DOI: https://doi.org/10.59294/HIUJS.VOL.7.2024.680

Develop predictive model for direct treatment cost of acute coronary syndrome using a neural network algorithm

Pho Nghia Van[1], Dinh Hoang Yen[2], Huynh Hai Duong[3,*], Le Quan Nghiem[1]
[1]University of Medicine and Pharmacy, Ho Chi Minh City, Vietnam
[2]Blood Transfusion Hematology Hospital, Vietnam
[3]Health Technology Assessment and Application Research Institution, Ho Chi Minh City, Vietnam

ABSTRACT
Background: Acute coronary syndrome (ACS) accounts for half the global economic burden. Current models to predict the ACS treatment cost have low accuracy and high complexity. This study aimed to build a more accurate predictive model using a neural network algorithm. Objectives: 1) Survey the cost of treating ACS at research hospitals. 2) Analyze factors associated with total direct cost of treating ACS at research hospitals. 3) Build and assess a model that predicts the total direct cost of treating ACS at research hospitals. Subjects and methods: A cross-sectional descriptive analysis was conducted based on the electronic medical records of 496 ACS inpatients at Cho Ray and Bach Mai hospitals. Factors associated with the total direct cost were used as inputs to build the neural network model. The grid search tool and k-fold cross-validation were used to select the best set of hyperparameters.Results: Mean total direct cost per ACS patient per course was 75,443,006±52,443,599 VND. Gender, health insurance type, course duration, health status at discharge, and number of comorbidities influenced the cost and were used as model inputs. Regarding the best set of hyperparameters, the distribution was Laplace, the transfer function was rectifier with dropout, the loss function was Absolute, the number of neurons in each hidden layer was 40, the number of hidden layers was 2, the lasso value was 1.0E-5, the ridge value was 1.0E-3, and the rho value was 0.999. The training set root mean squared error (RMSE) (25,091,949 VND) was smaller than those of the validation and test sets (33,025,969 and 29,202,777 VND, respectively); the difference between total predicted and actual cost was not significant, indicating that the optimization and regularization criteria were reached. Conclusions: The predictive model has relatively high accuracy and may be applicable in real-world settings. The model should be continuously enhanced to improve predictions and expanded to other patient groups based on big medical data.

Keywords: *acute coronary syndrome, neural network, predictive model, machine learning*

1. INTRODUCTION

Acute coronary syndrome (ACS) has evolved as a useful operational term that refers to a spectrum of clinical presentations ranging from ST-segment elevation myocardial infarction and angina pectoris to non-ST-segment elevation myocardial infarction or unstable angina [1, 2]. In 2015, there were an estimated 7.4 million deaths due to coronary heart disease[3].

ACS remains a leading cause of death in the world, accounting for half of the global economic burden. This leads to follow-up difficulties and causes poverty among patients after they have paid the treatment costs [4, 5]. Having an estimate of the total direct cost could help patients to prepare the payment before treatment. Furthermore, such estimation could encourage the Thai government to use the state budget to subsidize hospital fees.

There are a number of models that can be used to predict the total direct cost, such as multivariate, lasso and ridge regression models. However, these models often exhibit poor performance when the relationships between variables are nonlinear [6].

Corresponding author: Huynh Hai Duong
Email: duonghuynh.htari@gmail.com

On the other hand, other complicated algorithms such as neural networks, which can have high accuracy and be optimized, have often been prohibitively computationally expensive. Nowadays, following the development of big data, especially in the medical field, highly configurable computers have become available, which makes the use of complicated algorithms (such as neural network algorithms) possible. Therefore, this study aimed to build a more accurate predictive model using a neural network algorithm.

An artificial neural network is a computational model that is inspired by the way biological neural networks in the human brain process information: receiving, processing and transmitting information [7]. The basic unit of computation in a neural network is a neuron, which receives information from other nodes or an external source and computes an output. Multiple nodes form a multiple-layer neuronal model. Artificial neural networks are the most well-known machine learning model and are used for many applications in a wide range of disciplines[8].

Objectives:

1) To survey the cost of treating ACS at research hospitals.

2) To analyse the factors associated with the total direct cost of treating ACS at research hospitals.

3) To build and assess a model that predicts the total direct cost of treating ACS at research hospitals.

2. METHODS

2.1. Data collection

The medical records of inpatients with ACS at Cho Ray and Bach Mai hospitals who satisfied the inclusion and exclusion criteria (Table 1) were examined.

Table 1. Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
- ACS patients treated at Cho Ray or Bach Mai hospitals - Medical records were available from the hospitals and had all the necessary information - Patients had health insurance	- Patients were voluntarily discharged or transferred to another hospital - Patients did not agree to disclose personal information

2.2. Statistical analysis

After data collection, the data were preprocessed and cleaned (using nonparametric tests). The data were imported into R language software. Statistical analysis and construction of the predictive model were performed using R language version 3.4.2 with RStudio Integrated Development Environment (IDE) version 1.1.453. The following packages were used: *xml* and *jsonlite* (for extracting data from the electronic medical records); *dplyr* and *tidyr* (for manipulating/tidying the data); *base* and *nortest* (for statistical analysis); and *h2o* (for building the predictive model).

The following statistical analyses were conducted:

1) Descriptive measures were calculated, comprising frequencies (with percentages) for describing and summarizing categorical variables, and means (with standard deviations and ranges) for describing and summarizing normally distributed quantitative data.

2) As the distribution of the total direct cost was non-normal, non-parametric tests were used, comprising the Mann–Whitney U test, Kruskal–Wallis H test, and Spearman's rank correlation analysis[9].

3) Feature engineering: the predictive model was built according to Figure 1[10].

As shown in Figure 1, first, qualitative variables were encoded as dummy variables while quantitative variables were scaled or normalized so that they had the same range. This ensures stable convergence of weight and biases [11]. Next, the processed dataset was split into three sets: training, validation and test sets. k-fold cross-validation was conducted based on the training and validation sets. In detail, the training set was used to train the model (corresponding to weights and biases in the neural network algorithm), while the validation set was used to conduct an unbiased evaluation of a model fit using the training set while tuning the parameters. The gridsearch tool was used to tune the hyperparameters. The list of hyperparameters is presented in Table 2. The model was built based on the set of hyperparameters and the process of tuning the weights and biases. The test set was used to conduct an unbiased evaluation of the final model fit. k-fold cross-validation was conducted throughout the process of grid searching and model building to evaluate performance and select the appropriate level of flexibility[12].

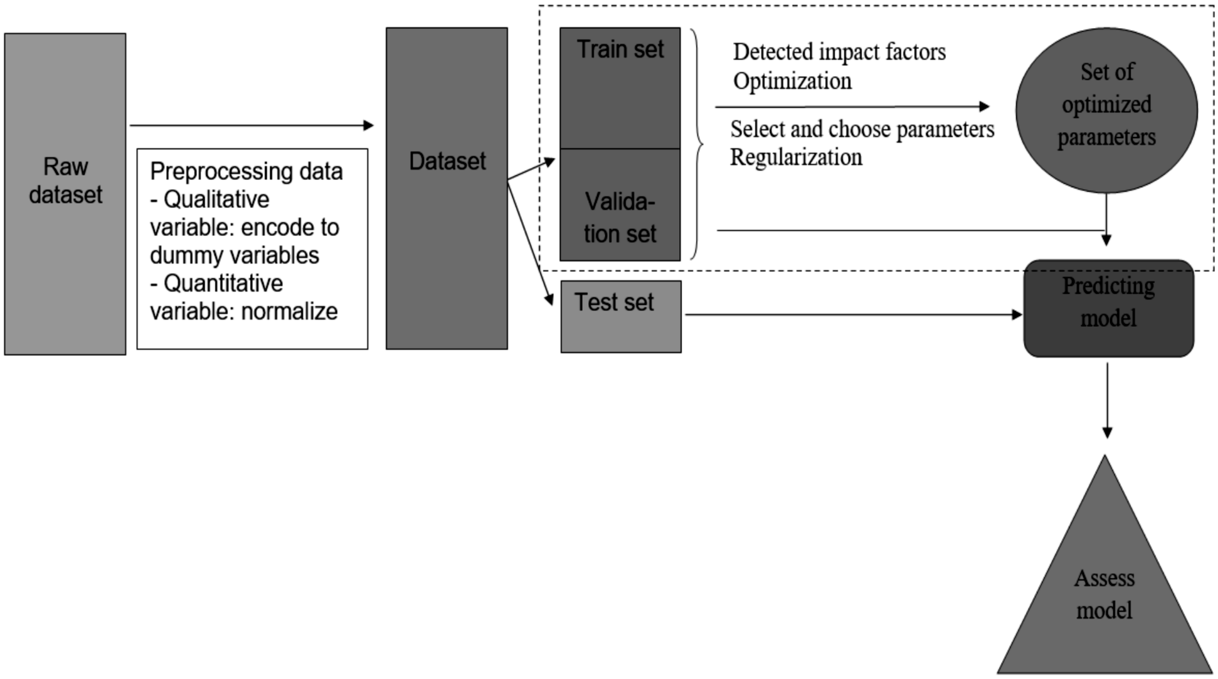

Figure 1. Feature engineering

Table 2. The list of hyperparameters

Parameter	Values
Distribution	AUTO, gaussian, Poisson, gamma, Laplace, quantile, Huber, Tweedie
Transfer function	Rectifier, tanh, maxout, rectifier with dropout, tanh with dropout, maxout with dropout
Loss function	Quadratic, absolute, Huber
Number of hidden layers	1, 2, 3, 4, ..., 10, 20, 30, ..., 100, 200, 300, ...
Number of neurons in hidden layers	1, 2, 3, 4, ..., 10, 20, 30, ..., 100, 200, 300, ...
Lasso	1e-1, 1e-2, 1e-3, 1e-4, 1e-5, 1e-6
Ridge	1e-1, 1e-2, 1e-3, 1e-4, 1e-5, 1e-6
Rho	0.9, 0.99, 0.999
Epsilon	1e-4, 1e-5, 1e-6, 1e-7, 1e-8, 1e-9, 1e-10

3. RESULTS AND DISCUSSIONS

3.1. Sociodemographic and clinical characteristics of study sample

The sociodemographic and clinical characteristics of the 496 participants are shown in Table 3.

The patients had a mean age of 66.3 ± 10.8 years, 352 (71.0%) were male, and 174 (35.1%) had health insurance paid by the state. More than a half of patients (54.8%) had a health insurance payment rate of 100% while the lowest proportion (10.3%) had a health insurance payment rate of 95%. Regarding diagnosis, 61.6% had acute myocardial infarction diagnosis during the index hospitalization while 38.9% had unstable angina. After each treatment course (mean duration: 6 ± 4.5 days), most patients had an improvement in health status at discharge (78.4%). The major comorbidities were diabetes, chronic renal failure and hypertension, and the major risk factors were family history of ACS, lipid disorders, smoking and alcoholism.

Table 3. Sociodemographic and clinical characteristics of study sample

Characteristic		Frequencies (percentage)/ Mean ± SD	Cumulative percentage/ min–max
Hospital	Cho Ray	315 (63.5%)	63.5%
	Bach Mai	181 (36.5%)	100.0%
Gender	Male	352 (71.0%)	71.0%
	Female	144 (29.0%)	100.0%
Age (years)		66.3 ± 10.8	34–92
Health insurance	Paid by employers and employees	32 (6.5%)	6.5%
	Paid by social insurance organization	155 (31.2%)	37.7%
	Paid by the state	174 (35.1%)	72.8%
	Partially paid by the state	10 (2.0%)	74.8%
	Paid by themselves	125 (25.2%)	100.0%
Payment rate of health insurance	80%	173 (34.9%)	34.9%
	95%	51 (10.3%)	45.2%
	100%	272 (54.8%)	100.0%
Diagnosis (ICD code)	Angina pectoris (I20)	193 (38.9%)	38.9%
	Myocardial infarction (I21)	303 (61.6%)	100.0%
Duration of a treatment course (days)		6.0 ± 4.5	1–40
Health status at discharge	Improved	389 (78.4%)	78.4%
	Unchanged	107 (21.6%)	100.0%
Number of comorbidities	0	99 (20.0%)	20.0%
	1	203 (40.9%)	60.9%
	2	130 (26.2%)	87.1%
	≥3	64 (12.9%)	100.0%
Number of risk factors	0	124 (25.0%)	25.0%
	1	261 (52.6%)	77.6%
	≥2	111 (22.4%)	100.0%

Note: ICD – International code disease

3.2 Survey of the cost of treating acute coronary syndrome at research hospitals

Outliers were removed based on a non-parametric method (boxplot) (Figure 2).

Among the 496 medical records, eight medical records that had an extreme total direct cost were removed, and 488 were retained.

The Shapiro–Wilk normality test was used to analyse the distribution of the total direct cost. The result is presented in Table 4.

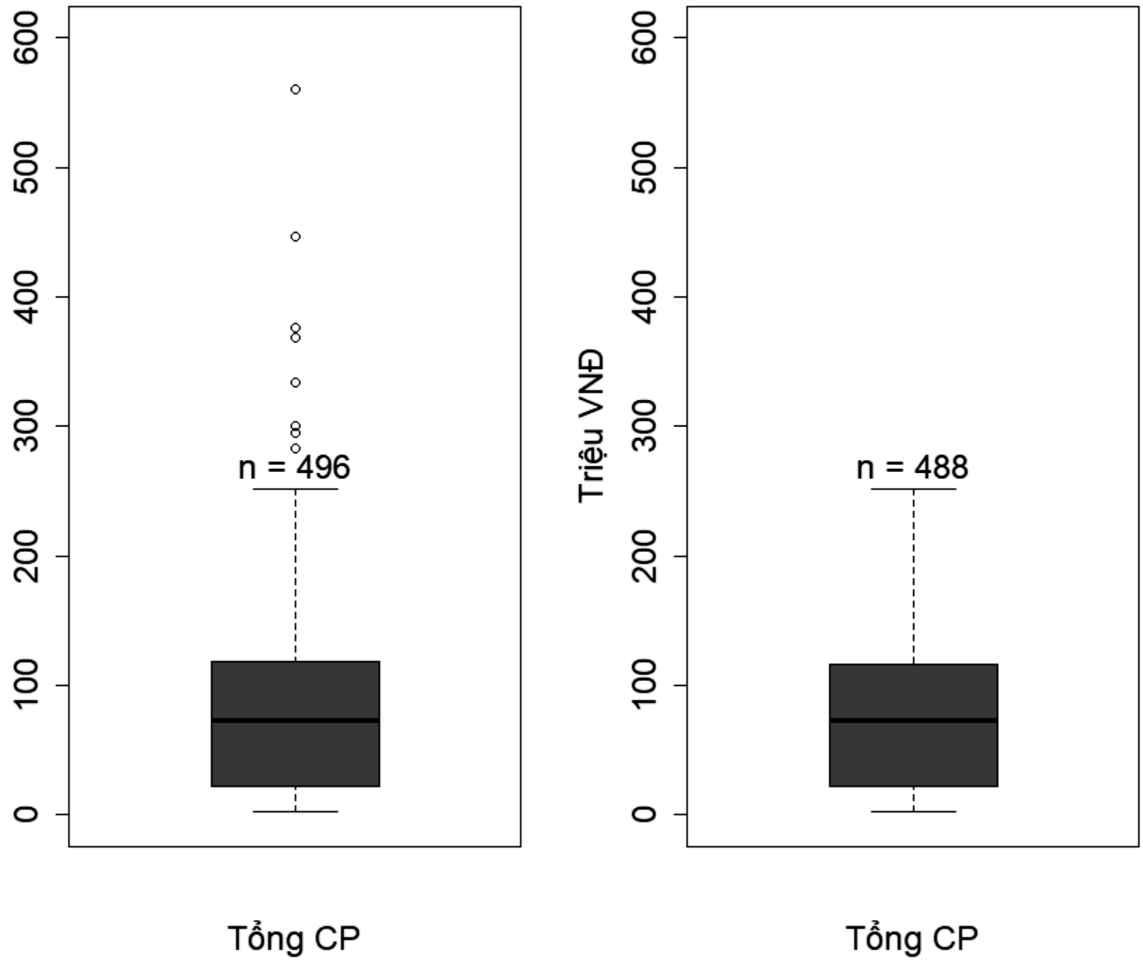

Figure 2. Total direct cost before (left) and after (right) removing outliers

As shown in Table 4, the distribution of the total direct cost was non-normal. Therefore, non-parametric tests comprising the Mann-Whitney U test, Kruskal-Wallis H test and Spearman's rank correlation analysis were used to analyse the factors associated with the total direct cost.

Table 4. Shapiro-Wilk normality test

W	p-value
0.94112	<5.497e-13

Table 5. Direct cost of treating acute coronary syndrome

	Mean	SD	25%	50%	75%
Drugs	6,184,419	17,659,689	596,542	1,363,678	2,280,128
Diagnosis	4,009,194	8,654,999	1,685,075	2,240,000	3,295,425
Medical supplies	50,635,022	49,174,988	568,185	54,593,572	76,959,831
Bed days	2,227,871	2,175,043	1,050,000	1,634,950	2,553,138
Surgical procedures	12,124,763	25,633,434	402,800	5,876,900	7,593,875

	Mean	SD	25%	50%	75%
Other	261,737	276,227	57,000	214,000	366,000
Total direct cost (VND)	75,443,006	52,443,599	21,709,693	72,865,782	116,032,016

As shown in Table 5, the mean total direct cost per ACS patient per course was 75,443,006 ± 52,443,599 VND. The mean length of hospital stay was 6.0 days. Besides the "other" costs incurred during hospitalization (261,737 VND), the total expenditure for one treatment course comprised the costs of drugs, diagnosis, medical supplies, bed days and surgical procedures. Medical supplies had the highest cost, at 50,635,022 ± 49,174,988 VND, or 67.1% of the total direct cost. Next, 16.1% of the total direct cost was due to surgical procedures (12,124,763 ± 25,633,434 VND). The remaining components accounted for <7 million VND per treatment course. The interquartile range of the total direct cost was around 95 million VND, which reveals that the distribution of the total cost was large.

3.3. Factors associated with the total direct cost of treating acute coronary syndrome at research hospitals

As shown in Table 6, there was no significant difference between the total direct cost of treating ACS at the two hospitals (77,578,408 VND – Cho Ray and 71,821,081 VND – Bach Mai; p=0.071). However, there was a significant difference in the total direct cost between male and female patients (78,479,827 VND compared with 68,116,410 VND; p=0.043). Patients partially supported by state health insurance tended to have the highest overall cost (98,964,420 VND) compared with patients with other insurance payments; the Kruskal-Wallis H test indicated that the difference was significant (p=0.049). The duration of the treatment course had a weak correlation with the total direct cost (r=0.168; p=0.000). Additionally, patients with ≥3 comorbidities had a significantly higher cost (90,679,548 VND; p=0.013), which was 1.33 times higher than that for patients without comorbidities (68,077,911 VND). Health status at discharge was also associated with the total direct cost; in detail, patients with a better health status at discharge had a higher cost than patients with an unchanged health status (77,203,588 VND compared to 69,174,018 VND; p=0.044). There was no significant difference in the total direct cost between patients with different numbers of risk factors (p=0.636).

Table 6. Factors associated with the total direct cost of treating acute coronary syndrome

Characteristic		Mean/r	Mean ranks	p-value
Hospital	Cho Ray	77,578,408	253.35	0.071
	Bach Mai	71,821,081	229.50	
Gender	Male	78,479,827	252.81	**0.043**
	Female	68,116,410	224.46	
Age (years)		r=0.055		0.225
Health insurance	Paid by employers and employees	78,919,752	250.90	**0.049**
	Paid by social insurance organization	83,526,369	270.89	
	Paid by the state budget	68,393,138	225.54	
	Partially paid by the state budget	98,964,420	273.20	
	Paid by themselves	72,575,182	234.58	

Characteristic		Mean/r	Mean ranks	p-value
Duration of a treatment course (days)		r=0.168		**0.000**
Health status at discharge	Improved	77,203,588	251.39	**0.044**
	Unchanged	69,174,018	219.97	
Number of comorbidities	0	68,077,911	230	**0.013**
	1	71,512,625	230	
	2	79,462,944	254	
	≥3	90,679,548	292	
Number of risk factors	0	73,292,225	239	0.636
	1	72,586,598	235	
	≥2	84,495,304	272	

3.4. Model that predicts the total direct cost of treating acute coronary syndrome at research hospitals

From the above analysis, gender, health insurance, duration of treatment course, health status at discharge, and number of comorbidities were associated with the total direct cost. These factors were used as input values to build the predictive model using a neural network algorithm.

Using gridsearch tool to select the appropriate parameters.

To select the most appropriate parameters based on the lowest root mean squared error (RMSE), the gridsearch tool in the *h2o* library was used. The data were split in a 0.6:0.2:0.2 ratio into three sets comprising training, validation and test sets. The training and validation sets were used to tune the hyperparameters using the gridsearch tool and sixfold cross-validation. Table 7 shows the top five sets of hyperparameters that were used to build the best models with the lowest RMSE values.

Table 7. Best sets of hyperparameters

Model	Distribution	Transfer function	Loss function	Hidden layers	Lasso	Ridge	Rho
1	Laplace	Rectifier with dropout	Absolute	[40, 40]	1.0E-5	1.0E-3	0.999
2	Gamma	Maxout with dropout	Quadratic	[60, 60, 60]	1.0E-4	1.0E-2	0.99
3	Tweedie	Rectifier with dropout	Huber	[40, 40, 40]	1.0E-3	1.0E-6	0.9
4	Laplace	Rectifier with dropout	Quadratic	[70, 70, 70]	1.0E-4	1.0E-5	0.9
5	Gaussian	Tanh	Huber	[100, 100, 100]	1.0E-6	1.0E-2	0.9

As shown in Table 7, the distribution of model 1 (the selected model) was Laplace, the transfer function was rectifier with dropout, the loss function was Absolute, the number of neurons in each hidden layer was 40, the number of hidden layers was 2, the lasso value was 1.0E-5, the ridge value was 1.0E-3, and the rho value was 0.999. This set of hyperparameters was used to build the model that could predict the total direct cost of treating ACS based on the neural network algorithm. The set of optimizing weights is presented in the appendix.

Assessment of the predictive model

The model was assessed based on optimization and regularization using the training, validation and test RMSE values (Table 8).

Table 8. Performance of the model

Error	Training set	Validation set	Test set
MSE	6.3e+14	1.1e+15	8.5e+14
RMSE	25,091,949	33,025,969	29,202,777

As shown in Table 8, the RMSE value of the training set (25,091,949 VND) was smaller than those of the validation and test sets (33,025,969 and 29,202,777 VND, respectively). This means that the optimization and regularization criteria were reached[12].

Ten cases in the test set were predicted, and the results are shown in Table 9.

Table 9. Comparison between predicted and actual costs

Case	Predicted cost (VND)	Actual cost (VND)		Case	Predicted cost (VND)	Actual cost (VND)
1	58,079,310.2	57,779,548.0		6	74,060,480.2	76,878,612.0
2	101,402,236.1	101,440,323.0		7	78,027,867.1	73,413,764.0
3	75,621,866.7	70,796,046.0		8	71,550,743.1	61,911,700.0
4	114,208,266.2	121,091,493.0		9	74,586,962.5	81,564,667.0
5	115,151,395.7	119,389,150.0		10	73,588,441.9	67,504,817.0

As shown in Table 10, the difference between the total predicted and actual costs was not significant, which means that the predictive model had relatively high accuracy.

4. DISCUSSION

The proportions of male and female patients concurs with the findings of Page et al. [13] and Johnston et al. [14], which showed that the proportions of males diagnosed with ACS were 68% and 72%, respectively. These proportions were higher than that reported by Anh et al. [15]. The mean age in this study was 66.3 ± 10.8 years, ranging from 34 to 92 years. This result is similar to findings in studies by Anh et al. (65.64 ± 13.5) [15], Johnston et al. (54.3 ± 6.7) [14] and Hyun et al. (65.0 ± 13.0) [16].

The estimated total direct cost for one episode was 75,443,006 ± 52,443,599 VND, which is approximately the same as in a study in Thailand by Truong et al. (69,899,677.3 ± 3,607,417.3 VND) [15]. In a study in Thailand by Tran et al. (2012), the total direct cost was lower, at 46.9 million VND [17], which is equivalent to 56.1 million VND (based on the 2017 and 2012 Consumer Price Index, which were 154.9 and 129.5, respectively[18]).

Most studies have built predictive models based on multivariate linear regression, lasso regression or ridge regression. However, these models have poor performance. This is the first study to provide a relatively accurate predictive model. The result of the assessment of optimization and regularization was good. Moreover, the development of big data in the medical field will allow more cases and variables to be used to train and enhance the accuracy of the model.

However, for application in real-world settings, an application should be built based on the core model. The user interface should be friendly and suitable for many common devices such as desktop computers, smartphones, and tablets. The database should be designed to collect data from many online resources, especially electronic medical records.

5. CONCLUSION

The accuracy of the predictive model was moderate. Therefore, for application in real-world settings, the accuracy of the model should be improved by training with more cases and variables. Furthermore, there is a need to build a user interface to turn the model in to a real-time application with a friendly interface that could be displayed on many devices such as desktop computers, smartphones, and tablets.

REFERENCES

[1] M. T. Hagan, H. B. Demuth and M. H. Beale, O. D. Jesús, *Neural network design*, Martin Hagan, 2014.

[2] Z. A. C. Alice, *Feature Engineering for machine learning: Principles and Techniques for data scientists*, O'Reilly, 2018

[3] E. A. Amsterdam, N. K. Wenger,…and D. E. Casey, "2014 AHA/ACC Guideline for the management of patients with non–ST-elevation acute coronary Syndromes: Executive Summary", *Circulation*, 130(25), 2014, pp. 2354-2394.

[4] A. Candel, V. Parmar, E. LeDell and A. Arora. *Deep learning with* H2O, 2017.

[5] M. Y. Chan, X. Du, D. Eccleston, …and M. Ogita, Y. H Jeong. "Acute coronary syndrome in the Asia-Pacific region", *International Journal of Cardiology*, 202, pp. 861-869, 2016.

[6] Chollet F. *Deep learning with python*, Manning Publications Co, 2017.

[7] J. L. Anderson, N. K. Morrow, "Acute coronary syndrome," *The New England Journal of Medicine,* vol. 376, no. 21, pp. 2035–2043, 2017. DOI: 10.1056/NEJMra1606915.

[8] J. D. Gibbons and S. Chakraborti. "Nonparametric statistical inference", *International encyclopedia of statistical science*, Springer, 2011, pp. 977-979.

[9] K. K. Hyun, B. M. Essue, and D. Chew, C. Astley. "The household economic burden for acute coronary syndrome survivors in Australia", *BMC health services research*, 16(1), 2016 pp. 636.

[10] G. James, D. Witten, T. Hastie and R. Tibshirani, *An introduction to statistical learning*, Vol. 112, Springer, 2013.

[11] S. S. Johnston, S. Curkendall,… and R. Maclean. "The direct and indirect cost Burden of acute coronary syndrome", *Journal of occupational and environmental medicine*, 53(1), 2011, pp. 2- 7.

[12] S. Lek, M. Delacoste, P. Baran,...and S. Aulagnier, "Application of neural networks to modelling nonlinear relationships in ecology", *Ecological modelling*, 90(1), 1996 pp. 39-52.

[13] T. Ohira and H. Iso. "Cardiovascular disease epidemiology in Asia", *Circulation Journal*, 77 (7), pp. 1646-1652, 2013.

[14] R. Paudel, K. Lee and J. K. Singh. Prevalence of behavioral risk factors of cardiovascular diseases and associated socio-economic factors among pregnant women in a rural area in Southern Nepal. *BMC Pregnancy Childbirth* 18, 2018, pp. 484.

[15] R. L. Page R, II, V. Ghushchyan, J. Van Den Bos, T. J. Gray, G. L. Hoetzer, D. Bhandary, K. V. Nair. "The cost of inpatient death associated with acute coronary syndrome", *Vascular health and risk management*. 12, 2016 pp. 13.

[16] V. T. Tran, Q. T. Tang, T. N. Nguyen, H. H. D. T. Nguyen. "Total average cost of treatment about a gas acute myocardial infarction at Nguyen Trai Hospital from 01/01/2009 to 31/12/2009", *MedPharmRes*. 16 (1), 2012, pp. 142-150.

[17] P. B. A. Truong, T. T. T. Nguyen, T. N. T Le. "Analyze treatment cost of myocardial infarction in Heart Institute in Ho Chi Minh City", *MedPharmRes,* 21 (1), 2017 pp. 301-305.

[18] S. Tsakalidis, V. Doumpiotis, W. Byrne. "Discriminative linear transforms for feature normalization and speaker adaptation in HMM estimation", *IEEE Transactions on Speech and Audio Processing*. 13 (3), 2005, pp. 367-376.

Xây dựng mô hình dự đoán chi phí trực tiếp trong điều trị hội chứng mạch vành cấp sử dụng thuật toán mạng nơ-ron

Phó Nghĩa Văn, Đinh Hoàng Yến,
Huỳnh Hải Dương và Lê Quan Nghiệm

TÓM TẮT

Đặt vấn đề: Hội chứng mạch vành cấp (ACS) chiếm một nửa gánh nặng kinh tế toàn cầu. Các mô hình hiện tại để dự đoán chi phí điều trị ACS có độ chính xác thấp và độ phức tạp cao. Nghiên cứu này nhằm xây dựng một mô hình dự đoán chính xác hơn bằng cách sử dụng thuật toán mạng nơ-ron. Mục tiêu: 1) Khảo sát chi phí điều trị ACS tại các bệnh viện nghiên cứu. 2) Phân tích các yếu tố liên quan đến tổng chi phí trực tiếp điều trị ACS tại các bệnh viện nghiên cứu. 3) Xây dựng và đánh giá một mô hình dự đoán tổng chi phí trực tiếp điều trị ACS tại các bệnh viện nghiên cứu. Đối tượng và phương pháp: Một phân tích mô tả cắt ngang đã được thực hiện dựa trên bệnh án điện tử của 496 bệnh nhân nội trú mắc ACS tại bệnh viện Chợ Rẫy và Bạch Mai. Các yếu tố liên quan đến tổng chi phí trực tiếp được sử dụng làm đầu vào để xây dựng mô hình mạng nơ-ron. Công cụ tìm kiếm lưới (grid search) và kiểm tra chéo k-lần (k-fold cross-validation) đã được sử dụng để chọn tập hợp siêu tham số tốt nhất. Kết quả: Chi phí trực tiếp trung bình mỗi bệnh nhân ACS cho mỗi lần điều trị là 75,443,006 ± 52,443,599 VND. Giới tính, loại bảo hiểm y tế, thời gian điều trị, tình trạng sức khỏe khi xuất viện và số lượng bệnh lý đi kèm đã ảnh hưởng đến chi phí và được sử dụng làm thông số đầu vào cho mô hình. Về tập hợp siêu tham số tốt nhất, phân phối là Laplace, hàm truyền là rectifier với dropout, hàm mất mát là Absolute, số nơ-ron trong mỗi lớp ẩn là 40, số lớp ẩn là 2, giá trị lasso là 1.0E-5, giá trị ridge là 1.0E-3, và giá trị rho là 0.999. Sai số trung bình căn bậc hai (RMSE) của tập huấn luyện (25,091,949 VND) nhỏ hơn so với các tập xác thực và kiểm tra (lần lượt là 33,025,969 và 29,202,777 VND); Sự chênh lệch giữa chi phí dự đoán và chi phí thực tế không có ý nghĩa thống kê, cho thấy rằng các tiêu chí tối ưu hóa và điều chỉnh đã đạt được. Kết luận: Mô hình dự đoán có độ chính xác tương đối cao và có thể áp dụng trong thực tiễn. Mô hình cần được cải tiến liên tục để nâng cao độ chính xác dự đoán và mở rộng cho các nhóm bệnh nhân khác dựa trên dữ liệu y tế lớn.

Từ khóa: *hội chứng mạch vành cấp, mạng nơ-ron, mô hình dự đoán, máy học*

Received: 30/09/2024
Revised: 14/10/2024
Accepted for publication: 15/10/2024