Jurnal Mantik, 5 (3) (2021) 1759-1765

Published by:Institute of Computer Science (IOCS)

Jurnal Mantik

Journal homepage: www.iocscience.org/ejournal/index.php/mantik/index

Classification Of Sexually Transmitted Infectional Diseases Using Artificial Neural Networks

Agung Mustika Rizki[1], Hendra Maulana[2], Dhian Satria Yudha Kartika[3], Gusti Eka Yuliastuti[4]

[1] Department of Informatic, Faculty of Computer Science, Universitas Pembangunan Nasional (UPN) Veteran Jawa Timur Surabaya, Indonesia
[23] Department of Information System, Faculty of Computer Science, Universitas Pembangunan Nasional (UPN) Veteran Jawa Timur Surabaya, Indonesia
[4]Teknik Informatika, Fakultas Teknik Elektro dan Teknologi Informasi, Institut Teknologi Adhi Tama Surabaya, Indonesia

E-mail: agung.mustika.if@upnjatim.ac.id

ARTICLEINFO

Article history:

Received: September 14, 2021
Revised: October 20, 2021
Accepted: November 02, 2021

Keywords:

Sexually Transmitted Infections,
Classification,
Artificial Neural Network.

ABSTRACT

Many diseases are caused by bacteria, some of which can be easily noticed by ordinary people for immediate treatment. However, not with this one disease, namely sexually transmitted infections (STIs). This STI disease can be spread mainly through sexual intercourse, both vaginal, anal and oral sex. Some STI diseases can also be transmitted in non-sexual ways, such as through needles, blood or other blood products. Indonesia is one of the countries whose handling can be said to be not optimal as in several other countries. This is the result of a lack of education on STI diseases in the community. Based on this background, it can be concluded that there is a need for an intelligent system to classify STI diseases based on their symptoms. Therefore, the authors propose this research topic by applying the Artificial Neural Network method. Based on the test results, the application of the Artificial Neural Network method shows that 80% of the predicted data is in accordance with the actual data.

1. Introduction

At this time the whole world including Indonesia is facing the COVID-19 pandemic which it is not known when it will end. All aspects of daily life are changing and even tend to be disturbed. One aspect that is very disturbed is the health aspect. With this pandemic, we are asked to take better care of ourselves from all things related to viruses and bacteria. We are required to always maintain cleanliness and health in order to avoid all diseases and be able to continue our daily activities.

Many diseases are caused by bacteria, some of which can be easily recognized by ordinary people to be treated immediately. However, not with this one disease, namely sexually transmitted infections (STIs). This disease, which is in everyone's private area, can occur not only in sexually active people, but also in inactive people. Several factors can cause sexually transmitted diseases, namely the presence of bacterial infections [1].

More than 30 types of bacteria, viruses and parasites are known to be transmitted through sexual contact. Some of them cause an increase in the spread of STIs. Some viral infections that can be cured include: Syphilis, Gonorrhea, Chlamydia and Trichomoniasis. Besides that, there are also incurable viral infections, namely Hepatitis V, Herpes Simplex Virus (HSV), HIV and Human Papillomavirus (HPV). Some of these incurable infections can only be reduced with special treatment [2].

This STI can be spread mainly through sexual intercourse, both vaginal, anal and oral sex. Some STIs can also be transmitted non-sexually, such as through needles, blood or other blood products. Some STI diseases such as syphilis, hepatitis B, HIV, chlamydia, gonorrhea, herpes and HPV can also be transmitted from mother to child during pregnancy and childbirth.

According to the World Health Organization (WHO), the number of people with STIs in the Southeast Asian region is very high. It is estimated that 78 million people are infected with this STI. A total of 7.2 million cases were infected with C. trachomatis, 25.4 million cases were infected with N. gonorrhoeae, 3 million cases were infected with syphilis and 42.9 million cases were infected with T. vaginalis. This situation will continue to increase due to several factors, including: the increasing number of sexually active people, changes in behavior, social changes to changes in patterns in the practice of preventing and treating STIs in each country [3].

Indonesia is one of the countries whose handling can be said to be not optimal as in several other countries. This is a result of the lack of education about STIs in the community. It is necessary to increase IEC efforts (Communication, Information, Education) related to this STI disease in order to reduce the number of people infected with STIs. The thing to remember is that STIs are very closely related to social change, so the main goal of IEC efforts is to help survivors to treat STIs they suffer so that they can change risky sexual behavior into safe sexual behavior [4].

The public's lack of understanding of the symptoms experienced when infected with an STI disease has resulted in not optimal early detection steps to handling STI survivors. In addition to the IEC efforts carried out by all elements of society, adequate facilities are also needed in terms of early detection of STI diseases [5].

The provision of services to STI survivors carried out by these doctors often encounters a problem, namely the symptoms that become the input data for survivors are incomplete. This happens because the survivors do not know for sure which are the so-called symptoms of this STI. So doctors often have difficulty diagnosing and making decisions for the right treatment

There are several previous studies that have been done related to STIs. This STI disease does not only attack adults, but can also attack at the age of children. This can happen due to the increasing number of cases of sexual violence against children which indirectly increases the risk of contracting this STI disease in children [6]. Some children experience pain due to sexual violence but not all are aware of the risk of transmitting STIs. Ignorance of the risk of STI transmission can not be separated from the lack of sexual education in children and also the trauma experienced so that children cannot retell what happened for an initial diagnosis.

According to Tuntun [5], if STIs are not treated properly, they can result in reproductive infections, prolonged illness, infertility and even death. In diagnosing an STI disease, complete symptom input data is needed.

Research related to the classification of STIs has been carried out previously by Yuliastuti et al (2018) by applying several data mining methods, namely: Naïve Bayes, K-Means and K-Nearest Neighbors. There are 139 sample data then divided into 109 training data and 30 testing data. The results of the application of the three methods obtained accuracy of 76.67%, 10% and 90%, respectively.

There is a study conducted by Suhartanto et al [8] related to skin diseases in children by applying a backpropagation artificial neural network. The conclusion of this research is that the backpropagation method is considered successful. Based on the input data obtained from the past, the backpropagation method is able to diagnose skin diseases in children with a fairly high accuracy of 90%.

The backpropagation method in artificial neural networks is a general form of the Least Mean Square algorithm where the network weight is entered to minimize errors in the desired output (2014). Widodo et al applied the backpropagation method to predict dengue fever. The results obtained are in the form of a prediction accuracy value of 99%.

Based on this background, it can be concluded that there is a need for an intelligent system to classify STIs based on the symptoms that arise. Therefore, the author proposes this research topic by applying the artificial neural network method. Artificial sound network is one method that can produce quite high accuracy.

2. Method

Artificial neural networks have the ability to learn from the existing environment. The workings of this artificial neural network are almost the same as humans, where the learning process that occurs includes observation and repetition until the desired condition is achieved [8]. Artificial neural networks have the ability to process all the inputs obtained to provide an output [10].

The backpropagation algorithm is one of the many algorithms in artificial neural networks. The backpropagation algorithm has more than 1 hidden layer. The function of this hidden layer is to adjust the weights during the training process. This algorithm is quite popular due to its computationally efficient implementation [11].

The backpropagation algorithm has a unique characteristic, namely after doing feed forward, back propagation will be carried out. The process aims to make changes in weight. The process will stop until it reaches the specified iteration or gets the minimum error.

The data that became input in this study were 31 variables consisting of 28 symptoms of sexually transmitted diseases, 1 gender variable, 1 risk factor variable and 1 bias. The various risk factors referred to are shown in Table 1.

Table 1

Risk Factor

Variable	Risk Factor
R1	Female Sex Worker
R2	Male Sex Worker
R3	Sex Worker Customers
R4	Woman Man
R5	Male Sex Man
R6	Prison Inmates
R7	Injection Drug User
R8	Others

As for the input variable in the form of symptoms being experienced by the patient as shown in Table 2.

Table 2

Disease Symptoms

Variable	Disease Symptoms
S1	Itching of the genital organs
S2	Pain when urinating
S3	Swelling in the groin
S4	Swelling in the scrotum
S5	Swelling of the glands in women
S6	Bleeding after sexual intercourse
S7	Lower abdominal pain
S8	Blisters in the genital area
S9	Painful nodules
S10	Wound (ulcer)
S11	Genital warts
S12	Bump
S13	DTV (vaginal discharge)
S14	DTS (cervical discharge)
S15	DTU (urethra discharge)
S16	DTA (fluid discharge from the anus)
S17	DTM (fluid discharge from eye)
S18	Wounds that don't hurt
S19	Smelly discharge
S20	Pain in the vagina during sexual intercourse
S21	Single wound on genital organs
S22	Wound more than 1 week
S23	Watery nodules
S24	After breaking leaves a wound
S25	Bad smell from vagina
S26	Vulva swelling redness
S27	Vaginal bleeding
S28	Puffy eyes

By these 31 input variables, the data is processed using Artificial Neural Network Backpropagation to produce output in the form of disease types. In this study, there were 16 types of diseases that became the output variables as shown in Table 3.

Table 3
Type of Disease

Variable	Type of Disease
D1	Syphilis
D2	Genital Herpes
D3	Chlamydia
D4	Vaginal Trichomoniasis
D5	Genital Warts
D6	Chancroid
D7	Lymphogranuloma Venereum
D8	Granuloma Inguinale
D9	Urethritis
D10	Chlamydiosis
D11	Servictis
D12	Candidiasis
D13	Bacterial Vaginosis
D14	Trichomoniosis
D15	Condyloma acuminata
D16	Moluskum Kontagiosum

2.1 Backpropagation Algorithm

In general, the steps of the Backpropagation Algorithm are as follows:

1. Initialize all weights using small random numbers.
2. Perform steps 3 to 9 if the desired condition has not been reached.
3. Perform steps 4 to 8 for each training data.

Stage I: Feed Forward

4. Calculate the result on the hidden layer $z_j$where $j = 1, 2, 3, \dots , p$.

$$z\_net_j = v_{j0} + \sum_{i=1}^{n} x_i v_{ji} \ (1)$$

Calculating activation with binary sigmoid function:

$$z_j = f\left(z_{net_j}\right) = \frac{1}{1+\exp\left(-z_{net_j}\right)} \ (2)$$

The output of the activation function is sent to all hidden layers.

5. Calculate the result at the output y_k where $k = 1, 2, 3, \dots , m$.

$$y\_net_k = w_{k0} + \sum_{j=1}^{p} z_j w_{kj} \ (3)$$

$$y_k = f(y\_net_k) = \frac{1}{1+\exp\left(-y\_net_k\right)} \ (4)$$

Stage II: Back Propagation

6. Calculate the factor value δ at the output based on the error value in each output y_k.

$$\delta_k = (t_k - y_k) f'\left(y_{net_k}\right) = (t_k - y_k) y_k (1 - y_k) \ (5)$$

δ_k weights are used to change the next layer.

7. Calculate the factor δ in the hidden layer based on the error value in hidden layer z_j.

$$\delta\_net_j = \sum_{k=1}^{m} \delta_k w_{kj} \ (6)$$

Factor δ for hidden layer:

$$\delta_j = \delta\_net_j f'^{z_{net}j} = \delta_{net_j} z_j (1 - z_j) \ (7)$$

To calculate the weight layers:

$$\Delta v_{ji} = \alpha\ \delta_j x_i \ (8)$$

$$\Delta w_{kj} = \alpha\ \delta_k z_j \ (9)$$

Stage III: Weight Changes

8. Change the output weight:

$$w_{kj}(baru) = w_{kj}(lama) + \Delta w_{kj} \ (10)$$

Change the output bias weight:

$$v_{ji}(baru) = v_{ji}(lama) + \Delta v_{ji} \ (11)$$

9. Calculate the error value with MSE:

$$MSE = \frac{\Sigma(target - y)^2}{jumlah\_record} \quad (12)$$

10. Test the stop condition when the maximum iteration or condition has reached the error limit.expected

2.2 Calculation Accuracy

The test data in the feed forward needs to be calculated for accuracy using the following MAPE values:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{target_i - y_i}{target_1} \times 100\% \quad (13)$$

3. Result and Analysis

The data used in this research case study amounted to 139 patient data, where the data was divided into 2 parts, namely 109 data as training data and 30 data as test data. The first test is to test the number of hidden nodes. With a learning rate of 0.1 and an epoch of 1000, the test results are as shown in Figure 1.

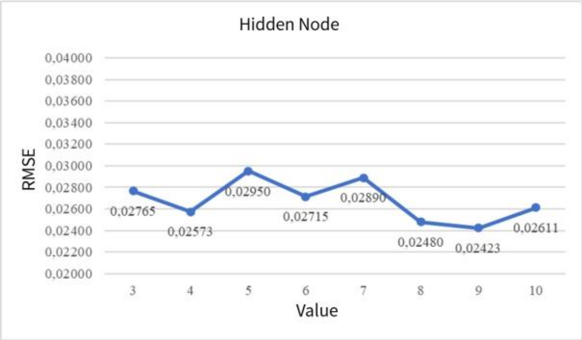

Fig 1. Hidden Node Test Results

Based on the test results in Figure 1, the smallest RMSE value is obtained with 9 hidden nodes. The results of this test will be used for further testing, namely testing the learning rate value. With 9 hidden nodes and 1000 iterations, the test results are obtained as shown in Figure 2.

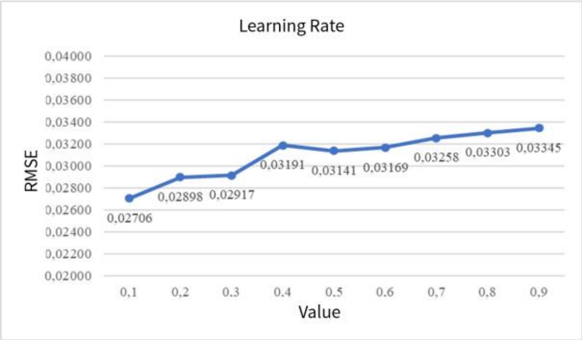

Fig 2. Learning Rate Test Results

Based on the test results in Figure 2, the smallest RMSE value was obtained using a learning rate of 0.1. The results of this test will be used for further testing, namely the number of iterations. With 9 hidden nodes and a learning rate of 0.1, the test results are as shown in Figure 3.

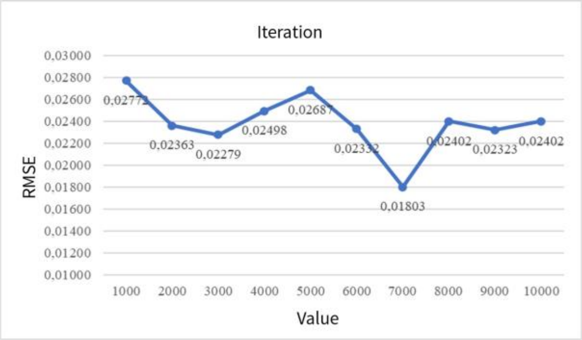

Fig 3.Iteration Test Results

Based on the test results in Figure 3, the smallest RMSE value is obtained with 7000 iterations. The three best parameters that have been obtained based on previous tests, will be used to perform the final test using 30 test data. The test results obtained are as shown in Table 4.

Table 4
Classification Results

Data	Actual	Results
Data1	D14	D14
Data2	D16	D16
Data3	D16	D11
Data4	D15	D15
Data5	D8	D8
Data6	D13	D13
Data7	D10	D10
Data8	D4	D4
Data9	D1	D1
Data10	D5	D5
Data11	D6	D6
Data12	D6	D6
Data13	D12	D12
Data14	D1	D1
Data15	D14	D14
Data16	D8	D8
Data17	D11	D11
Data18	D11	D11
Data19	D13	D13
Data20	D12	D12
Data21	D11	D11
Data22	D2	D2
Data23	D4	D4
Data24	D9	D4
Data25	D9	D4
Data26	D3	D3
Data27	D2	D2
Data28	D15	D15
Data29	D7	D14
Data30	D7	D11

The test results listed in Table 4 show that there are 6 predicted data that do not match the actual data. This can happen because the symptoms are similar to one another. In addition, the discrepancy between the prediction results with the actual data can occur because there are many classes of disease.

4. Conclusion

Based on the tests that have been carried out by the author, it can be concluded that the Artificial Neural Network Backpropagation Algorithm is suitable to be applied to predict sexually transmitted diseases and other similar case studies. It is proven by the application of the Artificial Neural Network Backporpagation Algorithm, the prediction results that match the actual data are 80%.

For further research, researchers need to modify or optimize the Backpropagation Algorithm for Artificial Neural Networks so that the prediction results can increase.

References

[1] B. Stoner, “Reproductive Health: Sexually Transmitted Infections – Overview,” *International Encyclopedia of Public Health*, vol. 6, pp. 306–314, 2017.

[2] D. McCormack and K. Koons, “Sexually Transmitted Infections,” *Emerg. Med. Clin. North Am.*, vol. 37, no. 4, pp. 725–738, 2019, doi: 10.1016/j.emc.2019.07.009.

[3] World Health Organization, “Media Centre-Sexually Transmitted Infections(STIs),” 2016. http://www.who.int/mediacentre/factsheets/fs110/en/.

[4] T. Y. Aditama, *Pedoman Nasional Penanganan Infeksi Menular Seksual*. Jakarta: Kementrian Kesehatan Republik Indonesia, 2011.

[5] M. Tuntun, “Faktor Resiko Penyakit Infeksi Menular Seksual (IMS),” *J. Kesehat.*, vol. 9, no. 3, p. 419, 2018, doi: 10.26630/jk.v9i3.1109.

[6] G. Hornor, “Sexually Transmitted Infections and Children: What the PNP Should Know,” *Journal of Pediatric Health Care*, vol. 31, no. 2, pp. 222–229, 2017.

[7] G. E. Yuliastuti, A. N. Alfiyatin, A. M. Rizki, A. Hamdianah, H. Taufiq, and W. F. Mahmudy, “Performance analysis of data mining methods for sexually transmitted disease classification,” *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, pp. 3933–3939, 2018, doi: 10.11591/ijece.v8i5.pp3933-3939.

[8] S. R. Suhartanto, C. Dewi, and L. Muflikhah, “Implementasi Jaringan Syaraf Tiruan Backpropagation untuk Mendiagnosis Penyakit Kulit pada Anak,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 7, pp. 555–562, 2017.

[9] W. Widodo, A. Rachman, and R. Amelia, “Jaringan Syaraf Tiruan Prediksi Penyakit Demam Berdarah dengan Menggunakan Metode Backpropagation,” *J. IPTEK*, vol. 18, no. 1, pp. 296–304, 2014, [Online]. Available: http://digilib.batan.go.id/e-prosiding/File Prosiding/Informatika/lkstn/LKSTN2012/Arie-Q2012.pdf.

[10] N. P. Sakinah, I. Cholissodin, and A. W. Widodo, “Prediksi Jumlah Permintaan Koran Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 7, pp. 2612–2618, 2017.

[11] N. R. Sari, W. F. Mahmudy, and A. P. Wibawa, “Backpropagation on Neural Network Method for Inflation Rate Forecasting in Indonesia,” *Int. J. Adv. Soft Comput. Its Appl.*, vol. 8, no. 3, 2016.