International J. of Engg. Research & Indu. Appls. (IJERIA).
ISSN 0974-1518, Vol.6, No. II (May 2013), pp 169-190

UTILIZATION OF DATA MINING TECHNIQUES FOR PREDICTION AND DIAGNOSIS OF AIDS/HIV DISEASE SURVIVABILITY

K. R. LAKSHMI[1], M. VEERA KRISHNA[2] AND S. PREM KUMAR[3]
[1] Research Scholar, Department of CSE, CMJ University, Shillong, Meghalaya, India
[2] Department of Mathematics, Rayalaseema University, Kurnool, Andhra Pradesh, India
[3] Professor & Head, Department of CSE&IT, G.Pullaiah college of Engineering & Technology, Nandikotkur Road, Kurnool, Andhra Pradesh, India.

Abstract

Human immunodeficiency virus (HIV) is a lent virus that causes acquired immunodeficiency syndrome (AIDS). The main drawback in HIV treatment process is its sub type prediction. The sub type and group classification of HIV is based on its genetic variability and location. HIV can be divided into two major types, HIV type 1 (HIV-1) and HIV type 2 (HIV-2). Many classifier approaches have been used to classify HIV subtypes based on their group, but some of cases are having two groups in one; in such cases the classification becomes more complex. The methodology used is this paper based on the HIV sequences. For this work several classifier approaches are used to classify the HIV1 and HIV2. For implementation of the work a real time patient database is taken and the patient records are experimented and the final best classifier is identified with quick response time and least error rate. A typical confusion matrix is furthermore displayed for quick check. The study describes algorithmic discussion of the HIV dataset from AVERT's Community database, on line repository of large datasets. The Best results are achieved by using Tanagra tool. Tanagra is data mining matching set. The accuracy is calculate based on addition of true positive and true negative followed by the division of all possibilities.

Keywords : HIV, SVM, ID3, LDA, C4.5, *k*-NN, BLR, MLR, PLS-DA, *k*-means and Apriori, Sensitivity and Specificity.