by identifying anomalous sequences. Different HMM models were investigated in terms of the dictionary of symbols, number of hidden states, and different training sets. It was found that good performances can be attained by using dictionary of symbols made up of all symbols in the training set, and adding a NaS (not-a-symbol) symbol in account of symbols in the test set that are not represented in the training set. Performances can be further improved by combining different HMM. As the size of training sets in an intrusion detection application is typically large, it was proposed to split the training set in a number of parts, training different HMM and then combining the output probabilities by three well known combination.

2.2 Naive Bayes Classifier

A Naïve Bayes Classifier is a simple probabilistic classifier which uses the Bayes' theorem with an assumption that the occurrence of an event is totally unrelated to the occurrence of another.

Despite their naive design and apparently oversimplified assumptions, Naive Bayes classifiers have worked quite well in many complex real-world situations. In 2004, an analysis of the Bayesian classification problem showed that there are sound theoretical reasons for the apparently implausible efficacy of naive Bayes classifiers.[14]

2.2.1 Bayes Theorem

Bayes theorem is used by the Naïve Bayes Classifier. It is particularly suited when the dimensionality of the inputs is high. The formula for the Bayes theorem is represented by the figure below.

$$P(W|L) = \frac{P(L|W)P(W)}{P(L)} = \frac{P(L|W)P(W)}{P(L|W)P(W) + P(L|M)P(M)}$$

where,

$P(W|L)$ →Probability of outcome W given L or "posterior probability"

$P(L|W)$ →Probability of outcome L given W or "likelihood"

$P(W)$ →Independent probability of W

$P(L)$ →Independent probability of L or "prior probability"