



ELSEVIER

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

A comprehensive identification-evidence based alternative for HIV/AIDS treatment with HAART in the healthcare industries

You-Shyang Chen *

Department of Information Management, Hwa Hsia University of Technology, 111, Gongzhuan Rd., Zhonghe Dist., New Taipei City 235, Taiwan

ARTICLE INFO

Article history:

Received 29 September 2015

Received in revised form

3 March 2016

Accepted 1 April 2016

Keywords:

Human immunodeficiency virus (HIV)

Acquired immune deficiency syndrome (AIDS)

Highly active anti-retroviral therapy (HAART)

Linear–nonlinear feature selection

Hybrid model

ABSTRACT

Background and Objective: The HIV/AIDS-related issue has given rise to a priority concern in which potential new therapies are increasingly highlighted to lessen the negative impact of highly active anti-retroviral therapy (HAART) in the healthcare industry. With the motivation of “medical applications,” this study focuses on the main advanced feature selection techniques and classification approaches that reflect a new architecture, and a trial to build a hybrid model for interested parties.

Methods: This study first uses an integrated linear–nonlinear feature selection technique to identify the determinants influencing HAART medication and utilizes organizations of different condition-attributes to generate a hybrid model based on a rough set classifier to study evolving HIV/AIDS research in order to improve classification performance.

Results: The proposed model makes use of a real data set from Taiwan’s specialist medical center. The experimental results show that the proposed model yields a satisfactory result that is superior to the listed methods, and the core condition-attributes PVL, CD4, Code, Age, Year, PLT, and Sex were identified in the HIV/AIDS data set. In addition, the decision rule set created can be referenced as a knowledge-based healthcare service system as the best of evidence-based practices in the workflow of current clinical diagnosis.

Conclusions: This study highlights the importance of these key factors and provides the rationale that the proposed model is an effective alternative to analyzing sustained HAART medication in follow-up studies of HIV/AIDS treatment in practice.

© 2016 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

The acquired immune deficiency syndrome (AIDS) is the most severe expression of a spectrum of related disorders due to infection by the human immunodeficiency virus (HIV) [1,2]. HIV infections are caused by two major types of retrovirus: HIV type 1 (HIV-1) and HIV type 2 (HIV-2) [3]. HIV-1 originated in

West-Central Africa in the first half of the 20th century and has progressed to become the most widespread AIDS disease due to its lengthy 10 to 12 years latent period. HIV-2 mainly infects less developed countries. HIV causes AIDS, and the HIV virus infects T-helper lymphocytes (i.e., CD4, cluster of differentiation 4, which is a glycoprotein found on the surface of immune cells). The HIV infection entry is a complex and intricate process. The only viral surface protein, Envelope (Env),

* Department of Information Management, Hwa Hsia University of Technology, 111, Gongzhuan Rd., Zhonghe Dist., New Taipei City 235, Taiwan. Tel.: +886 2 8941 5100; fax: +886 2 2941 5730.

E-mail address: ys_chen@cc.hwh.edu.tw

<http://dx.doi.org/10.1016/j.cmpb.2016.04.001>

0169-2607/© 2016 Elsevier Ireland Ltd. All rights reserved.

is composed of a trimer of glycoprotein 120 (gp120, a viral envelope protein) and gp41 (another viral protein) heterodimers. HIV-1 binds with glycoprotein 120 to CD4 T cells and uses gp41 to enter the host T-cells; consequently, the binding to CD4 T cells shifts the conformation of gp120 to allow HIV-1 to bind to co-receptors (two chemokine co-receptors, either CCR5 or CXCR4) expressed on the host cells, and the gp41 sub-unit starts fusion of the virion into the target cell after disassociation of the gp120 [4]. The gp41 includes three domains: an ectodomain (ECD), a transmembrane domain (TMD), and a cytoplasmic domain (CTD), and the gp41 also mediates the processing of the gp160 intermediate as well as the gp120/gp41 complex formation [5]. Furthermore, the HIV viral proteins contain the matrix (MA), capsid (CA), and nucleocapsid (NC) proteins [6], and HIV is spread between CD4 T cells by the three ways of cell-free spread, cell-to-cell spread, and hybrid spreading mechanisms against antiretroviral therapies [7]. In practice, five types of treatments are available to HIV/AIDS patients, including type 1: nucleoside reverse-transcriptase inhibitors (NRTIs), type 2: non-nucleoside reverse-transcriptase inhibitors (NNRTIs), type 3: protease inhibitors (PIs), type 4: fusion inhibitors (glycopeptides)-gp120/gp41 (CCR5 or CXCR4), and type 5: integrase inhibitors.

Since the cause of HIV infection was first identified in early 1984, HIV/AIDS [8,9] has become one of the most devastating epidemics. Based on the literature [10], the transmission sources of HIV primarily include unprotected sexual intercourse, hypodermic needles, contaminated blood transfusions, and disease transmission from mother to child during pregnancy, delivery, or breastfeeding. In Taiwan, the two primary sources are homosexual acts and drug injections, and few people have acquired HIV/AIDS via blood transfusion. This suggests that the majority of those affected [11] are male patients. Importantly, the increasing trend of HIV/AIDS patients is still an advanced persistent threat to Taiwan's regulators. Thus, HIV/AIDS issues have had a great negative societal impact as both an illness and a source of discrimination, and these issues also have significant adverse economic effects; HIV/AIDS patients will exhaust large amounts of medical resources and increase the national healthcare burden. Thus, prevention of HIV infection is a key strategy to control the spread of disease and to avoid wastage of medical resources. This issue has attracted international medical and political attention and large-scale funding over time [12]. Considerable studies and practices have further formulated HIV treatment guidelines and established clinical prognostic models for tracing the natural history and curative effects of HIV disease [13]. A recent method, known as highly active anti-retroviral therapy (HAART) [14–16], was proposed in 1995 and has performed well in treating HIV opportunistic infections to improve the patient's condition. Analytical studies on HAART revealed that the best time to start the therapy on infected HIV-1 patients is when the measurement of CD4 lymphocytes is over 350 cells/ml, which will result in a high survival rate. There are many key factors that have major effects on the clinical treatment of HIV/AIDS [17]. Thus, complete differentiation of the determinants is an interesting issue. Such meaningful work motivates this study.

In Taiwan, there is a lack of medical analysis and curative effects of HIV/AIDS disease. This shortage may cause inaccurate reactions to the clinical features of AIDS patients, and HIV/AIDS physicians only passively quote literature information on

examples studied from abroad. In addition, medical works have encountered the problem of many spurious data features, and the determination of their feature properties as well as medical activities is a time-consuming and costly process and is lagging severely behind [18]. The originated medical data may contain demographic information (registration files), original claim data for reimbursement, and information on health care utilization, across inpatients, outpatients, and physicians, such as inpatient expenditures by admissions records, registry for contracted medical facilities, registry for board-certified specialists, and details of inpatient orders. These data are high-dimensional and de-identified by scrambling the identification codes of both patients and medical facilities to form the original files of the database used for some specific purposes. These complex non-homogeneous data make it complex to apply the data mining tools used; thus, suitable feature selection methods are necessary in an effort to lower the data dimensionality. Feature selection techniques have been developed and can be simply classified as linear and nonlinear models [19]. The conventional linear method of statistics with data projections is easy to implement and has shown its effectiveness for data visualization and classification in a wide variety of application fields. Recently, numerous nonlinear methods of machine learning have been proposed to properly handle data with complex nonlinear structures; the family of nonlinear algorithms has attracted much attention because it exhibits noteworthy performance in many real-world data sets. On the one hand, the linear feature selection methods are an example of a linearly predictive feature. Swets and Weng [20] proposed a so-called content-based query technique of automatic feature selection for a hierarchical image database at the object recognition level. They demonstrated the applicability of the approach even for large variability within a particular object class. Cheng et al. [21] applied two common feature selection approaches, called data-driven automatic feature selection mechanisms and knowledge-driven expert judgment on a risk prediction of cardiovascular disease data set. Their empirical results suggested that the automatic feature selection mechanism improves the predictive power of a classifier on the majority class. Fauvel et al. [22] used a preprocessing step to extract relevant feature in hyperspectral images analysis for further classification and to prevent the Hughes phenomenon and obtain a satisfactory result. Kuo and Chang [23] proposed a regularized feature extraction (RFE) method and illustrated the importance of the hyperspectral image experiment for a small sample size classification problem, and the best feature extraction was by RFE with nonparametric weighted scatter matrices. Dhir et al. [24] compared information gain with the Fisher criterion for recognition of facial images with different illumination and proved that the information gain criterion gives a slight performance improvement when ICA-based features are used. Xu et al. [25] applied the F-score method of statistics to provide an effective solution for a good accelerating segmentation result on the liver. In the study by Sharma and Paliwal [26], the linear discriminant analysis (LDA) technique was proven to be a very popular method in pattern recognition for the reduced feature space and significantly improved the classification performance. On the other hand, nonlinear feature selection methods usually provide better results and more cost-effective identifiers for a nonlinear cause-effect relationship than linear feature

Download English Version:

<https://daneshyari.com/en/article/469072>

Download Persian Version:

<https://daneshyari.com/article/469072>

[Daneshyari.com](https://daneshyari.com)