Corruption of one or more cells of ANN does not prevent it from generating output. This feature makes the networks fault-tolerant.	The difficulty of showing the problem to the network: ANNs can work with numerical information. Problems have to be translated into numerical values before being introduced to ANN. The display mechanism to be determined here will directly influence the performance of the network. This depends on the user's ability.
Parallel processing ability: Artificial neural networks have numerical strength that can perform more than one job at the same time.	Assurance of proper network structure: There is no specific rule for determining the structure of artificial neural networks. The appropriate network structure is achieved through experience and trial and error.