Advances and Applications in Statistics

https://pphmjopenaccess.com
https://doi.org/10.17654/0972361725045
Volume 92, Number 7, 2025, Pages 1023-1030
P-ISSN: 0972-3617
E-ISSN: 3049-2165

HYBRID METHODOLOGY FOR PREDICTING HYPERTENSION IN PATIENTS WITH DYSLIPIDEMIA AND TYPE 2 DIABETES MELLITUS

Mohamad Nasarudin Adnan[1], Wan Muhamad Amir W Ahmad[1,*], Farah Muna Mohamad Ghazali[2,*], Nor Azlida Aleng[3], Mohamad Shafiq Mohd Ibrahim[4] and Nurfadhlina Abdul Halim[5]

[1]School of Dental Sciences
USM
Malaysia

[2]Faculty of Computer and Mathematical Sciences
UiTM
Malaysia

[3]Faculty of Computer Science and Mathematics
UMT
Malaysia

Received: March 26, 2025; Accepted: May 8, 2025
2020 Mathematics Subject Classification: 62P10, 92B20.
Keywords and phrases: hypertension, dyslipidemia, multiple logistic regression, type 2 diabetes mellitus, MLFNN, surface plot.
*Corresponding authors
Communicated by K. K. Azad

How to cite this article: Mohamad Nasarudin Adnan, Wan Muhamad Amir W Ahmad, Farah Muna Mohamad Ghazali, Nor Azlida Aleng, Mohamad Shafiq Mohd Ibrahim and Nurfadhlina Abdul Halim, Hybrid methodology for predicting hypertension in patients with dyslipidemia and type 2 diabetes mellitus, Advances and Applications in Statistics 92(7) (2025), 1023-1030. https://doi.org/10.17654/0972361725045

Published Online: June 3, 2025

[4]Kulliyyah of Dentistry
IIUM
Malaysia

[5]Risk and Analytics Research Group
Faculty of Science and Technology
USIM
Malaysia

Abstract

This study uses statistical computational methods to model hypertension in patients with dyslipidemia and type 2 diabetes. A retrospective analysis of 39 patients from Hospital Universiti Sains Malaysia identified key factors like blood pressure, glucose, and cholesterol. A hybrid model combining bootstrap, logistic regression, and neural networks achieved 99.99% accuracy with a MAD of 0.0001. Eight factors were significantly associated with hypertension, demonstrating the model's high predictive power and reliability for risk assessment.

1. Introduction

Hypertension is a major risk factor for cardiovascular diseases and is often linked with dyslipidemia and type 2 diabetes mellitus, which exacerbate hypertension through mechanisms like endothelial dysfunction and increased vascular resistance [1]. This synergistic effect elevates the risk of cardiovascular events like stroke and myocardial infarction [2]. The global prevalence of hypertension is rising, contributing to over 9 million deaths annually [7]. However, managing hypertension remains challenging, particularly in advanced stages [3-5]. Current predictive models often focus on single-method approaches, which do not fully account for complex interactions between risk factors [6]. This study proposes a hybrid methodology combining bootstrap sampling, logistic regression, and Multi-Layer Feed-Forward Neural Networks (MLFFNNs) to improve hypertension

risk prediction [3]. This multi-technique approach enhances model reliability and accuracy, achieving a 99.99% accuracy rate and a Mean Absolute Deviation (MAD) of 0.0001, providing a comprehensive tool for clinical decision-making and better patient outcomes [8].

2. Materials and Methods

The study combines bootstrap, logistic regression, MLFFNN, and surface/contour plots to enhance prediction accuracy. Logistic regression models categorical outcomes and identifies risk factors, while MLFFNNs capture complex, non-linear data relationships. This hybrid approach improves hypertension risk prediction, especially in patients with dyslipidemia and type 2 diabetes. The study analyzed data from 39 participants at Hospital Universiti Sains Malaysia, using a retrospective approach and advanced computational methods, including logistic regression and MLFFNN. Methodology evaluation included MSE, predicted values, and accuracy metrics. Contour and surface plots enhanced the analysis. Ethical approval was granted by the USM Research Ethics Committee (USM/JEPeM/16050184). The dependent variable was defined dichotomously: Hyper = 0 for non-hypertensive individuals and Hyper = 1 for hypertensive individuals. The estimated logit is given by

$$\hat{g}(x) = \beta_0 + \beta_1(\text{Marital}) + \beta_2(\text{Sysbp}) + \beta_3(\text{Fbs}) + \beta_4(\text{Tc}) + \beta_5(\text{Hdl}) + \beta_6(\text{Alt}) + \beta_7(\text{Alp}) + \beta_8(\text{Urea}). \quad (1)$$

3. Result and Discussion

Table 1 and Figure 1 show logistic regression results (under the combining method) for hypertension status with a MAD of 0.01121405, indicating high model accuracy. The dataset is split 70:30 for training and testing. Table 1 summarizes the analysis, and Figure 1 displays the model architecture.

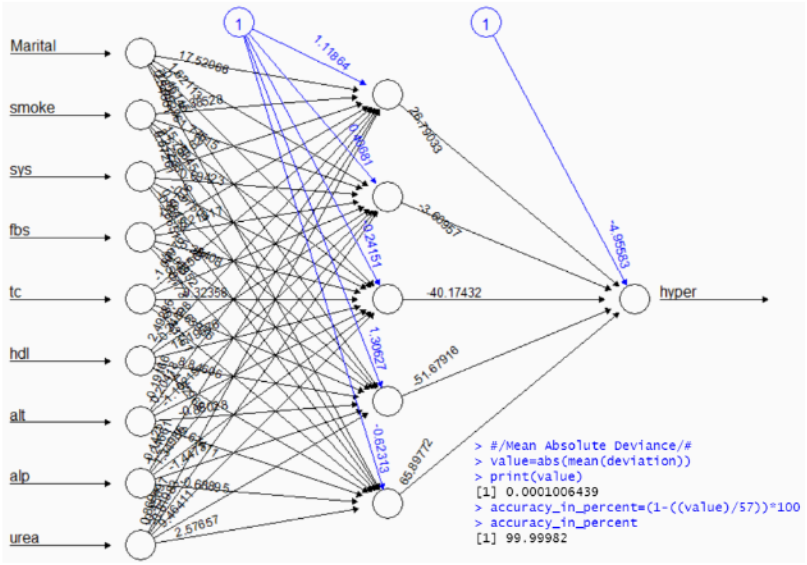

Figure 1. Result of multiple logistic regression by combining the bootstrap method with neural network with separation of training and testing dataset.

Table 1. Coefficient of multiple logistic regression

Variable	Estimate	Std. error	*t*-value	*P*-value
(Intercept)	2.48	1.28×10^{-01}	19.36	2.20×10^{-16} **
Marital	1.95×10^{-1}	2.55×10^{-02}	7.68	3.74×10^{-14} **
Smoke	1.25×10^{-02}	1.50×10^{-02}	-0.83	0.40×10^{-05} **
Sys	-1.57×10^{-02}	8.59×10^{-04}	-18.36	2.20×10^{-16} **
Fbs	1.47×10^{-04}	5.46×10^{-05}	2.69	0.0070631 **
Tc	3.76×10^{-02}	9.82×10^{-03}	3.83	0.0001347 **
Hdl	-3.91×10^{-01}	4.44×10^{-02}	-8.80	2.2×10^{-16} **
Alt	-4.35×10^{-04}	4.09×10^{-04}	-1.06	0.2880626
Alp	-1.57×10^{-03}	3.67×10^{-04}	-4.27	2.13×10^{-05} **
Urea	1.00×10^{-02}	9.49×10^{-04}	10.62	$< 2.20 \times 10^{-16}$ **

Multiple logistic regression was applied.
**Significant at the level of the 0.01

Table 1 presents the results of a logistic regression analysis for the training dataset, with hypertension status as the dependent variable. The

model, validated using the bootstrap method, yields a mean absolute deviance (MAD) of 0.01121405, indicating minimal spread and high consistency among the data points. This supports the model's accuracy and reliability. Nine variables were assessed: marital status, smoking, systolic blood pressure, fasting glucose, cholesterol, hdl, alanine transferase, alkaline phosphatase, and urea reading. Eight of these were significant predictors of hypertension. Significant variables include marital status ($\beta_1 : 1.95;\ p < 0.05$), systolic blood pressure ($\beta_3 : -1.57;\ p < 0.05$), smoking ($\beta_2 : 1.25;\ p < 0.05$), fasting glucose ($\beta_4 : 1.47;\ p < 0.05$), cholesterol ($\beta_5 : 3.76;\ p < 0.05$), hdl ($\beta_6 : -3.91;\ p < 0.05$), alkaline phosphatase ($\beta_8 : -1.57;\ p < 0.05$), and urea ($\beta_9 : 1.00;\ p < 0.05$). The dataset is split into a 70:30 train-to-test ratio, with 70% for model development and 30% for validation. Figure 1 visually illustrates the model architecture.

Figure 2 provides a summary of all the studied variables. Hypertension levels rise with increasing total cholesterol (tc) and systolic blood pressure (sys), as well as with rising urea and systolic blood pressure (sys), with both contour and surface plots highlighting their compounded effects on hypertension risk. Similarly, hypertension increases with higher urea and fasting blood sugar (fbs), higher urea and total cholesterol (tc), and rising urea and high-density lipoprotein (hdl), all of which are emphasized by contour and surface plots showing their combined impact. Additionally, hypertension levels rise with increasing urea and alkaline transferase (alt), as well as with rising urea and alkaline phosphatase (alp), with contour and surface plots reinforcing the joint effects of these factors on hypertension risk. This study introduces a hybrid methodology combining multiple logistic regression (MLR) and a Multi-Layer Feed-Forward Neural Network (MLFFNN) to improve hypertension risk prediction in patients with dyslipidemia and type 2 diabetes. By integrating linear and non-linear components, the model effectively captures complex relationships and uses bootstrap sampling to enhance generalizability. Significant predictors include marital status, systolic blood pressure, cholesterol, hdl, alp, and urea, though smoking and alt were not significant. The hybrid model achieved

exceptional performance with a 99.99% accuracy and low Mean Absolute Deviance (MAD) of 0.0001, making it highly suitable for clinical decision-making. Despite challenges like a small sample size, the model's potential for early hypertension detection is promising. Future work should focus on larger cohorts, additional variables, and more advanced machine-learning techniques.

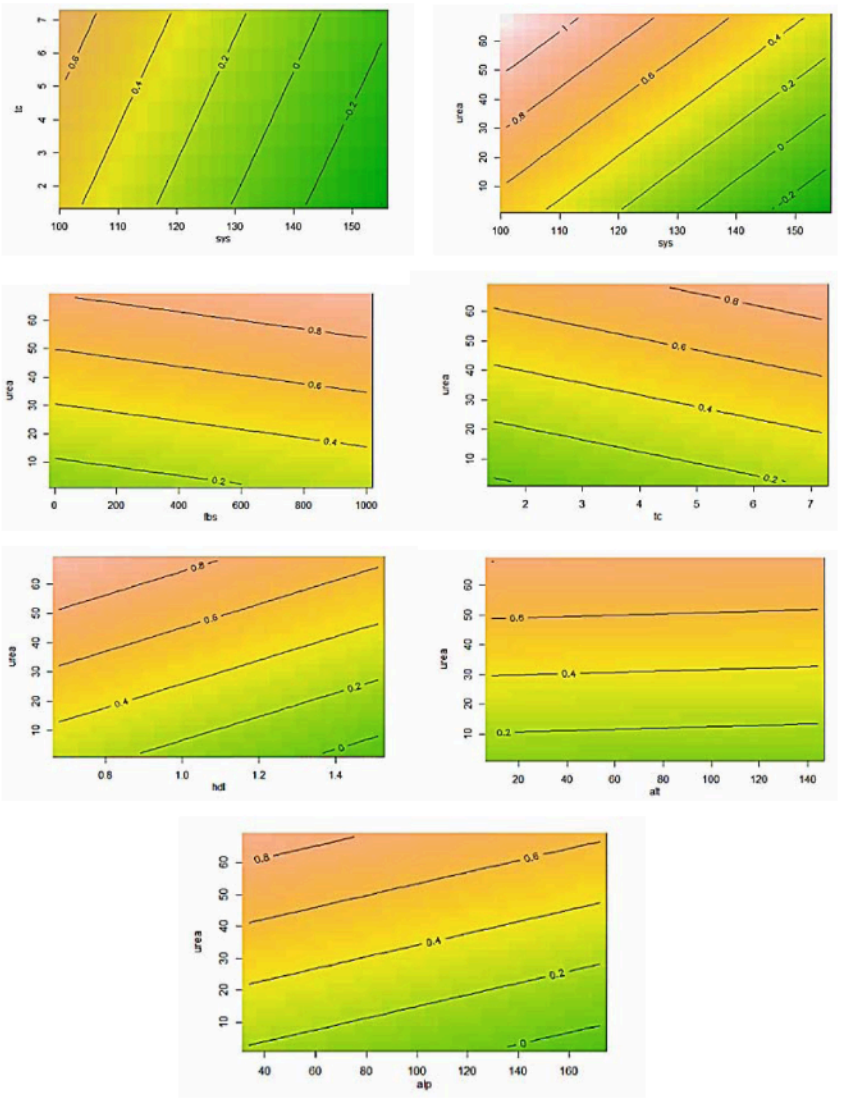

Figure 2. Contour plot for studied variables.

4. Conclusion

This study developed a predictive model for hypertension risk in patients with dyslipidemia and type 2 diabetes, combining logistic regression and MLFFNN. With bootstrap sampling, it achieved 99.99% accuracy and a low MAD of 0.0001, identifying key predictors like marital status, systolic blood pressure, cholesterol, hdl, alp, urea, and smoking. Alt was not significant. This model could enhance early detection and personalized healthcare. Future research should use larger populations and explore additional machine-learning techniques to refine the model for broader clinical applications.

Acknowledgements

The authors express their gratitude to Universiti Sains Malaysia (USM) for their support in funding this study through the Ministry of Higher Education (MOHE) Fundamental Research Grant Scheme (FRGS/1/2022/STG06/USM/02/10).

References

[1] I. Usui, Common metabolic features of hypertension and type 2 diabetes, Hypertension Research 46(5) (2023), 1227-1233. https://doi.org/10.1038/s41440-023-01233-x.

[2] M. Yousaf, V. Razmovski-Naumovski, M. Zubair, D. Chang and X. Zhou, Synergistic effects of natural product combinations in protecting the endothelium against cardiovascular risk factors, Journal of Evidence-Based Integrative Medicine 27 (2022). https://doi.org/10.1177/2515690X221113327.

[3] B. Farran, R. Alwotayan, H. Alkandari, D. Al-Abdulrazzaq, A. Channanath and T. A. Thanaraj, Use of non-invasive parameters and machine-learning algorithms for predicting future risk of type 2 diabetes: a retrospective cohort study of health data from Kuwait, Frontiers in Endocrinology 10 (2019), 624. https://doi.org/10.3389/fendo.2019.00624.

[4] O. T. Kee, H. Harun and N. Mustafa, Cardiovascular complications in a diabetes prediction model using machine learning: a systematic review, Cardiovasc. Diabetol. 22 (2023), 13. https://doi.org/10.1186/s12933-023-01741-7.

[5] H. Kim, D. H. Lim and Y. Kim, Classification and prediction on the effects of nutritional intake on overweight/obesity, dyslipidemia, hypertension and type 2 diabetes mellitus using deep learning model: 4-7th Korea national health and nutrition examination survey, International Journal of Environmental Research and Public Health 18(11) (2021), 5597. https://doi.org/10.3390/ijerph18115597.

[6] C. X. Ma, X. N. Ma, C. H. Guan, Y. D. Li, D. Mauricio and S. B. Fu, Cardiovascular disease in type 2 diabetes mellitus: progress toward personalized management, Cardiovasc. Diabetol. 21(1) (2022), 74. https://doi.org/10.1186/s12933-022-01516-6.

[7] A. Sharma, S. Mittal and R. Aggarwal, Diabetes and cardiovascular disease: inter-relation of risk factors and treatment, Futur. J. Pharm. Sci. 6 (2020), 130. https://doi.org/10.1186/s43094-020-00151-w.

[8] Z. Xie, O. Nikolayeva, J. Luo and D. Li, Building risk prediction models for type 2 diabetes using machine learning techniques, Preventing Chronic Disease 16 (2019), E130. https://doi.org/10.5888/pcd16.190109.