Short Communication Journal of Diabetes & Metabolism 2020
Vol.11 No.12

Support Vector Machine Analysis Leads to a Proposal of Cost-Effective Two-Step Screening Strategy for Metabolic Syndrome in Low-Income Countries

Wen Jie Zhang

CHS, China

Machine Learning is reflected as a subfield of Artificial Intelligence and it is concerned withdevelopment of methods which allow the computer to learn. Development of algorithms also enable the machine to learn and perform tasks and activities. Machine learning overlaps with statistics in many ways. Many techniques and methodologies were advanced for machine learning works. Support vector machines (SVMs) are a set of related controlled learning methods used for ordering and reversion. They fit in to a set of generalized linear classifiers. SVM is a classification and regression prediction device that uses machine learning theory to make the most of predictive accuracy while automatically to avoid data over-fit. SVMs are the systems which use theory space of linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias resulting from statistical learning theory. Support vector machine was firstly popular with the NIPS community and now is a part of the machine learning research globally. SVM are great when using pixel maps as input; it gives comparable accuracy to sophisticated neural networks with expanded features in a handwriting recognition task. This is used for many applications, such as face analysis, hand writing analysis and specially for pattern classification and regression based applications. The formulation uses the Structural Risk Minimization (SRM) principle to traditional Empirical Risk Minimization (ERM) principle, used by conventional neural networks. SRM minimizes an upper bound on the expected risk, whereas ERM minimizes the error on the training data. It is the difference which prepares SVM with a better ability to generalize, which is the goal in statistical learning. Support Vector Machines were developed to solve the classification problem, recently they have been extended to solve regression problems.

Metabolic syndrome (MS) is common in low-income communities where primary healthcare facilities often have no biochemical blood tests. According to the Joint-Interim-Statement (JIS) criteria, MS is diagnosed if there present a combination of 3 abnormal findings out of 5 bio-indicators including 2 non-blood indicators of waist circumference (WC) and blood pressure (BP) and 3 blood indicators of fasting blood glucose (GLU), triglycerides (TG), and high-density lipoprotein cholesterol (HDL-C). Support Vector Machine (SVM) is a computer learning algorism capable of recognizing combinational patterns and differentiating them by sensitivity and specificity. Employing SVM's ability of recognizing the best combinations of bio-indicators, this study aimed at establishing a sensitive and cost-effective screening strategy for MS in low-income communities. From 2012-2014 in a typical low-income rural township in China's far-western Xinjiang Province, 3,276 individuals (1,590 males and 1,686 females) aged ≧18 years without prior diagnoses of MS were physically examined and blood biochemistry tested. MS was first diagnosed as an end result among these individuals based on the JIS criteria. Following the Support vector machines analysis revealed that two non-blood bio-indicators, WC and BP together, were able to detect MS individuals with 57.0% sensitivity and 88.4% specificity. We then used SVM algorism to analyze abilities of diagnosing MS by WC+BP+TG, WC+BP+HDL-C, and WC+BP+GLU combinations. WC+BP+TG added the finest sensitivity of 72.5% with 88.1% specificity; WC+BP+HDL-C had a lesser sensitivity of 58.6% with 90.6% specificity; and WC+BP+GLU showed the least sensitivity of 56.6% with 90.9% specificity, respectively. Based on these observations, we propose a cost-effective "two-step screening strategy" for MS in low-income developing countries by which the "first screening step" is to use only non-blood indicators, WC+BP, which can be feasibly examined by paramedical personnel in "village/community clinic" and, individuals with abnormal WC+BP are then recommended to go through the "second screening step" of blood testing in higher level hospitals. We recommend TG and HDL-C, other than GLU, as the first-line blood tests for low-income populations.

This work is partly presented at 28th European Diabetes Congress, June 10-11, 2019, Edinburgh, Scotland