ORIGINAL ARTICLE

Unveil the Features Influencing Hypertension Adults in Malaysia using Machine Learning Models

Ridwan Sanaudi[1], Zainul Azhar Zakaria[2], Aiman Adlan Khairulisam[3], Nurain Ibrahim[3,4], Ahmad Zia Ul-Saufie[3]

[1] Sector of Biostatistics and Data Repository, National Institutes of Health, Ministry of Health Malaysia, 40100 Shah Alam, Selangor, Malaysia
[2] Foundation in Medical Sciences, Faculty of Medicine, Royal College of Medicine Perak, Universiti Kuala Lumpur, 30450 Ipoh, Perak, Malaysia
[3] School of Mathematical Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[4] Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Kompleks Al-Khawarizmi, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia.

ABSTRACT

Introduction: The number of people affected by hypertension is staggering, with an estimated one billion people living with the disease worldwide. It has been shown that machine learning (ML) models surpass clinical risk; nevertheless, there isn't much research using ML to predict hypertension in Malaysia. **Materials and methods:** A study is being conducted using ML analyses to predict hypertension using secondary data from population-based surveys, such as the National Health & Morbidity Survey (NHMS) 2015. The dependent or target variable was hypertension status and 24 features. Three standard ML-based classifiers, which are logistic regression (LR), decision tree (DT) and artificial neural network (ANN), were used to predict hypertension and the associated factors that influence hypertension were obtained from filter-based feature selection, which are feature weight by information gain, feature weight by information gain ratio and feature weight by correlation. **Results:** Out of 11,520 respondents, 4,175 (36.24%) adults had hypertension. LR is the best model to predict hypertension since LR has the highest accuracy (76.73%) compared to DT and ANN (73.02%). In terms of odd ratio explanation, a person who does not have diabetes mellitus is 2.05 odds likely to have hypertension, and a person who does not have hypercholesterol has 1.67 odds of having hypertension, and with an increase in the age of adults, 6.0% are less likely to have hypertension. **Conclusion:** From LR model, the essential features that influence hypertension in adults were diabetes mellitus, hypercholesterolemia status, age, waist circumference, marital status, occupation, education, and total household income.
Malaysian Journal of Medicine and Health Sciences (2024) 20(6): 167-174. doi:10.47836/mjmhs20.6.22

Keywords: Artificial neural network, Decision tree, Hypertension, Logistic regression, Machine learning, Prediction

Corresponding Author:
Nurain Ibrahim, PhD
Email: nurain@tmsk.uitm.edu.my
Tel : +6011-39328310

INTRODUCTION

The disorder known as hypertension, or high blood pressure, is characterized by an excessively high blood pressure against the arterial walls. Two numbers are used to represent blood pressure. The first (systolic) number represents the blood vessel pressure during a heartbeat or contraction. The second (diastolic) number indicates the pressure in the arteries during the heart's resting heartbeat. If a person's diastolic blood pressure is 90 mmHg or higher on two different occasions, and

their systolic blood pressure is 140 mmHg or higher, they are diagnosed with hypertension (1). Risk assessment for hypertension, or high blood pressure, can be quite complex and depends on a variety of factors. These can include genetics, lifestyle choices, and environmental conditions that can contribute to an increase in blood pressure readings. The World Health Organization (WHO) reported that high blood pressure is a leading cause of death worldwide, responsible for one in every eight deaths (2). It is considered the third leading cause of death globally. The number of people affected by hypertension is staggering, with an estimated one billion people worldwide living with the condition.

Malaysia's Ministry of Health (MOH) conducted the National Health and Morbidity Survey (NHMS). The