LVC 1 - Glossary of Notations

$\mathcal{X}$ = A vector of categorical data

y = Outcome class (categorical)

f: $\mathcal{X} \rightarrow \mathrm{y}$ = Decision Rule i.e. f is a function that is mapping the independent features with the target values

$\boldsymbol{x}_{\mathrm{i}}$ = i$^{\text{th}}$ row of the vector $\mathcal{X}$

y_{i} = i$^{\text{th}}$ row of vector y

N = Natural number

$\in$ = Belongs to

Σ =The summation

$\neq$ = Not equal to

$R(f)$ = Empirical Error (generalization error) of a Decision Rule

$R^*(f)$ = Probabilistic Error of a Decision Rule

$\frac{1}{N}\sum_{i}^{N} I(f(\boldsymbol{x}_{\mathrm{i}}) \neq y_{\mathrm{i}})$ = The average number of misclassifications. The $I()$ function is 1 in case of a misclassification and 0 otherwise

C = It is a sub class that maps with a function f to predict the y along with probability maximization

k = subset of all feature indices in the sub class

Z = Random Variable

$X,\ Y$ = X,and Y are the random variables

$P(Z)$ = Z is a random variable with the probability mass function

E = Expectation value

$P(x, y)$ = It represents the joint distribution of X and Y

$H(Z)$ = Entropy

$H(X, Y)$ = Joint Entropy of random variables X and Y

$H(X|\ Y)$ = Conditional Entropy of Y given X

$IG(Y \mid X)$ = Information Gain of Y given X

$X \perp Y$ = X is perpendicular to Y

$X(m)$ = a feature from the X

S_1 = {(y_i | $\boldsymbol{x}_i$ (m) = 0} = Splitting outcome based on class 0

S_2 = {(y_i | $\boldsymbol{x}_i$ (m) = 1} = Splitting outcome based on class 1