Theory Assignment Week 4

Q VC Dimension
Find VC dimension for following models :

i) logistic Regression
ii) SVM
iii) Decision Tree
iv) Neural Net

i) logistic Regression
for logistic regression we have real-valued functions for which we can find VC dimension with the help of VC dimension for indicator functions. The VC dimension for a class of real-valued functions $f(x, a)$ is defined to be VC-dim of indicator class $(f(x,a) - \alpha) > 0$

Hence we can find VC-dim of real-valued function by finding VC-dim of the class of indicator function that can be formed by thresholding that class of real-valued function.

ii) VC dimension for SVM
for SVM ; VC dim is equal to enumerative combinatorics of the num of nominals for a polynomial of degree k and n variables
thus $\binom{n+k}{k}$

PAGE No
DATE / /

iii) Decision Tree

Finding VC-dim of decision tree has been proposed in the research paper published in the year 2015 titled: "On the dimension of Univariate Decision Tree". Author has defined for and proved various theorems regarding the VC dim.

Theorem 1 : VC-dim of degenerate univariate decision tree with N nodes that classifies dimensional data is at least $\lfloor \log_2 (d-N+2) \rfloor + N$

Theorem 2 : The VC-dim of univariate decision tree with binary features that classifies 'd' dim of its left & right subtree of those classifying d-1 dimensional data.

Theorem 3 : The VC-dim of single node L-ary decision tree that classifies 'd' dimensional data is $\lfloor \log_2 (\Sigma (2^{L(i-1)}_{-1}) + 1 \rfloor + 1$

Theorem 4:

The VC-dim of L-ary decision tree that classifies 'd' dimensional ~~tree~~ data is at least the sum of the VC-dim of its subtree those classifying d-1 dimensional data.

iv) VC-dimension for Neural Network

A network with 'W' weights & L layers & relu activations ; its VC-dimension can be defined as

$$c W L \log(W/L) \leq V.C \leq c W L \log(W.L)$$

for constant 'c'.