International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Heart Disease Prediction using Graph Neural Network

Rakhi Wajgi[1], Tushar Champaneria[2], Dipak Wajgi[3], Yogesh Suryawanshi[4], Dinesh Bhoyar[5], Ajinkya Nilawar[6]

Submitted: 20/11/2023 **Revised**: 29/12/2023 **Accepted**: 09/01/2024

Abstract: Heart is an important organ playing vital role in the life of living organisms**.** Heart and circulatory disease encompasses a range of conditions affecting the heart and blood vessels, including coronary artery disease, arrhythmias, and heart failure mechanism. Early detection of malfunctioning before failure of heart is necessary. This paper deals with the model built using Graph Neural Network (GNN) to predict heart disease so that mortality rate caused due to sudden heart failure can be reduced. In order to improve the accuracy of GNN-based model, different optimizers are used. They are help to optimize or improve the model's performance by iteratively updating its parameters to reach the optimal values that minimize the difference between predicted and actual outputs. The proposed model is applied on a real dataset from kaggle containing 14 features. Out of all optimizers, RMSprop outperforms otherswith accuracy of 91% and MSE of 48%.

Keywords*: Cardiovascular disease, Graph Neural Network, Optimizer*

1. Introduction

In computer discipline, graph is an important data structure consisting ofnodes and edges connecting those nodes. The Graph element may or may not affect the associated graph element. The feature of graph element can depend upon features of another graph element.GNN stands for Graph Neural Network, which is a type of neural network designed to handle data structure like graphs. Unlike traditional neural networks that process data organized in a grid or sequence (such as images or text), GNNs are specialized for dealing with graph-structured data, where entities (nodes) are connected by relationships (edges).A Graph neural network operates on Graph Structure which has nodes and edges where nodes represents the entities and edges represent the relationships between the entities and provide easy way to do some node-level, edge-level and graph-level prediction tasks.GNNs aim to learn and understand the underlying patterns, interactions, and representations of nodes and edges in a graph. They perform operations that aggregate information from a node's neighbourhood (connected nodes) to update node representations, allowing for powerful learning and inference on graph-structured data. Among various life threatening disease, heart disease has gained a great attention in research. The diagnosis of heart disease is daunting task, automatic prediction about the heart condition to be carried out so that it can be effective to do the further treatment. Numerous habits elevate the likelihood of heart disease, including smoking, a family background with heart disease, elevated cholesterol levels, obesity, hypertension, and insufficient physical activity.Cardiovascular disease are the major cause of mortality in developing countries [7,8,9,10,11] due to random and busy life style. Logistic Regression (LR), Back Propagation Neural Network (BPNN), and Multilayer Perceptron (MLP) have been effectively employed as decision-making tools for predicting heart disease using individual-specific information[12].

Past literature suggest that hybrid models performs better in predicting heart related diseases such as Random Forest, Multilayer Perceptron, SVM, Bayes Net[13]. The organization of paper is as follows:

- Section 2 briefs about literature survey and important terminologies and notations related to GNN.
- Section 3 deals with dataset used and methodology of implementation and section 4 discusses about result followed by conclusion and future scope.

2. Literature Survey

The graph neural network is proposed in [1] for drug-disease association prediction (GNDD) framework to tackle the existing challenge of drug-disease prediction which rely on assembling multiple drug related

1 Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra (India) rakhiwajgi17@gmail.com
2, Department of Computer Engineering, Governemnt Engineering College, Modasa, Gujarat,(India)tushar.champaneria@gecmodasa.ac.in
3 Department of Computer Engineering, St. VincntPallotti College of Engineering and Technology, Maharashtra (India) dipak.wajgi@gmail.com
4 Department of Electronics Engineering, YeshwantraoChavan College of Engineering , Nagpur, INDIA yogesh_surya8@rediffmail.com
5 Department of Electronics and Telecommunication Engineering, Yeshwantrao Chavan College of Engineering, Nagpur, INDIA dinesh.bhoyar23@gmail.com
6 Department of Electronics and Communication Engineering, Shri Ramdeobaba College of Engineering and Management, Maharashtra (India) nilawarap1@rknec.edu

biological information. Assuming that user with similar characteristics would interact with similar items is widely used in recommender system to eliminate the dependency on multi-data. Author in [2] aims to extend the data representation and classification capabilities of convolutional neural network using Graph Convolutional Neural Network (GCNN). The classifier uses structural connectivity inputs in the form of graph Laplacian to generate cognitive status category label as its output. For the purpose of predicting next-period prescriptions, the author in [3] suggests a hybrid RNN and GNN method termed RCNN. RNN is used to describe patient status sequences, while GNN is utilized for presenting periodic medical event graphs. RNN is popular for patient longitudinal medical data representation but it cannot represent complex interaction of different medical information so this temporal graphs can be represented by Graph neural network. [4] Spatio-temporal graph neural networks, or STGNNs, are a type of graph with applications across multiple fields that represent node connections as a function of time and place. The key idea of STGNN considers spatial dependency and temporal dependency at the same time.STGNNs is used to evaluate on the US country level COVID-19 dataset. [5] Propose Knowledge Graph Neural Network (KGNN) to resolve drug disease interaction prediction. This framework can effectively capture drug and its potential neighbourhoods by mining their associated relation in knowledge graph. Author in paper [6] described graph-based deep learning model deep2Conv to systematically conclude new drug disease relationships for SARS-COV-2 drug repositioning. The fundamental idea behind deep2Conv involves combining varied information from networks related to SARS-COV-2, including the drug-drug network, drug-disease network, and drug-target network. This amalgamation aims to deduce potential drugs for SARS-COV-2 through a collective graph convolutional network.

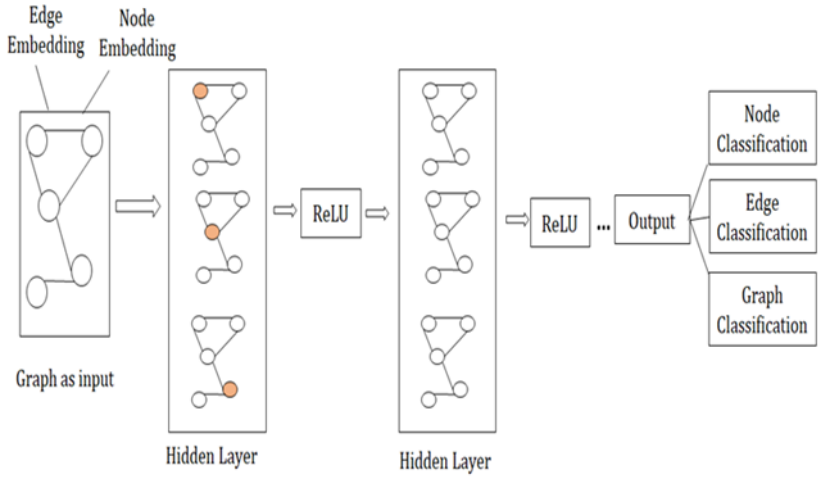

Fig. 1. General working of GNN

2.1 Terminologies and Notations

Graph: Graph is defined as an ordered set and it can be denoted by G = (V, E) where V represents the vertices and E represents the edges which is used to connect the vertices.GNNs are being used by an increasing number of businesses to enhance recommendation systems, fraud detection, and medication discovery. A detailed survey of GNN can be found in [14] Finding patterns in the relationships between data pieces is essential to these and many other applications. GNN applications in computer graphics, cybersecurity, genomics, recommendation systems, and materials science etc. are being investigated by researchers. In a recent work, it was shown that GNNs improved arrival time forecasts by using transportation maps as graphs [15]. GNN uses message passing to encapsulateinformation in a node. This information can be about neighboring nodes or GNNs are The general pipeline for GNN is shown in Fig. 1.The input to the neural network is in terms of graph which can have node or edge embedding and the output is in the form of node classification or edge classification or graph classification.

3. Dataset and Methodology

The Dataset used in this research is the public health dataset which is dated from 1988 and it consist of four dataset which is Cleveland,Hungary, Switzerland, and Long Beach V. This dataset has 76 attributes including ground truth. Figure 2 shows the overall distributions of all 14 features.

Table 1. Notations used in GNN

Notation	Description
G	A graph
V	The vertex set of G
E	The edge set of G
A	The graph adjacency matrix.
A^T	The transpose of the matrix A.
v	The node v ϵ V
N (v)	The adjacent nodes of node v
n	The number of nodes, n = |V|.
m	The number of edges, m = |E|.
d	The dimension of a node feature vector.
b	The dimension of a hidden node feature vector.
c	The dimension of an edge feature vector.
h_v	Hidden state of node v
o_v	Output of node v
X_v	Features of node v
$X_{co[v]}$	Features of its edges
h_{nv}	The states
X_{nv}	Features of the nodes in the neighbourhood of v
k	The layer index

but all the published research used only 14 attributes which are listed in table 2. The predicting attribute "target" field refers to heart disease patient and non heart disease patient. If the value is 0 then no heart disease and 1 denotes heart disease. edges so the machine learning algorithms can be benefited from it. significantly more effective because they carry out graph classification directly using the retrieved graph representations [16].

Table 2. Features used for heart disease prediction

Sr. No.	Attribute	Description
1	Age	Age of patient in year
2	Sex	Gender of patient
		0 = female
		1 = male
3	cp	Chest pain
		0 = typical angina (decreases blood supply to heart)
		1 = not related to heart
		2 =non- heart related
		3 = no sign of disease

4	trestbp	Resting blood pressure. 120-80 = normal range
5	chol	Serum cholesterol shows the amount of triglycerides present. It should be <170mg/dL
6	fbs	Fasting blood sugar. 1 = 120mg/dL <100 = normal 100-125 = prediabetes.
7	restecg	Resting electrocardiographic results. 0 = nothing 1= Can range from mild to severe 2= possible left ventricular hypertrophy
8	thalach	Maximum heart rate achieved is 220 minus your age
9	exang	Exercise induced angina. Angina is caused due to less blood flow.
10	oldpeak	ST depression induced by exercise relative to rest.
11	slope	Slope of the peak exercise ST segment. 0 = up-sloping. It is uncommon 1 =indicates healthy heart 2 = indicates unhealthy heart
12	ca	Colored vessel means the blood passing through. If the blood has movement then there is no clots. Number of major vessels which varies from 0-3 colored by fluroscopy.
13	thal	Thalassemia stress if 1,3= normal, if value 6: fixed defect, and 7 = reversible defect
14	target	Target denotes 0 = no heart disease and 1 = Heart disease presence

It observed that there are more number of male in the age group of 52-68 in the dataset. Seaborn python library is used for visualization of data.

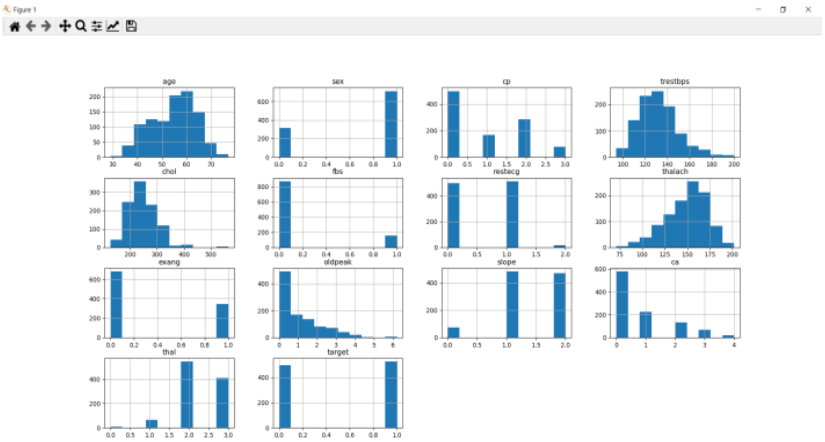

Fig 2. Distribution of all 14 attributes

3.1 Methodology

Following are the steps used for classification of data into heart disease and no-heart disease.

Step 1 :Define GNN Model

Create a GNN model that consists of several parameters

Step 2: Initialization

Set up the initial configuration of the GNN modelwhich includes specifying the dimensions for input, hidden, and output layers. Also initialize weights and biases for the connections between these layers.

Step 3: Node Embedding

Define a function of message passing between nodes in the graph which computes message based on node features and graph structures

Step 4: Aggregation and Transformation

Compute combined message at a node by aggregating messages of neighboring nodes and apply transformation which applies weights and biases to this message to generate output.

Iterate through step 4 and use backpropagation to update the weights based on the calculated loss and the all five optimizers.

Repeat this process for all batches and continue for the n number of epochs.

1. Prepare: the input node representation will be processed using Feed Forward network to produce a message. Using linear transmission we can simplify the process.

2. Aggregate: The messages originating from every node's neighbors are combined based on edge weights through various combinations and permutations, such as mean, max, and sum for every node.

3. Update: The Node-representation and aggregated message are combined and processed to produce a new state of node representation. If the combination-type is a GRU layer then the node representation and aggregated message will be stacked in a queue for the process of GRU Layer. Otherwise the node representation and aggregated message are added and processed using feed forward neural network.

The GNN Classification model follows the following approach:

a. To generate initial node representation we will use preprocessing on node feature using FFN.

b. To produce node embedding use one or more than one layer with skip connections.

c. Apply post processing using FFN to generate final node embedding.

d. Then feed the node embedding into Softmax layer to predict the node class.

Every GCL add the information which is captured from the further level of neighbors. Adding more GCL can result into over smoothing that means it can produce similar embedding for all nodes.

3.2 Optimizers

Optimizers are crucial part in neuralnetwork.To choose the correct optimizer for our application we should know how the optimizers work.

1. RMSProp Optimizer: An optimization method called RMSprop (Root Mean Square Propagation) is used to train neural networks, particularly in the context of stochastic gradient descent (SGD) and its variations. RMSprop's main idea is to scale each parameter's learning rate inversely proportionate to the moving average of the squared gradients in order to adaptively modify it. This means that the learning rate changes overtime. The value of momentum is denoted by beta which is usually set to 0.9. The below equation shows the updating rule of RMSprop optimizer.

    a. ADAM optimizer: ADAM optimizer is First-order-gradient based algorithm. It is known for its efficiency in handling sparse gradients, dealing with noisy or non-stationary problems, and providing good convergence properties for a wide range of neural network architectures and problems. The direction of update is given by the first moment normalized by the second moment. The below equation shows the updating rule of ADAM optimizer.

$$\theta_{n+1} = \theta_n - \frac{\alpha}{\sqrt{\hat{v}_n + \epsilon}} \hat{m}_n$$

    b. Adadelta optimizer: Adadelta is extension of Adagard. In Adadelta instead of summing all past squared gradient we are restricting it to the window size.

    c. Adagard optimizer: Adaptive Gradient Algorithm is using different learning rate for each parameter based on iterations. The reason behind this is we need learning rate for sparse feature parameter need to be higher compared to dense feature learning rate.

    d. SGD: SGD stands for Stochastic Gradient Descent algorithm. The problem with SGD is that we can't increase its learning rate because of the high oscillation. SGD is slow converge because it needs forward and backward propagation for every record.

4. Results and Discussions

This section shows the test accuracy for this given model. Table shows the accuracy for various optimizers, losses and metrics which is used to compile this model.

We have used Adam, RMSprop, Adadelta, Adagard and SGD optimizers. Binary Cross-entropy, Categorical Cross-entropy, Mean Absolute Error Loss and Accuracy, Binary Accuracy, Categorical Accuracy Metrics.

Table 3. Comparison of all five optimizers

Optimizer	***Binary Crossentropy***	***Categorical Cross-entropy***	***Mean Absolute Error***
	Binary Accuracy		
Adam	85%	87%	54%
RMSprop	**87%**	**92%**	**44%**
Adadelta	48%	48%	58%
Adagrad	48%	48%	68%
SGD	48%	84%	69%

Binary Cross-Entropy, also known as Binary Logarithmic Loss or Binary Cross-Entropy Loss, is a loss function used primarily in binary classification tasks within machine learning. It evaluates how well a classification model predicts the possibility that an input will belong to a particular class, usually represented by the number 1 in the model's output, which is a probability value between 0 and 1.The formula for Binary Cross-Entropy is as follows:

$$Binary\ Cross-Entropy = \frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1 - y_i).\log(1 - p_i)]$$

Where N is total number of samples or instances, y_i represents the truth label of the ith sample belongs to class 1(heart disease) and p_i represents predicted probability that the i[th] sample belongs to class 1. Binary accuracy is a metric used to evaluate the performance of a binary classification model in machine learning. It is calculated using following formula:

$$Binary\ accuarcy = \frac{no.\ of\ correct\ predictions}{total\ no.\ of\ predictions}$$

Which is mathematically expressed as follows:

$$Binary\ Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Where TP : True Positive; TN: True Negative; FP : False Positive; FN: False Negative. From the above table it is observed that RMSProp is performing better than other optimizers with accuracy of 92%. Following screenshot1shows the confusion matrix and screenshot 2 shows the overall summary of model with number of parameters generated during training.

```
Test Accuracy: 92.0%
confusion matrix
[[100   6]
 [ 10  89]]
classification report
              precision    recall  f1-score   support

           0       0.91      0.94      0.93       106
           1       0.94      0.90      0.92        99

    accuracy                           0.92       205
   macro avg       0.92      0.92      0.92       205
weighted avg       0.92      0.92      0.92       205
```

Fig 3. Confusion Matrix

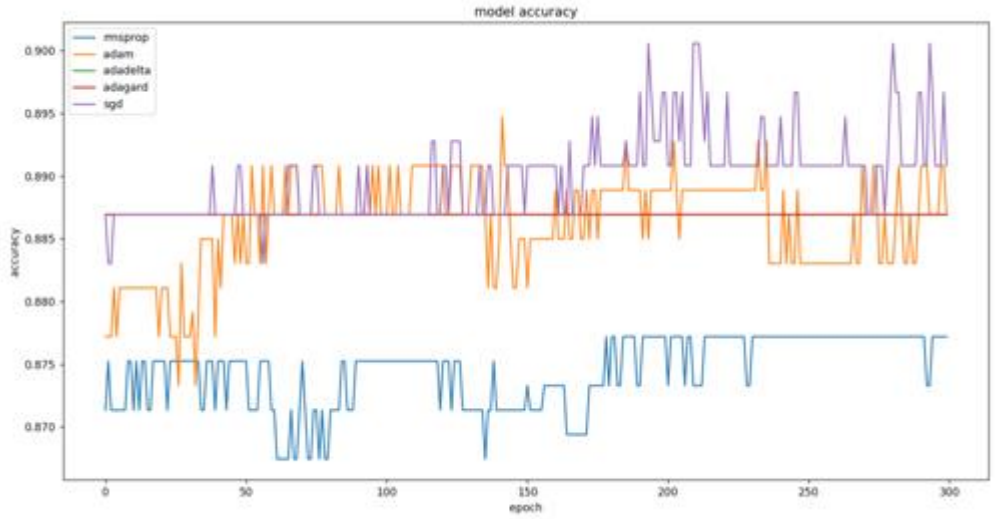

Fig 4. Model Summary

```
Model: "gnn-model"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 preprocess (Sequential)     (1, 10, 64)               20689

 graph_conv1 (GraphConvLayer  multiple                 24628
 )

 graph_conv2 (GraphConvLayer  multiple                 24628
 )

 postprocess (Sequential)    (14, 64)                  16189

=================================================================
Total params: 86,134
Trainable params: 86,134
Non-trainable params: 0
_________________________________________________________________
```

Fig 5. shows the comparison of all optimizers in terms of accuracy.

5. Conclusion

The main objective of this research was to predict presence of heart disease or not in a given sample of attributes using GNN. For this the dataset which is used is combination of four different datasets i.e Cleveland, Hungary, Switzerland, and Long Beach V created in 1988. Various optimizers are used to improve accuracy of model but RMS-Prop perform better than other optimizers with accuracy of 92%. We aim to evaluate and contrast the performance of eight Graph Neural Network (GNN) models—AGNN, ChebNet, GAT, GCN, GIN, GraphSAGE, SGC, and TAGCN—using the public health heart disease dataset. Additionally, we plan to compare their effectiveness among themselves and against traditional machine learning models like decision tree, gradient boosting, multi-layer perceptron, naive Bayes, and random forest, which will serve as our baseline models for comparison.

References

[1] Bei W. , Xiaoqing L. , Jingwei Q. , Haowen S., Zehua P., Zhi T., GNDD: A Graph Neural Network-Based Method for Drug-Disease Association Prediction, *IEEE International conference on Bioinformatics and Biomedicine* (BIBM) (2019)

[2] Tzu-An S., Samadrita C., Fan Y. Heidi J., Georges F., Quanzheng L., Keith J., Joyita D. Graph Convolutional neural network for Alzheimer's disease classification. *IEEE 16th international symposium on biomedical imaging (ISBI 2019)*

[3] Sicen L., Tao L., Haoyang D., Buzhou T., Xiaolong W., Qingcai C., Jun Y., Yi Z., A hybrid method of recurrent neural network and graph neural network for next-period prescription prediction, *International journal of machine learning and cybernetics (2020)*.

[4] Amol K. Xue B., Luyang L., Bryan P., Matt B., Martin B., Shawn O., Examining COVID-19 Forecasting using spatio-temporal Graph Neural Network, *arXiv:2007.03113v1 (2020)*

[5] Tao L., Weihua P., Qingcai C., Xioaolong W., Buzhou T., KeoG: a knowledge-aware edge-oriented graph neural network for documnet-level relation extraction, *IEEE international conference on Bioinformatics and biomedicine (2020)*

[6] Haifeng L., Hongfei L., Chen S., Liang Y., Yuan L., Bo X., Zhihao Y., Jian W., Yuanyuan S., Drug Repositioning for SARS-CoV-2 Based on Graph neural network, *IEEE International conference on Bioinformatics and Biomedicine (2020).*

[7] Waigi, R.; Choudhary, S.; Fulzele, P.; Mishra, G. Predicting the risk of heart disease using advanced machine learning approach. *Eur. J. Mol. Clin. Med.* ***2020***, *7*, *1638–1645.*

[8] Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.

[9] Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the KDD '16: 22nd *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 13–17 August 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794.

[10] Gietzelt, M.; Wolf, K.-H.; Marschollek, M.; Haux, R. Performance comparison of accelerometer calibration algorithms based on 3D-ellipsoid fitting methods. *Comput. Methods Programs Biomed.* **2013**, *111*, 62–71.

[11] K, V.; Singaraju, J. Decision Support System for Congenital Heart Disease Diagnosis based on Signs and Symptoms using Neural Networks. *Int. J. Comput. Appl.* **2011**, *19*, 6–12.

[12] Nagavelli U, Samanta D, Chakraborty P. Machine Learning Technology-Based Heart Disease Detection Models. J *Healthcare Eng*. 2022 Feb 27;2022:7351061. doi: 10.1155/2022/7351061. PMID: 35265303; PMCID: PMC8898839.

[13] Sogancioglu E., Murphy K., Calli E., Scholten E. T., Schalekamp S., Van Ginneken B. Cardiomegaly detection on chest radiographs: segmentation versus classification. *IEEE Access* . 2020;8 doi: 10.1109/access.2020.2995567.94631

[14] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Philip, S. Y. (2020). A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, *32*(1), 4-24.

[15] https://blogs.nvidia.com/blog/what-are-graph-neural-networks/ accessed on 10th Dec. 2023

[16] Kriege, N. M., Johansson, F. D., & Morris, C. (2020). A survey on graph kernels. *Applied Network Science*, *5*(1), 1-42.