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 to the target values

x_i = i^{th} row of the vector $\mathcal{X}$

y_i = i^{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

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

C = It is a subclass of data points

k = Subset of all feature indices in the subclass

Z = Random Variable

$X,\ Y$ = X represent the independent features and Y represents the target feature

$P(Z)$ = Probability mass function of the random variable Z

E = Expected value

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

$H(Z)$ = Entropy of Z

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

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

$IG(Y \mid X)$ = 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 | x_i (m) = 0} = Splitting outcome based on class 0

S_2 = {(y_i | x_i (m) = 1} = Splitting outcome based on class 1