Graph Application in Artificial Neural Network for Medical Diagnosis

Muhammad Farid Adilazuarda, 13518040
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
[1]*author@itb.ac.id*

***Abstract*— *Artificial Neural Networks* are one of the branch in *Artificial Neural Network* that mainly used as a tool for image recognition and Pattern Matching. Artificial neural networks are one of the main tools used in machine learning. As the "neural" part of their name suggests, they are brain-inspired systems which are intended to replicate the way that we humans learn. *Artificial neural network* is a system that can adapt to some problem by changing its structure an solve the problem based on information that flows though the system both internally or externally. Neural networks consist of input and output layers, as well as (in most cases) a hidden layer consisting of units that transform the input into something that the output layer can use. They are excellent tools for finding patterns which are far too complex or numerous for a human programmer to extract and teach the machine to recognize. By using neural network, computer can learn how to figure its own solution and an sole complex problems as we discuss in this paper is Graph Application in Artificial Neural Network for Medical Diagnosis.**

***Keywords*— *Artificial Intelligence, Artificial Neural Network. Artificial Neural Network,* Graph.**

I. Introduction

Artificial Intelligence, specifically speaking *Artificial Neural Network* as a branch of *Artificial Neural Network* is a branch of Computer Science that learn to how to make a machine or computer to think like human being. Therefore, *Artificial Intelligence* is expected to solve upcoming complex problems that hardly solved by human. One of the techniques that is a branch of *Artificial Intelligence* is *Artificial Neural Network.*

Artificial Neural Network are finding many uses in the medical diagnosis application. The goal of this paper is to evaluate *Artificial Neural Network* in medical disease diagnosis. Two cases are studied. The first one is breast cancer; data is the dataset that contains attributes of a normal breasts compared to cancerous breasts. The second is the diabetic retinopathy; data is on fundus photography, which uses a specialized camera to take a picture of a patient's retina. Each patient classified into two categories: infected and non-infected. Classification is an important tool in medical diagnosis decision support. Feed-forward back propagation neural network is used as a classifier to distinguish between infected or non-infected person in both cases. The results of applying the *Artificial Neural Network* methodology to acute nephritis diagnosis based upon selected symptoms show abilities of the network to learn the patterns matching and corresponding to symptoms of the person.

II. Theory of Graph and Neural Networks

2.1. Graph

2.1.1. Graph Definition

A graph consists of vertices and edges. Each edge has two vertices that connect each other. This pair of vertices are related to each other with a set of rule. In neural networks, graph are used to represent neurons that looks like brain. There are many types of graph based on different properties Conceptually speaking, graphs are data structures comprised of two primary components: vertices and edges. They are also sometimes referred to as networks. Graphs express relationship between pairs of items. The items are represented through vertices, while the relationships are represented through edges.

Formally, a graph is a pair of sets (V, E), where V is the set of vertices and E is the set of edges, formed by pairs of vertices.

Let graph G=(V,E)
V is a set of vertices.
E is a set of edges
V isn't allowed to be an empty set but E is allowed.

2.1.2. Graph Classification

Based on the direction of the edge, graphs can be classified in two types:

1. Non-Directed Graph
   The vertices of non-directed graphs are connected by edges that have no specific direction.
2. Directed Graph
   The vertices of non-directed graphs are connected by edges that have specific direction(s).

Graph Terminology :

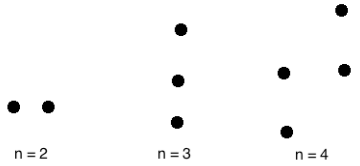
1. *Adjacent Graph*

If there are edges that connect two vertices, then those two vertices are now adjacent.

2. *Incidency Graph*
If there is a edge that connect two verices, then that edge is now incindental with those vertices.

3. *Isolated Vertex*
Vertices that had no edge that incidental with it.

4. *Null Graph*
Graph that doesn't have any edges.



Source of image : http://mathonline.wdfiles.com

5. *Degree*
Number of edges that incidental with a vertices.

6. *Path*
Path in a graph is a rank that contains names from the origin vertices and next vertices until the aimed vertices.

7. *Cycle* and *Circuit*
Line that starts and ends in the same vertices.

8. *Connected Graph*
If two vertices has a line, then they are connected. If a graph is connected, then every vertices in the set V has a line. If not all vertices are connected by that line, then those graphs is non connected graph.

9. *Subgraph*
If there is graph G = (V, E) and there is a graph G_1 = (V_1, E_1) then G_1 is a subgraph if V_1 is a subset of V and E_1 is a subset of E .

10. *Weighted Graph*
Graph that has weight in its edges.

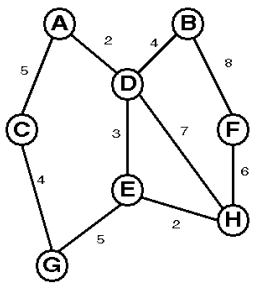


Image source : stackoverflow.com

There are different kinds of graphs, depending on whether edges have directions, whether multiple edges can connect the same pair of vertices, and whether loops are allowed.

1. A graph with directed edges is called a directed graph.
2. A graph without directed edges is called an undirected
graph.
3. A graph in which each edge connects two different
vertices and where no two edges connect the same pair of vertices is called a simple graph.
4. A graph that may have multiple edges connecting the same vertices is called a multigraph.
5. Finally, a graph that may include loops, and possibly multiple edges connecting the same pair of vertices or a vertex to itself, is called a pseudograph.

Graphs could have paths or circuits. Paths are sequences of edges that begins at a vertex of a graph and travels from vertex to vertex along edges of the graph. Meanwhile, circuits are paths which end at the vertex it begins.

Graph is considered as among the most useful mathematical models, because it's widely applicable in almost any problems, especially those related to connectivity or relationship between objects. It is used in many fields, not strictly limited to mathematics, but also includes computer science, biology, and economics.

There is a lot of representation of the graph, such as adjacency matrix, incidency matrix, and adjacency list. Because we will discuss about the image, and the image represents the matrix, adjacency matrix will relevant to the representation.

The adjacency matrix is a matrix that represent the relational between the vertex. In the position (u, v) denoted by 0 for no relation between vertex u and vertex v or there no edge between that vertex. Otherwise, if the position (u, v) denoted by 1 it has relation between vertex u and vertex v or there is an edge between those vertices. The relation between them can be represented as a weighted graph, if and only if the element in position (u, v) denoted as an integer.

The incidency matrix is a matrix that shows the relationship between the vertex has size $n \; x \; m$. The rows denoted as vertex's label, and the columns denoted as edge's label. Call the matrix as M. If the element of M at the position (i, j) have value, then the edge of j is vertex-i with the that value (if weighted graph) Otherwise, if the element of M at the position (i, j) is zero, no relation between them.

2.1.3. *Artificial Neural Networks*

Artificial neural networks are a system inspired by a biological neural network that constitute our brain. Neural network is not an algorithm, but a framework for many machine learning algorithm to process complex data input.

Machine learning algorithm that uses neural network usually doesn't need to be programmed with a specific rule that define what to expect from the input but instead the algorithm learns from the data or "answers" to learn the characteristics of the input needed to construct the correct solution or output. When a sufficient number of data has been given, the neural networks can process a new input that differs from the data given to return the correct solution. The more data given to the neural networks, the more accurate the solution would be, especially when the program is given the feedback whether the answer is correct or not.

Using Neural Network has very promising future in upcoming years, especially in medical diagnosis and applied biomedicine field of study. This is because Neural Network can be trained to solve very complex problems and give a very accurate diagnosis by data that given to them. Neural Network, especially Artificial Neural Network can be used in some particular analysis, including one that requires scanning and pattern matching to identify a disease based on the symptoms.

III. Artificial Neural Network in Breast Cancer Diagnosis

This case is aimed to know whether some condition of breasts is a cancerous symptom or non cancerous symptom from images of breast biopsy.

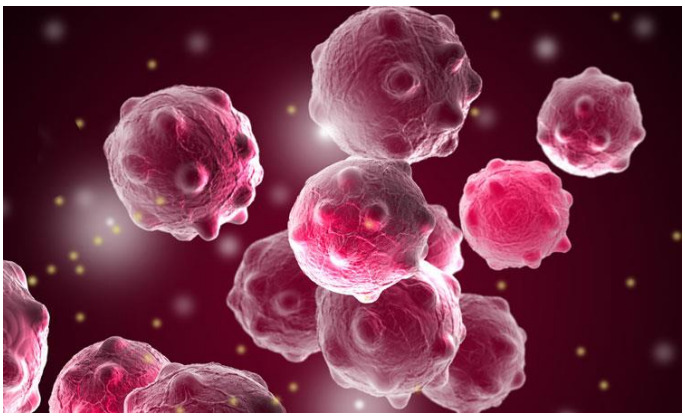


Image Source : drugtargetreview.com

A. Datasets

In this case, we used data that examine the medical diagnosis of a breast named breast_cancer.csv. This dataset contains 9 attributes that is needed. The following list summarizes the attributes information:

- clump_thickness : (1-10) Non-cancerous cells usually groped in monolayer tissue, while cancerous cells usually grouped in multilayer tissue.
- cell_size_uniformity : (1-10) Cancerous cells vary in different size and shape that scaled on 1-10.
- cell_shape_uniformity : (1-10) Cancerous cells vary in different size and shape that scaled on 1-10.
- marginal_addhesion : (1-10) Normal cells usually stick together while cancerous cells are solitary.
- single_epithelial_cell_size : (1-10) cancerous cells usually has larger epithelial cell size.
- bare_nuclei : (1-10) term used for nuclei that is not surrounded by cytoplasm (the rest of the cell).
- bland_chromatin : (1-10). Describes a uniform "texture" of the nucleus seen in cells whether the cell is coarse or not. Cancerous cells tends to be more coarse.
- normal_nucleoli : (1-10) In cancer cells the nucleoli become more prominent, and sometimes appears more often.
- mitoses : (1-10) basically, cancerous cells is the one that has uncontrolled mitosis.
- diagnose : (0 or 1) the final result f diagnosis.

Each attribute that used in this dataset contains the input and target variables of a different patient. The dataset is separated into a training(60%), generalization(20%), and testing(20%)

We can draw table and calculate the statistics by the provided data with the minim, maximum, and standard deviations of all attributes. Also, we can calculate the distributions for all variable. The following figure shows a pie chart with the numbers of cancerous and non cancerous lump in the dataset.

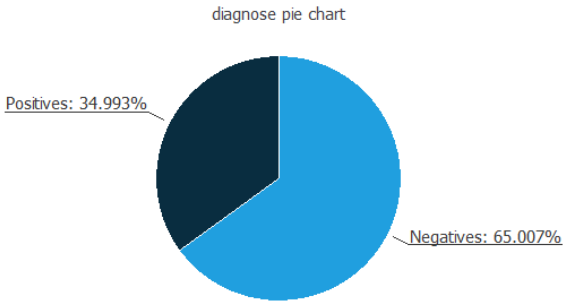


Image source : neuraldesigner.com

The inputs-targets correlations might indicate us what factors are most determining for a tumor to be cancerous or non cancerous.

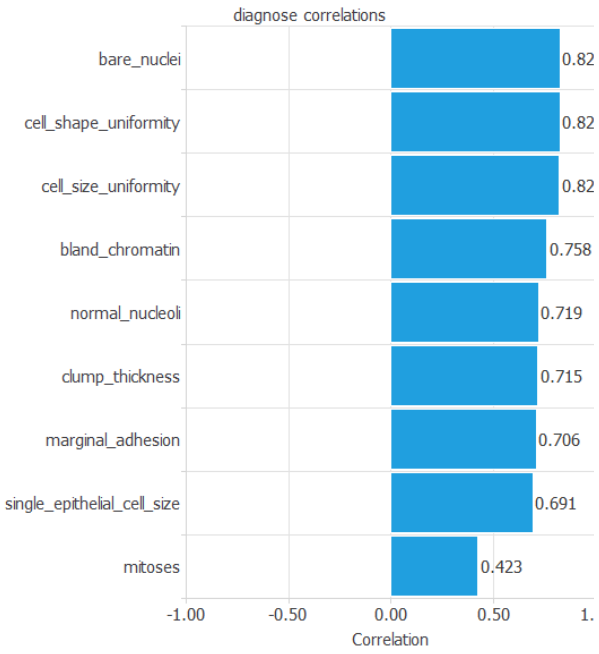


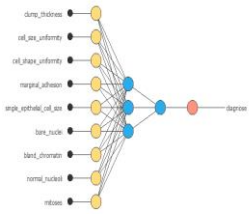
Image source : neuraldesigner.com

Here, the most correlated variables with malignant tumors are bare nuclei, cell shape uniformity and cell size uniformity.

B. Neural Network

The second step is to set a neural network to represent the classification function. For this class of applications, the neural network is composed by:

- Scaling layer. (Contains statistics and inputs)
- Perceptron layers. (Contains logistics hidden layer and logistics output layer)
- Probabilistic layer. (contains the method for interpreting the outputs as probabilities)



C. Training Strategy and Model Selection

The objective of model selection is to find the network architecture with best generalization properties, that is, that which minimizes the error on the selection instances of the data set.

More specifically, we want to find a neural network with a selection error less than 0.072 WSE, which is the value that we have achieved so far.

Order selection algorithms train several network architectures with different number of neurons and select that with the smallest selection error.

The incremental order method starts with a small number of neurons and increases the complexity at each iteration. The following chart shows the training error (blue) and the selection error (orange) as a function of the number of neurons.

We want to find a neural network with the minimum selection error. So, we need an Order selection algorithms that trains several network architectures with different number of neurons and select that with the smallest selection error.

The incremental order method starts with a small number of neurons and increases the complexity at each iteration. The following chart shows the training error (blue) and the selection error (orange) as a function of the number of neurons.

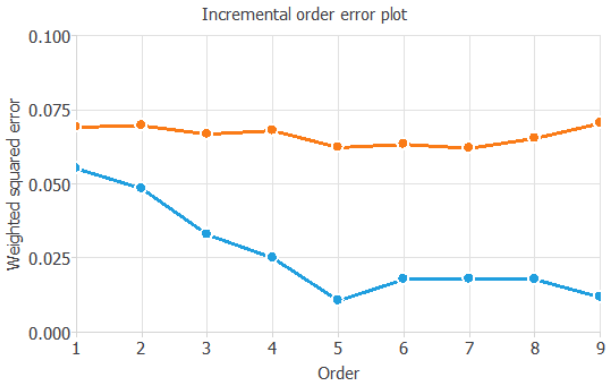


Image source : neuraldesigner.com

final architecture for the neural network:

Image source : neuraldesigner.com

D. Testing and Analysis

The main goal of testing analysis is to validate the generalization performance of the trained neural network. To validate a classification technique, we need to compare the values provided by this technique to the actually observed values.

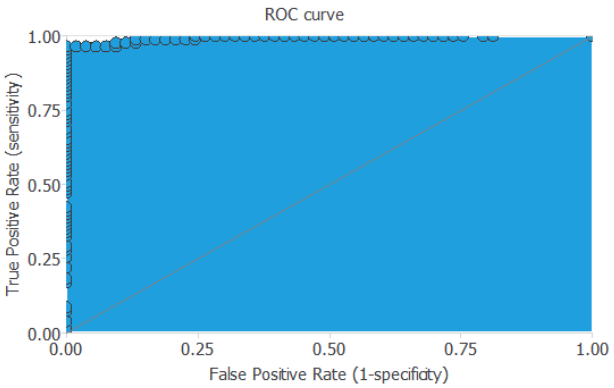


Image source : neuraldesigner.com

The following table contains the elements of the confusion matrix. This matrix contains the true positives, false positives, false negatives, and true negatives for the variable diagnose.

	Predicted positive	**Predicted negative**
Real positive	129	3
Real negative	1	37

The binary classification tests are parameters for measuring the performance of a classification problem with two classes:

- **Classification accuracy** (ratio of instances correctly classified): 97.6%
- **Error rate** (ratio of instances misclassified): 2.4%
- **Sensitivity** (ratio of real positive which are predicted positive): 99.2%
- **Specificity** (ratio of real negative which are predicted negative): 92.5%

E. Model Deployment

Once the generalization performance of the neural network has been tested, the neural network can be saved for future use in the so-called model deployment mode.

We can diagnose new patients by calculating the neural network outputs. For that we need to know the input variables for them. An example is the following:

clump_thickness (1-10): 5
cell_size_uniformity (1-10): 4
cell_shape_uniformity (1-10): 2
marginal_adhesion (1-10): 3
single_epithelial_cell_size (1-10): 3
bare_nuclei (1-10): 3
bland_chromatin (1-10):2
normal_nucleoli (1-10): 4
mitoses (1-10): 1
diagnose: Negative

The mathematical expression represented by the neural network is written below. It takes the inputs clump_thickness, cell_size_uniformity,cell_shape_uniformity,marginal_adhesion ,single_epithelial_cell_size,bare_nuclei,bland_chromatin,normal_nucleoli and mitoses to produce the output diagnose.

```
scaled_clump_thickness = (clump_thickness-4.44217)/2.82076;
scaled_cell_size_uniformity = (cell_size_uniformity-3.15081)/3.06514;
scaled_cell_shape_uniformity = (cell_shape_uniformity-3.21523)/2.98858;
scaled_marginal_adhesion = (marginal_adhesion-2.83016)/2.86456;
scaled_single_epithelial_cell_size = (single_epithelial_cell_size-3.23426)/2.22309;
scaled_bare_nuclei = (bare_nuclei-3.54466)/3.64386;
scaled_bland_chromatin = (bland_chromatin-3.4451)/2.4497;
scaled_normal_nucleoli = (normal_nucleoli-2.86969)/3.05267;
scaled_mitoses = (mitoses-1.60322)/1.73267;

y_1_1 = Logistic (-1.35621+ (scaled_clump_thickness*-2.54409)+ (scaled_cell_size_uniformity*-5.01572)
+ (scaled_cell_shape_uniformity*-3.39576)+ (scaled_marginal_adhesion*-0.278873)+ (scaled_single_epithelial_cell_size*-2.61646)
+ (scaled_bare_nuclei*-5.51018)+ (scaled_bland_chromatin*-0.979982)+ (scaled_normal_nucleoli*-1.71412)+ (scaled_mitoses*0.410197));
non_probabilistic_diagnose = Logistic (3.94959+ (y_1_1*-9.14654));

diagnose = Probability(non_probabilistic_diagnose);
```

IV. Experiment and Results

The neural network and analysis is created by using neural designer software. The software is provided and developed by neuraldesigner.com, and uses no special neural network libraries. But this short program can recognize breast cancer by training error = 0.054 WSE and selection error = 0.072 WSE, without human intervention. The breast cancer database here was was obtained from the University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg. The database contains 684 data from different people.

V. Conclusion

The neural network used to recognize a breast cancer is just one of breakthrough that could improvise the future of human race. Imagine that people could develop technology that instead just detecting or diagnosing cancer, it can cure it too. Beside that, there are so much more Neural Network application that could make human life easier and healthier.

VI. Appendix

The software used in this paper can be accesed on https://www.neuraldesigner.com/free-trial.

VII. Acknowledgment

The The author of this paper would like to thank Allah SWT because of the blessings given so that the author can finish this paper to complete Discrete Mathematics assignment and of course to Dr. Ir. Rinaldi Munir because of his guidance and the knowledge given in the class. And also the author would like to thank all of people who helped author and every references that was used in this paper.

REFERENCES

[1] [1] Rinaldi Munir, Matematika Diskrit, Bandung : Penerbit Informatika, Palasari

[2] https://www.neuraldesigner.com/solutions/medical-diagnosis

[3] https://towardsdatascience.com/medical-diagnosis-with-a-convolutional-neural-network-ab0b6b455a20

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 3 Desember 2017

Muhammad Farid Adilazuarda