Abstract

Botswana was one of the first countries to establish a national Antiretroviral therapy (ART) programme in Africa. “Masa,” a native word depicting “a new dawn,” is the programme name. The AIDS epidemic has led to the emergence of several disease entities which in the pre-AIDS era seemed innoxious. Some HIV AIDS positive patients under the ART programme continue to be at risk of contracting related Opportunistic Infections (OIs) and little evidence based research work has been carried out so as to apply preventative or mitigating factors. The use of Data Mining (DM) techniques is becoming more popular for investigating subtle relationships in Clinical data. This paper proposes to build a robust classification and prediction model by mining the historical data stored in a data-warehouse to determine which patients might be at risk of contracting the afore-mentioned infections. Four supervised learning algorithms viz Generalized Linear Model (GLM), Support Vector Machine (SVM), Decision Tree (DT) and Naïve Bayes (NB) were used for building the models. Their performances were analyzed and evaluated for their efficacy against the Confusion Matrix analytical performance, the Receiver Operating Characteristic (ROC) curve, the LIFT Cumulative and Profit analysis. Experimental results proved that the SVM exhibited superior performance and was therefore
deployed in building the HIV related OI prediction model.

References

- Data Mining: Know It All By Soumen Chakrabarti, Earl Cox, Eibe Frank, Ralf Hartmut Güting, Jiawei Han, Xia Jiang, Micheline Kamber, Sam S. Lightsome, Thomas P. Nadeau, Richard E. Neapolitan, Dorian Pyle, Mamdouh Refaat, Markus Schneider, Toby J. Teorey, Ian H. Witten, 2009.
- Data Mining Know It All by AMSTERDAM • BOSTON • HEIDELBERG • LONDON NEW YORK • OXFORD • PARIS • SAN DIEGO SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO Morgan Kaufmann is an imprint of Elsevier et al. 2009.
- Granular Support Vector Machines for Medical Binary Classification Problems by Yuchun Tang, Bo Jin, Yi Sun, and Yan-Qing Zhang, Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2004.
- I. Rish, An empirical study of the naive Bayes classifier, T. J. Watson Research Center.
- Oracle® Data Mining Concepts 11g Release 2 (11. 2) E16808-06 July 2011.
- Granular Support Vector Machines for Medical Binary Classification Problems by Yuchun
Tang, Bo Jin, Yi Sun, and Yan-Qing Zhang, Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2004.

- City Innovative Capability Prediction Based on Support Vector Machine: Taking Thirteen Chinese Cities as the Example by Zhao Jing School of Economics and Management and Guo Hai-xing School of Management Xi'an University of Technology, XUT Xi'an, China. 2010 IEEE.

Index Terms

Computer Science Artificial Intelligence

Keywords

Antiretroviral Therapy Generalized Linear Model Support Vector Machine
Decision Tree and Naïve Bayes