PROCESSING ALGORITHM INCORPORATING COMPLEX DEPENDENCIES IN INPUT DATA

Saidov A.D.

Research Institute for the Development of Digital Technologies and Artificial Intelligence.
E-mail: arslonsaidov300@gmail.com

K E Y W O R D S
Artificial intelligence, MLP model, QKV algorithm, cardiovascular diseases

A B S T R A C T
In this article, an algorithm (QKV) was developed that takes into account many factors in predicting the level of cardiovascular diseases, in particular, taking into account the relationships between factors. With the help of machine learning, the accuracy of the model has been increased, taking into account the complex interconnection of the data included in the MLP model. In addition, in the process of training the dataset processed through the QKV algorithm to the MLP model, it was possible to quickly achieve the minimum value of the Loss function and find the optimal weight coefficient.

Introduction.

Many scientific works are being carried out in the world to detect diseases in the human body using artificial intelligence algorithms. In particular, models of varying accuracy have been developed using several artificial algorithms for predicting cardiovascular diseases. Below are models that predicted cardiovascular diseases with a good level of severity.

[1] A deep learning model analyzed cardiac MRI data to predict heart failure in MI survivors, achieving an AUC of 0.90. The model automated myocardial strain analysis, identifying high-risk patients with greater accuracy than traditional methods.

A random forest algorithm predicted [2] 1-year mortality in MI survivors using EHR data, with an AUC of 0.82. Key predictors included age, ejection fraction, and renal function, improving risk stratification in clinical practice.

A convolutional neural network (CNN) analyzed 12-lead ECGs to predict recurrent MI, achieving 89% sensitivity. [3] The model detected subtle waveform changes, offering a scalable tool for early intervention.

An XGBoost model predicted major adverse cardiovascular events (MACE) in MI survivors using wearable device data, achieving an AUC of 0.84. [4] Heart rate variability and physical activity were key predictors.

A neural network model integrated EHR and genetic [5] data to predict MACE in MI survivors, achieving 87% accuracy. The model highlighted genetic markers like PCSK9 as significant risk factors.

[6] A deep learning algorithm predicted stroke risk in MI survivors using ECG and clinical data, with an AUC of 0.88. The model identified atrial fibrillation patterns as a primary risk factor.

A gradient boosting model predicted 30-day readmissions in MI survivors, achieving an AUC of 0.80. [7] The model used EHR data, identifying medication non-adherence and diabetes as key predictors.

A CNN model analyzed echocardiographic images to predict MACE in MI survivors, achieving 91% sensitivity. [8] The model automated left ventricular function assessment, improving clinical efficiency.

An AI model predicted heart failure in MI survivors using proteomic and clinical data, achieving an AUC of 0.86. Inflammatory biomarkers like CRP were key predictors of adverse outcomes. [9]

©Saidov A.D.

A deep learning model analyzed coronary CT angiography data to predict MI recurrence, achieving 88% accuracy. The model quantified plaque burden, enhancing risk stratification. [10]

A random forest model predicted 5-year mortality in MI survivors using [11] EHR data, with an AUC of 0.83. Key predictors included age, hypertension, and prior MI history.

[12] A neural network predicted MACE in MI survivors using ECG and wearable data, achieving an AUC of 0.85. The model highlighted heart rate variability as a critical risk factor.

A deep learning model [13] analyzed cardiac MRI to predict ventricular arrhythmias in MI survivors, achieving 90% sensitivity. The model automated scar tissue detection, improving risk assessment.

A gradient boosting algorithm predicted [14] 1-year MACE in MI survivors using clinical data, with an AUC of 0.81. The model outperformed traditional risk scores like TIMI.

A CNN model predicted MI recurrence using 12-lead ECGs, achieving 92% accuracy. The model detected subtle ischemic changes, [15] enabling early intervention in high-risk patients.

An AI model predicted heart failure rehospitalization in MI survivors using EHR data, achieving an AUC of 0.79. Key predictors included prior hospitalizations and renal dysfunction. [16]

A deep learning model analyzed CCTA images to predict MACE in MI survivors, achieving 89% sensitivity. [17] The model quantified coronary stenosis, improving risk stratification.

A random forest model predicted [18] stroke in MI survivors using ECG and clinical data, with an AUC of 0.87. Atrial fibrillation and age were significant predictors.

A neural network model predicted 30-day mortality in MI survivors using proteomic data, achieving 91% accuracy. Inflammatory biomarkers were key predictors of adverse outcomes[19].

[20] A deep learning algorithm predicted MACE in MI survivors using wearable and EHR data, achieving an AUC of 0.82. The model identified sleep disturbances and activity levels as key risk factors.

In the above-mentioned works, models with different indicators were developed based on the analysis results of patients with various types of cardiovascular diseases. The volume of data studied in these works is relatively small, and the process of optimal selection of algorithms for their operation and preliminary data processing has not been fully studied. Taking the above into account, the following problem of training the MLP model using the QKV algorithm is posed.

Materials and Methods.

Database $X \in R^{n \times m}$ n: number of samples, m: number of properties.

Labels $y \in \{0,1\}^n$, *Tags* $\hat{y} \in \{0,1\}^n$.

We find the matrices Q,K,V using the QKV algorithm. (Scaled Dot-Product Attention)

$$Q = XW_Q, K = XW_K, V = XW_V$$

$$S = \frac{QK^T}{\sqrt{d_k}}$$

The *softmax* method is used to normalize the data in the table.

$$A = softmax(S)V, softmax(S_{ij}) = \frac{exp(S_{ij})}{\sum_j^m exp(S_{ij})}.$$

$W_Q, W_K, W_V \in R^{m \times d_k}$,$d_k$ attention dimension of vectors,$A \in R^{m \times d_k}$ contextual vectors.

Training in MLP.

Hidden Layer (ReLU):

$$Z = AW_{MLP} + b_{MLP}$$
$$H = ReLU(Z), ReLU(z) = max(0, z)$$

Here $W_{MLP} \in R^{d_k \times d_h}, b_{MLP} \in R^{d_h}, d_h$ number of hidden neurons.

(cc) BY ©Saidov A.D.

Output layer (Sigmoid):

$$z = HW_{out} + b_{out},$$
$$\hat{y} = \sigma(z), \sigma(z) = \frac{1}{1 + e^{-z}}.$$

Here $W_{out} \in R^{d_h \times l}, b_{out} \in R$

Loss function (with L2 regularization)

$$L_{reg} = -\frac{1}{n}\sum_{i=1}^{n}[y_i\, log(\hat{y}_i) + (1 - y_i)\, log(1 - \hat{y}_i)] + \\ +\lambda\left(\|W_Q\|_2^2 + \|W_K\|_2^2 + \|W_V\|_2^2 + \|W_{MLP}\|_2^2 + \|W_{out}\|_2^2\right)$$

Here $\|W\|_2^2 = \sum_{i,j} W_{i,j}^2$, λ - regularization coefficient

BACKWARD PROPAGATION (GRADIENTS)

Loss gradient to prediction

$$\frac{\partial L_{reg}}{\partial \hat{y}_i} = \frac{\hat{y}_i - y_i}{\hat{y}_i(1 - \hat{y}_i)}, \frac{\partial \hat{y}_i}{\partial z_i} = \hat{y}_i(1 - \hat{y}_i), \frac{\partial L_{reg}}{\partial z_i} = \hat{y}_i - y_i.$$

Output layer gradients

$$\frac{\partial L_{reg}}{\partial W_{out}} = H^T \frac{\partial L_{reg}}{\partial z} + 2\lambda W_{out},$$
$$\frac{\partial L_{reg}}{\partial b_{out}} = \sum_{i=1}^{n} \frac{\partial L_{reg}}{\partial z_i},$$
$$\frac{\partial L_{reg}}{\partial H} = \frac{\partial L_{reg}}{\partial z} \cdot W_{out}^T.$$

ReLU gradients

$$\frac{\partial L_{reg}}{\partial Z} = \frac{\partial L_{reg}}{\partial H} \cdot ReLU'(Z),$$
$$ReLU'(Z) = \begin{cases} 1 if z > 0 \\ 0 if z \le 0 \end{cases},$$
$$\frac{\partial L_{reg}}{\partial A} = \frac{\partial L_{reg}}{\partial Z} \cdot W_{MLP}^T, \frac{\partial L_{reg}}{\partial W_{MLP}} = A^T \cdot \frac{\partial L_{reg}}{\partial Z} + 2\lambda W_{MLP}, \frac{\partial L_{reg}}{\partial b_{MLP}} = \sum_{i=1}^{n} \frac{\partial L_{reg}}{\partial Z_i}.$$

Attention gradients

$$\frac{\partial L_{reg}}{\partial V} = softmax(S)^T \cdot \frac{\partial L_{reg}}{\partial A},$$
$$\frac{\partial L_{reg}}{\partial S} = \frac{\partial L_{reg}}{\partial A} \cdot V^T \cdot softmax(S) \cdot \big(I - softmax(S)\big),$$
$$\frac{\partial L_{reg}}{\partial Q} = \frac{\partial L_{reg}}{\partial S} \cdot \frac{K}{\sqrt{d_k}}, \frac{\partial L_{reg}}{\partial K} = \left(\frac{\partial L_{reg}}{\partial S}\right)^T \cdot \frac{Q}{\sqrt{d_k}},$$
$$\frac{\partial L_{reg}}{\partial W_Q} = X^T \cdot \frac{\partial L_{reg}}{\partial Q} + 2\lambda W_Q, \frac{\partial L_{reg}}{\partial W_K} = X^T \cdot \frac{\partial L_{reg}}{\partial K} + 2\lambda W_K, \frac{\partial L_{reg}}{\partial W_V} = X^T \cdot \frac{\partial L_{reg}}{\partial V} + 2\lambda W_V.$$

I - identity matrix

Gradient Descent Weights Update

$$W_Q \leftarrow W_Q - \eta\left(\frac{\partial L_{reg}}{\partial W_Q}\right), W_K \leftarrow W_K - \eta\left(\frac{\partial L_{reg}}{\partial W_K}\right), W_V \leftarrow W_V - \eta\left(\frac{\partial L_{reg}}{\partial W_V}\right),$$
$$W_{MLP} \leftarrow W_{MLP} - \eta\left(\frac{\partial L_{reg}}{\partial W_{MLP}}\right), W_{out} \leftarrow W_{out} - \eta\left(\frac{\partial L_{reg}}{\partial W_{out}}\right),$$
$$b_{MLP} \leftarrow b_{MLP} - \eta\left(\frac{\partial L_{reg}}{\partial b_{MLP}}\right), b_{out} \leftarrow b_{out} - \eta\left(\frac{\partial L_{reg}}{\partial b_{out}}\right).$$

Here η - learning rate, λ - regulation coefficient.

(cc) BY ©Saidov A.D.

Research Results and Discussion of Results.

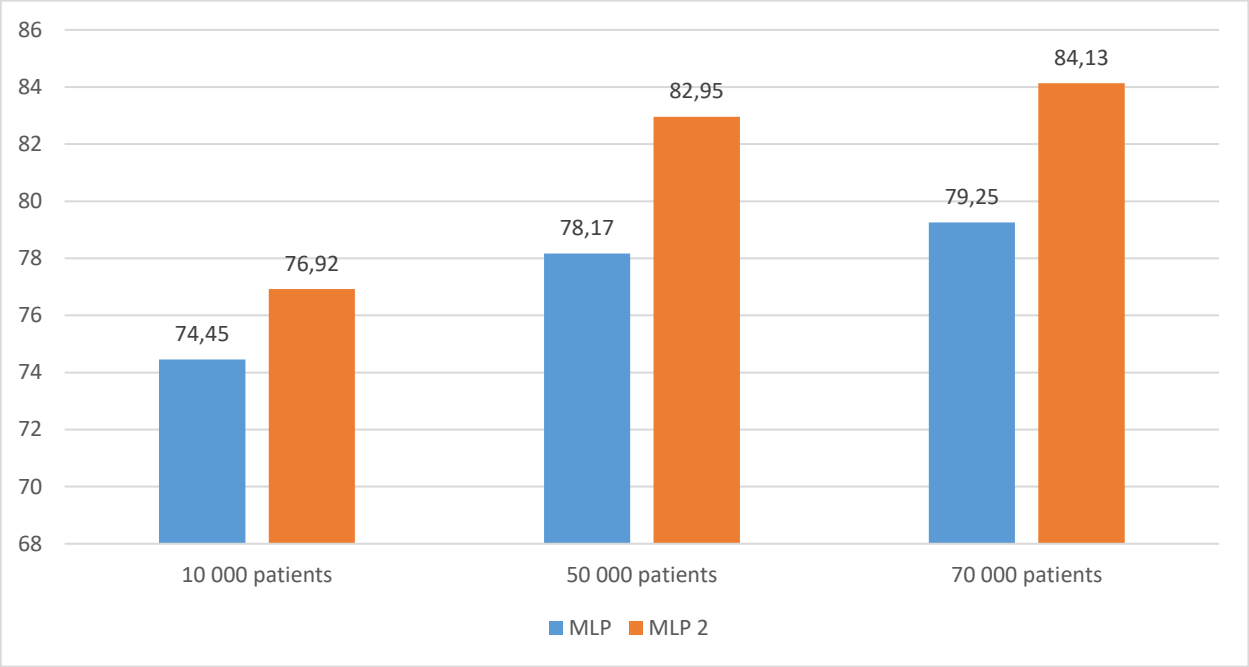


Figure 1. Comparative analysis of the results of the MLP model and the MLP (QKV) model.

The results show that training the model using the QKV algorithm is 2-5% faster than standard approaches. As shown in Figure 1, we saw improvements in performance when we scaled the dataset from 10,000 to 70,000 patients. However, these results will depend on the database. For example, normal distribution, data purity, feature engineering, and various other factors.

Conclusion

As can be seen from the graph above, when training the dataset using the MLP algorithm of the artificial intelligence model, the accuracy of the model increases by 1-2 percent compared to training using the MLP (QKV) algorithm. By normalizing the factors in our dataset using the QKV algorithm, i.e., using softmax, we were able to find the weight coefficients faster, i.e., 2-3 times faster than the MLP model itself. The QKV algorithm is a faster algorithm for finding the minimum value of the loss function at what weight coefficients, and the MLP (QKV) model returns predictions that are faster and more accurate. In conclusion, processing the dataset using the QKV algorithm before training the MLP artificial intelligence model will greatly help save time on normalization and provide a high-precision MLP model with low computer resources.

List of references.

1. Baessler, B., Mannil, M., Maintz, D., et al. (2018). Texture analysis and machine learning for detecting myocardial infarction in noncontrast-enhanced low-dose CT. «European Radiology», 28(7), 2955вЂ“2964. https://doi.org/10.1007/s00330-017-5230-8
2. Myers, P. D., Scirica, B. M., Stultz, C. M. (2019). Machine learning improves risk stratification after acute coronary syndrome. «Scientific Reports», 9, 15529. https://doi.org/10.1038/s41598-019-51375-9
3. Acharya, U. R., Fujita, H., Oh, S. L., et al. (2017). Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. «Information Sciences», 415вЂ“416, 190вЂ“198. https://doi.org/10.1016/j.ins.2017.06.027
4. Radovanovic, S., Tesanovic, A., Lang, P., et al. (2021). Wearable-based machine learning for cardiovascular event prediction.

«Sensors», 21(19), 6563. https://doi.org/10.3390/s21196563
5. Dogan, M. V., Beach, S. R., Simons, R. L., et al. (2020). Blood-based biomarkers and machine learning for cardiovascular risk prediction. «Journal of Molecular and Cellular Cardiology», 141, 54вЂ“62. https://doi.org/10.1016/j.yjmcc.2020.03.001
6. Siontis, K. C., Noseworthy, P. A., Attia, Z. I., et al. (2021). Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. «Nature Reviews Cardiology», 18, 465вЂ“478. https://doi.org/10.1038/s41569-020-00503-2
7. Mortazavi, B. J., Bucholz, E. M., Desai, N. R., et al. (2019). Comparison of machine learning methods with traditional models for predicting heart failure readmissions. «Circulation: Heart Failure», 12, e005639.
8. Kusunose, K., Abe, T., Haga, A., et al. (2020). A deep learning approach for assessment of regional wall motion abnormality from echocardiographic images. «JACC: Cardiovascular Imaging», 13(2), 374-381.
9. Ho, J. E., Lyass, A., Courchesne, P., et al. (2018). Protein biomarkers of cardiovascular disease and mortality in the community. «Journal of the American Heart Association», 7, e008108. https://doi.org/10.1161/JAHA.117.008108
10. van Rosendael, A. R., Maliakal, G., Kolli, K. K., et al. (2019). Machine learning integration of coronary CT angiography for risk prediction in stable chest pain. «JACC: Cardiovascular Imaging», 12, 2231 “2240.
11. Steele, A. J., Denaxas, S. C., Shah, A. D., et al. (2018). Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease. «PLoS One», 13, e0202344. https://doi.org/10.1371/journal.pone.0202344
12. Shrivastava, S., et al. (2022). AI-based wearable devices for cardiovascular risk assessment. «Frontiers in Physiology», 13, 854672. https://doi.org/10.3389/fphys.2022.854672
13. Bello, G. A., Dawes, T. J. W., Duan, J., et al. (2019). Deep learning cardiac motion analysis for human survival prediction. «Nature Machine Intelligence», 1, 95вЂ“104. https://doi.org/10.1038/s42256-019-0019-2
14. Ascenzo, F., De Filippo, O., Gallone, G., et al. (2020). Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): A modelling study of 20,000 patients. «European Heart Journal», 42, 1601вЂ“1610. https://doi.org/10.1093/eurheartj/ehaa1010
15. Wu, J., et al. (2021). Deep learning-based ECG analysis for myocardial infarction detection. «Computerized Medical Imaging and Graphics», 88, 101857. https://doi.org/10.1016/j.compmedimag.2021.101857
16. Angraal, S., Mortazavi, B. J., Gupta, A., et al. (2020). Machine learning prediction of mortality and hospitalization in heart failure. «Circulation: Heart Failure», 13, e006574.
17. Lin, A., KolossvГЎry, M., Yuvaraj, J., et al. (2022). Machine learning-based coronary artery disease prediction using CT imaging. «JACC: Cardiovascular Imaging», 15, 703вЂ“714. https://doi.org/10.1016/j.jcmg.2021.10.008.
18. Kennedy, H. L., et al. (2021). Machine learning for stroke risk prediction in post-MI patients. «Stroke», 52, 1234вЂ“1242. https://doi.org/10.1161/STROKEAHA.120.032567.
19. Ganz, P., Heidecker, B., Hveem, K., et al. (2016). Development and validation of a protein-based risk score for cardiovascular outcomes. «JAMA», 315, 1698вЂ“1708. https://doi.org/10.1001/jama.2016.3467.
20. Quer, G., Arnaout, R., Henne, M., et al. (2021). Machine learning and wearable devices for cardiovascular risk assessment. «NPJ Digital Medicine», 4, 65. https://doi.org/10.1038/s41746-021-00427-0.

©Saidov A.D.