



Bayesian Network Model to Distinguish COVID-19 for Illness with Similar Symptoms

E Luthfi¹, A W Wijayanto^{1,2}

¹ Politeknik Statistika STIS, Jl. Otto Iskandardinata No.64C, Jakarta, Indonesia

² BPS-Statistics Indonesia, Jl. Dr. Sutomo 6-8, Jakarta, Indonesia

*Corresponding author's e-mail: 221810266@stis.ac.id, ariewahyu@stis.ac.id

Abstract. Numerous diseases and illnesses exhibit similar physical and medical symptoms, such as COVID-19 and its similar disguised illness (common cold, flu, and seasonal allergies). In this study, we construct a Bayesian Network model to distinguish such symptom variables in a classification task. The Bayesian Network model has been widely used as a classifier comparable to machine learning models. We develop the model with a scoring-based method and implement it using a hill-climbing algorithm with the Bayesian information criterion (BIC) score approach. Experimental evaluations using publicly available Mayo Clinic based data using this Bayesian Network model that present Directed Acyclic Graph (DAG) which can show the relationship between the similar symptoms and the type of disease with Conditional Probability Table (CPT). This model shows a promising accuracy performance up to 93.14% which is better than the performance of other machine learning classifiers, including the Support Vector Machine (SVM) and the ensemble approaches such as Random Forest (RF), while slightly smaller than that of the neural networks (NN).

1. Introduction

According to the World Health Organization (WHO), coronavirus disease (COVID-19) is an infectious respiratory disease caused by a SARS-CoV-2 virus infection that was recently discovered in the current pandemic era. Most people who contract COVID-19 will experience mild to moderate symptoms and will recover without special treatment. All corners of the world are very badly affected by the virus. Everyone has deep concerns and take various preventive ways that are done to protect themselves from the virus, one of which is by keeping a distance from people who have symptoms arising from the COVID-19 virus. The most common symptoms of COVID-19 are fever, cough, and fatigue. But there are many other possible signs and symptoms. However, COVID-19 cannot be detected easily because other conditions have similar symptoms. Several other illnesses have similar symptoms, namely common colds, seasonal allergies, and influenza (the flu) which certainly have differences in certain symptoms.

This problem can be solved by building a model for diagnosing COVID-19 and similar diseases. With this model, it can greatly assist doctors in making the initial diagnosis to patients can be solved, addition to the knowledge and experience of doctors can be maintained to provide the results of disease diagnosis appropriately and effectively.

The symptoms and types of this disease will be variable in making a diagnosis of the disease. This study aims to build a model that represents the relationship between variable symptoms and diseases using the Bayesian Network as well as we using the resulting model to classify COVID-19 disease and



similar diseases. The Bayesian Network approach is a probabilistic method to describe causal relationships among variables in an observed environment. Currently, a learning method has been developed that makes it possible to build Bayesian network structures directly from the database. These methods that will be used in this study are search and scored-based. In its application, Bayesian Network has been widely applied in the field of health, especially to diagnose various types of diseases.

In previous research, the Bayesian Network was used to make medical decisions on the diagnosis of psychiatric illness [1]. In the development of this model, the Bayesian Network was used to look at the relationship between factors and the results of diagnosis [2]. Based on previous research Bayesian Network produces good accuracy for the diagnosis of diseases, so in this study, Bayesian Network is used to diagnose COVID-19 disease and similar diseases, common colds, flu, and allergies. The research that will be conducted includes the construction of the Bayesian Network structure, determining the parameters of each symptom and disease, probabilistic inference, and comparisons to other classification algorithms, among others, Random Forest, SVM, and Neural Network as an evaluation of the adopted model.

2. Literature Review

2.1. Bayesian Network

Bayesian Network is one method to classify the simple probability that is mixed in the Bayes Theorem. This method became popular in the last decade because it was used as a learning machine and for various applications. Bayesian Network is also one of the Probabilistic Graphical Models (PGM) based on graph theory and probability. Graph theory is a theory that is directly related to the form of representation obtained while probability theory is a theory directly related to data [3].

Bayesian Network method is a good method in machine learning based on training data, with conditional probability of course. The Bayesian Network method can also be used to perform descriptive and predictive functions. Descriptive has a function to learn the relationships that occur in each variable or also called structural learning.

In building the Bayesian Network, the structure is built by applying the Bayes theorem. Conditional opportunity is a calculation of opportunities where an event A is known if event B has been known first, so it is denoted as $P(B)$. This theorem is used to calculate the chances of a data set is included in a particular class based on existing data references. The basic formula of the Bayes Theorem is:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

which can be denoted as follows:

$$P(A|B) = \frac{p(A)P(B|A)}{P(B|A)P(A)+P(B|\bar{A})P(A)} \quad (2)$$

which is a degree of belief (trust) or Probabilistic information. Prior Probability is denoted as $P(X)$ serves as a Probabilistic reference if no other biased information is used as an event or condition.

$$P = \frac{X}{A} \quad (3)$$

where P indicates the prior value, X is a number of data in each class, and A is the sum of data.

Conditional Probability Table (CPT) is a table containing a probabilistic possibility A and B. Conditional Probability itself is part of the quantitative component of the Bayesian Network Method. If in the process new information or new value has been obtained then the value or new information from the Probabilistic is used as a reference. This probability is referred to as conditional probability [4] which can be expressed as follows:

$$CPT = \frac{P(A|B)}{P(B)} \quad (4)$$



where CPT is the conditional probability, $P(A|B)$ represents the probability of event A occurring on the condition that the possibility of that B occurs, and $P(B)$ denotes the probability of occurrences of data B.

Normalizing Constant or also called Constant Normalization is a step used in the Bayes method to normalize conditional probability results. In the Bayes Theorem, it is mentioned that the comparison between Posterior probability is the same as Conditional Probability. Where $P(A)$ is the value of prior assuming the Prior value is true while $P(B|A)$ is the value of Conditional Probability of the data or symptoms inputted.

Bayesian networks consist of two components Directed Acyclic Graph (DAG) and Conditional Probability Table (CPT). DAG is the graphical structure of the Bayesian networks graph which is a directional graph without a directional cycle. The DAG consists of nodes and edges. Nodes represent random variables and edges represent a direct dependency relationship and can also be interpreted as direct influences (causations) between the variables they connect to. The absence of edge indicates a conditional freedom relationship between variables. This graph is used to represent the knowledge of an uncertain domain. In particular, each node in the graph is a random variable, while the tip between nodes represents a probabilistic that relies on among the corresponding random variables. These dependency conditions in graphs are often estimated using computational statistics. Therefore, this model combines the principles of graph theory, probability theory, computer science, and statistics. The set of parameters is defined as the conditional probability distribution for each variable. On the Bayesian network, nodes correspond to random variables. Each node is associated with a set of conditional opportunities, $p(x_i|A_i)$ so x_i is a variable associated with the node and A_i is the set of the parent in the graph [5].

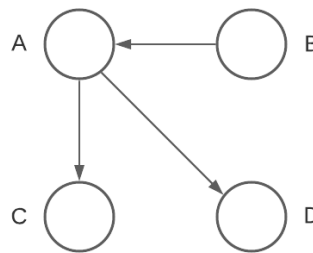


Figure 1. Examples of Simple Structure of Bayesian Networks [5].

In general, if there is node $U = U_1 \dots U_n$, the joint probability function for each Bayesian Networks is as follows:

$$P(U) = \prod_i^n P(P_k \vee \text{parents}(P_k)) \quad (5)$$

Based on equation 5, it can be known that you are a representation of a node contained in Bayesian Networks, then $P(U)$ is a product of all conditional probabilities contained in the graph structure, and parents (P_k) is the parent of the P_k node.

$$P(A, B, C, D) = P(A).P(B|A).P(C|B) \dots P(C \vee B) \quad (6)$$

Equation 6 is for calculating joint probability contained in Figure 1 graph A[5].

2.2. Random Forest

Random Forest is a compound tree method where we construct numerous decision trees which later form a colloquially known as "a forest", from which the majority vote is performed to infer the final decided class. Random Forest offers a potential benefit that prevents overfitting in data sets when it achieves maximum accuracy. However, in a class imbalance case, one of the problems in data mining, Random Forest is also facing a difficulty. This happens at a time when the sample size of the minority class is significantly smaller than that of the majority class [6].

2.3. Support Vector Machines (SVM)



SVM is a technique designed to find the optimal separator *hyperplane* that has a maximum margin [7]. SVM linear classification can be calculated by Equations 7 and 8 [8].

$$[(w^T, x_i) + b] \geq 1 \text{ for } y_i = +1 \quad (7)$$

$$[(w^T, x_i) + b] \leq -1 \text{ for } y_i = -1 \quad (8)$$

Equations 7 and 8 can be solved using the *Lagrange Multiplier* function. As for finding the best *hyperplane* can be used *Quadratic Programming (QP) Problem* method [9]. In non-linear cases, kernel methods can be used in the initial data *set* to feature. With the kernel, data features are mapped from low dimensions to new features with relatively higher dimensions [10].

2.4. Neural Network

Neural Network derives weight values for transforming values from *input nodes* to *output nodes* from a particular learning algorithm [11]. Weight value updates are generally formulated in Equation 9.

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n) \quad (9)$$

where $\Delta w_{ij}(n)$ is calculated by learning algorithms and $w_{ij}(n)$ is a randomly determined initial weight at the initiation stage. Neural Network can solve difficult and various problems because this method has fault tolerance capability to a certain extent, so this method can still produce output even though the data is incomplete [12].

2.5. Model Evaluation

Model evaluation is done after classifier classification of data testing. In this final task, the system performance will be measured by accuracy, precision, and recall values [13].

Accuracy is the ratio of the accuracy of the system in classifying documents to their appropriate class. Precision is a ratio of the number of relevant documents found to the total number of documents found by the classifier. Precision is a measure of how close is the predicted class to the genuine class. The recall is the ratio of the number of relevant documents recovered to the total number of documents in a document that are considered relevant. Precision and Recall can be calculated by creating a confusion matrix table first. Examples of confusion matrix tables can be seen in Table 1, and equations for calculating accuracy, precision, and recall values can be seen in equations 10, 11, and 12.

Table 1. Confusion matrix

Classifier\Actual	Actual Positive	Actual Negative
Classified Positive	True Positive	False Positive
Classified Negative	False Negative	True Negative

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

$$Precision = \frac{TP}{TP+FP} \quad (11)$$

$$Recall = \frac{TP}{TP+FN} \quad (12)$$

3. Research Methods

3.1. Stages of Research

Building a Bayesian Network model to classify COVID-19 disease and others consists of several stages, namely (1) Identification of problems, (2) Data collection, (3) Application of Bayesian Network method, and (4) Evaluation of models built with other models [14]. The research framework of this study is illustrated in Figure 2 based on Polancik [21] and Dawson [22] scheme where the schematic diagram the motivated problem, proposed method, indicators, research objective, and measurements are described.

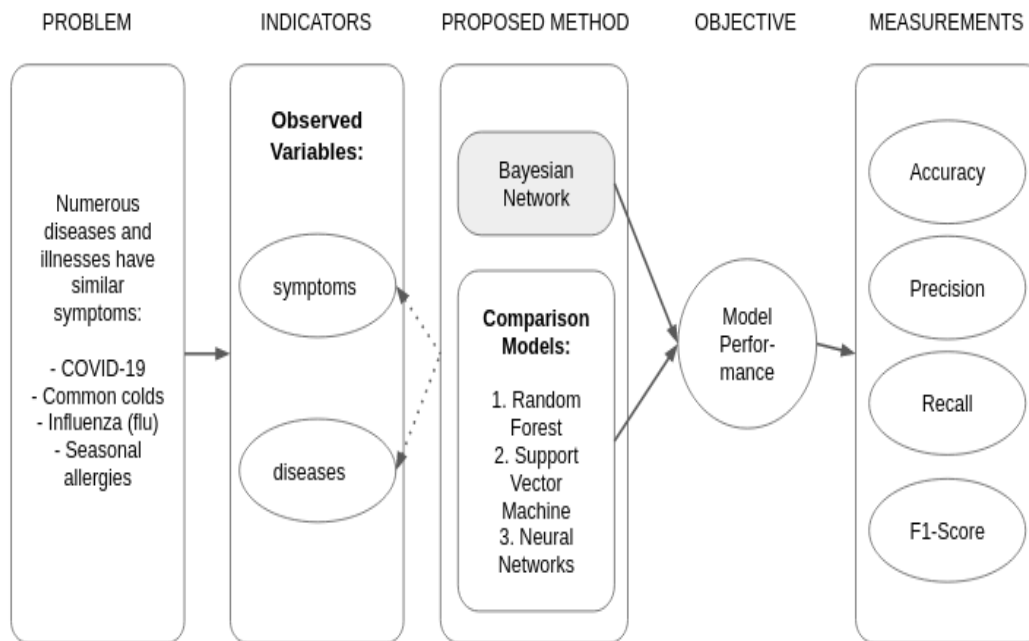


Figure 2. Research Framework.

3.2. Data Collection

The data obtained from Kaggle's website which is data on symptoms and types of diseases simulation results built on the theory from the article Mayo Clinic with 44.453 records of 21 variables, which provides a common symptom difference between COVID-19, common cold, flu, and seasonal allergies [15].

Variables consist of cough, muscle aches, fatigue, sore throat, runny nose, tiredness, fever, nausea, vomiting, diarrhea, shortness of breath, difficulty breathing, loss of taste, loss of smell, itchy nose, itchy eyes, itchy mouth, itchy ears, sneezing, pink eyes, and the types (categories of diseases) are sequentially defined as described in Table 2.

Table 2. List of variables

No.	Name of variable	Type of variable
1	Cough	Predictor variable
2	Muscle_Aches	Predictor variable
3	Fatigue	Predictor variable
4	Tiredness	Predictor variable
5	Sore_Throat	Predictor variable
6	Runny_Nose	Predictor variable
7	Stuffy_Nose	Predictor variable
8	Fever	Predictor variable
9	Nausea	Predictor variable
10	Vomiting	Predictor variable
11	Diarrhea	Predictor variable
12	Shortness_Of_Breath	Predictor variable



No.	Name of variable	Type of variable
13	Sneezing	Predictor variable
14	Difficulty_Breathing	Predictor variable
15	Loss_Of_Taste	Predictor variable
16	Loss_Of_Smell	Predictor variable
17	Itchy_Nose	Predictor variable
18	Itchy_Eyes	Predictor variable
19	Itchy_Mouth	Predictor variable
20	Itchy_Inner_Ear	Predictor variable
21	Type of disease/illness	Target variable

3.3. Application of Bayesian Network Method

In this study, we use 20 symptoms and the type of disease, each represented by variable. To model, the connectedness between these 21 variables, a Bayesian Network structure model was built from the data using Rstudio and Google Colaboratory software. The resulting model will be used to classify COVID-19 diseases, common colds, flu, and allergies. In general, the workflow of the model to be built consists of 6 stages as in Figure 2, namely preprocessing data, building the Structure of Bayesian Network with BIC score approach, visualization of Bayesian Network structure with output in the form of *Directed Acyclic Graph* (DAG), calculating learning parameters of Bayesian Network structure, and evaluation to get accuracy value from *confusion matrix* results for further use as a comparison of Bayesian Network model with other models based on accuracy obtained.

3.4. Comparison with other Classification Methods.

Existing machine learning classification methods have been investigated in the recent literatures to distinguish multiple symptoms of illnesses [16-20]. The classification method used in this study is a further evaluation in comparing the performance of models that have been built using the Bayesian Network, including Random Forest, Support Vector Machines (SVM), Neural Network.

4. Results and Discussion

In this study, there were 4 types of diseases, namely COVID-19, colds, flu, and allergies, and 20 symptoms of the disease. Due to related research, that they don't make causal relationship of the symptoms at all, each symptom only depends directly on the type of disease. Therefore, here is built a model using the Bayesian Network method against these symptoms so that this method can present a causal relationship between the symptoms that have not been defined in the related research. In addition, this method can also calculate the probability of symptoms of a disease suffered. On the Bayesian Network, all parameters are variables so that each parameter has a distribution value. The simulation data published on Kaggle is based on the theory in the Mayo Clinic article with many simulations of 44.453 rows data so that it can support to be applied to the Bayesian network method.

There are several steps in implementing this Bayesian Network method, the following are the steps and explanations of the methods mentioned.

4.1. Bayesian Network Model

We construct a Bayesian Network structure for the classification of COVID-19 diseases, colds, flu, and allergies. In this study, the Bayesian Network model was implemented using RStudio software. The model structure was constructed using the Hill-Climbing method which is a popular method for the formation of the Bayesian Network of data. By using a Bayesian Information Criterion (BIC) score to obtain a high-quality Bayesian Network structure with reasonable time. Iterations of the Scoring Based algorithm assume that each node is in a sorted state and starts from a structure that has no



connection/network at all, and the network creation process will stop if it meets the number of iterations. In that order, there is a rule that the nodes in the initial order are unlikely to be a child for the next sequence node. The formation of the Bayesian Network model uses a Scoring Based algorithm and the output is illustrated in Figure 3.

```

Bayesian network learned via Score-based methods

model:
[ITCHY_NOISE|TYPE|ITCHY_NOISE][COUGH|TYPE][MUSCLE_ACHES|TYPE]
[TIREDNESS|TYPE][RUNNY_NOSE|TYPE][STUFFY_NOSE|TYPE][FEVER|TYPE]
[DIARRHEA|TYPE][ITCHY_EYES|TYPE][ITCHY_MOUTH|TYPE]
[ITCHY_INNER_EAR|TYPE][SNEEZING|TYPE][PINK_EYE|TYPE]
[LOSS_OF_TASTE|COUGH:TYPE][SORE_THROAT|LOSS_OF_TASTE:TYPE]
[VOMITTING|LOSS_OF_TASTE:TYPE][LOSS_OF_SMELL|LOSS_OF_TASTE:TYPE]
[NAUSEA|LOSS_OF_SMELL:TYPE]
[SHORTNESS_OF_BREATH|LOSS_OF_SMELL:TYPE]
[DIFFICULTY_BREATHING|LOSS_OF_SMELL:TYPE]
nodes:                21
arcs:                  27
  undirected arcs:    0
  directed arcs:      27
average markov blanket size: 2.57
average neighbourhood size: 2.57
average branching factor: 1.29

learning algorithm:    Hill-Climbing
score:                 BIC (disc.)
penalization coefficient: 5.351094
tests used in the learning procedure: 750
optimized:             TRUE
    
```

Figure 3. Output result from Bayesian Network modeling

4.2. Bayesian Network Structure

It can be seen that by using a scoring-based algorithm with Hill climbing application and BIC score approach, 27 arcs formed from 21 variables, so that the Bayesian Network structure obtained is made into the plot visualization of the resulting DAG structure can be seen in Figure 4.

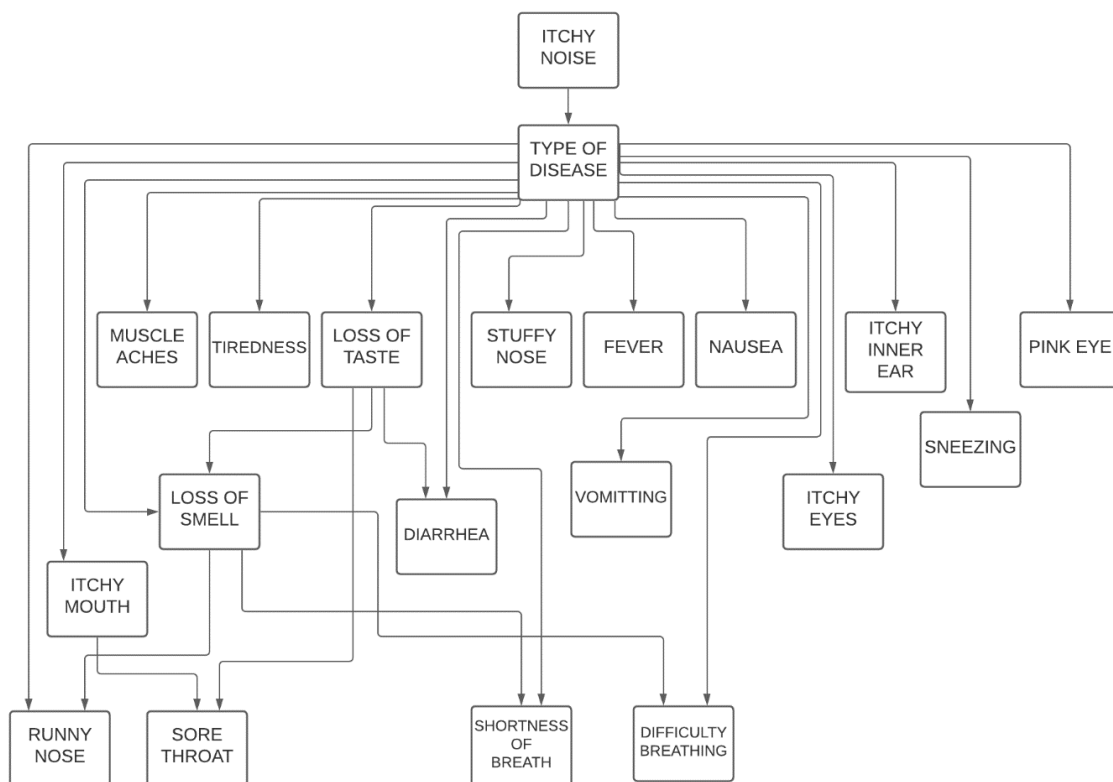


Figure 4. The Directed Acyclic Graph structure of our Bayesian Network model



4.3. Learning Parameters

Calculating the Learning Parameters of the DAG structure in Figure 4 as the interrelationship between data variable. After the Bayesian Network structure is formed, a value is generated for each parameter according to the probability. In the calculation of learning parameters itself using Bayesian Estimation, the calculation is done by using data to get the value of CPT (Conditional Probability Table).

For example, based on the structure in Figure 4, there can be identified an association between cough variables that indicate an association with LOSS_OF_TASTE (loss of taste) and ITCHY_MOUTH (itchy mouth) that is described in Table 3.

Table 3. Conditional Probability Table

Cough	Itchy Mouth			
	Yes		No	
	Loss of Taste	Loss of Taste	Loss of Taste	Loss of Taste
	Yes	No	Yes	No
Yes	0,455	0,493	0,500	0,502
No	0,545	0,507	0,500	0,498

4.4. Confusion Matrix

The result of the confusion *matrix* of the Bayesian Network model assisted by using package 'mltest' in RStudio obtained the following output results in Figure 5.

```
> CM2$accuracy
[1] 0.9313913
> CM2$F1
ALLERGY    COLD      COVID     FLU
0.9804099 0.6094808 0.5705458 0.9486231
> CM2$precision
ALLERGY    COLD      COVID     FLU
0.9751199 0.5601660 0.4902655 0.9719941
> CM2$recall
ALLERGY    COLD      COVID     FLU
0.9857576 0.6683168 0.6822660 0.9263496
```

Figure 5. Bayesian Network's Confusion Matrix

From the results of the Bayesian Network model that has been produced, then continued the evaluation by looking at the accuracy of the model with other classification methods namely Random Forest, SVM, and Neural Network. By using the help of a web application from Google Collaboratory obtained the output for the random forest method in Figure 6.



```
{'ALLERGY': 0, 'COLD': 1, 'COVID': 2, 'FLU': 3}
[[3241  81   2  14]
 [  82  73   6  50]
 [   3   5  42 360]
 [  22  67 343 4500]]
      precision    recall  f1-score   support

     0       0.97     0.97     0.97     3338
     1       0.32     0.35     0.33      211
     2       0.11     0.10     0.10      410
     3       0.91     0.91     0.91     4932

 accuracy          0.88     8891
 macro avg         0.58     8891
 weighted avg      0.88     8891
```

Figure 6. Random Forest’s Confusion Matrix

Performance result of the Support Vector Machine (SVM) and neural network methods are provided in Figures 7 and respectively in terms of the precision, recall, F1-score, and accuracy.

```
{'ALLERGY': 0, 'COLD': 1, 'COVID': 2, 'FLU': 3}
[[3317  18   3   0]
 [ 104  99   8   0]
 [   5   0 195 210]
 [  40  51 280 4561]]
      precision    recall  f1-score   support

     0       0.96     0.99     0.98     3338
     1       0.59     0.47     0.52      211
     2       0.40     0.48     0.44      410
     3       0.96     0.92     0.94     4932

 accuracy          0.92     8891
 macro avg         0.73     8891
 weighted avg      0.92     8891
```

Figure 7. SVM’s Confusion Matrix

```
---- Test data ----
Accuracy: 93.84
Classification Report:
      precision    recall  f1-score   support

     0       0.97     0.98     0.97     3338
     1       0.41     0.42     0.42      211
     2       0.42     0.70     0.53      410
     3       0.96     0.91     0.94     4932

 micro avg         0.91     0.91     0.91     8891
 macro avg         0.69     0.75     0.71     8891
 weighted avg      0.93     0.91     0.92     8891
 samples avg      0.91     0.91     0.91     8891
```

Figure 8. Neural Network’s Confusion Matrix



From the above results, we can see and compare the accuracy of the Bayesian Network method with the three classification methods compared further and compiled based on the accuracy value as follows.

- Bayesian Network : 93.14%
- Random Forest : 88%
- SVM : 92%
- Neural Network : 93.84%

From the experimental results, the highest accuracy value is achieved by using the neural network method while Random Forest obtained the lowest accuracy value. The proposed Bayesian Network model achieved better accuracy than RF and SVM while it is slightly worse than the Neural Network. Hence, in this case, the Bayesian Network method is quite good at classifying the respiratory diseases the directed acyclic graph structure of Bayesian Network model is quite representative in classifying the cases of COVID-19 among other similar counterparts.

5. Conclusion

From the Bayesian Network Modeling that has been proposed and constructed in this study, it can be concluded that the use of this model using scoring-based methods, with its application using hill-climbing algorithms is able to produce a fairly representative Bayesian Network. The obtained accuracy of 93.14% is quite good and better than other machine learning classification models of Random Forest and Support Vector Machine (SVM). In a sense, the method of the Bayesian Network is a promising approach in classifying the symptoms of diseases of the type of COVID-19 disease, common colds, flu, and seasonal allergies.

References

- [1] Daniel-Ioan, et.al. 2009. Bayesian Network Model for Diagnosis of Psychiatric Diseases. Proceedings of the International Conference on Information Technology Interfaces, ITI. 61 - 66. 10.1109/ITI.2009.5196055.
- [2] Seixas F, et. al. 2014. A Bayesian Network Decision Model for Supporting the diagnosis of Dementia, Alzheimer's disease and mild cognitive impairment. Computers in biology and medicine. 51C. 140-158. 10.1016/j.combiomed.2014.04.010.
- [3] Heckerman, D. 2008. A Tutorial on Learning with Bayesian Networks. In: Innovations in Bayesian Networks. SCI, vol. 156, pp. 33–82. Springer, Heidelberg.
- [4] Putra, A. E. 2018. Implementasi Metode Bayesian Network Untuk Diagnosis Penyakit Kambing (Studi Kasus: UPTD Pembibitan Ternak Dan Hijauan Mankanan Ternak Singosari Malang). Malang: Fakultas Ilmu Komputer, Universitas Brawijaya.
- [5] Judea Pearl. 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- [6] Syukron & Subekti. 2018. Penerapan Metode Random Over-Under Sampling dan Random Forest untuk Klasifikasi Penilaian Kredit. Jurnal Informatika Vol.5 No.2
- [7] Gorunescu, F. 2011. Data Mining: Concept, Models, and Techniques. Verlag Berlin Heidelberg: Springer.
- [8] Vapnik, V., & Cortes, C. 1995. Support-Vector Networks. Boston: Kluwer Academic Publishers. 20, pp. 273-297.
- [9] Liao, W. T. & Triantaphyllou, E., 2007. Recent Advances in Data Mining of Enterprise Data: Algorithms and Applications, Series on Computers and Operations Research. Vol 6. Singapore: World Scientific Publishing Co. Pte. Ltd.
- [10] Prasetyo, E., 2014. Data Mining Mengolah Data Menjadi Infomasi Menggunakan Matlab. Yogyakarta: Andi.
- [11] Nur, A. I. T., Setiawan, N. Y., Bachtiar, F. A. 2019. Perbandingan Performa Metode Klasifikasi Svm, Neural Network, Dan Naïve Bayes Untuk Mendeteksi Kualitas Pengajuan Kredit Di Koperasi Simpan Pinjam. JTIK, Vol. 6 No. 4.



- [12] Raras, B. 2014. Prediksi Data Menggunakan Neural Network dengan Algorithma Backpropagation. [online] Tersedia di: https://www.academia.edu/7382142/Prediksi_Data_Menggunakan_Neural_Network_dengan_Algorithma_Backpropagation. Accessed on June 9, 2021
- [13] D. C. Manning, P. Raghavan and H. Schutze. 2008. Introduction to Information Retrieval, Cambridge: Cambridge University Press.
- [14] Pradasari, N., & Atimi, R. 2019. Pemodelan Bayesian Network untuk Prediksi Penyakit Saluran Pernapasan. Petir: Jurnal Pengkajian Dan Penerapan Teknik Informatika, 12(2), 292 - 302. <https://doi.org/10.33322/petir.v12i2.637>
- [15] Conway, W. Covid, Flu, Cold Symptoms Dataset. Kaggle. <https://www.kaggle.com/walterconway/covid-flu-cold-symptoms>. Accessed on June 9, 2021.
- [16] P. Dutta, S. Paul, and A. Kumar. 2021. Chapter 25 - Comparative analysis of various supervised machine learning techniques for diagnosis of COVID-19. Electronic Devices, Circuits, and Systems for Biomedical Applications. 521-540
- [17] M. Pahar, M. Klopper, R. Warren, and T. Niesler. 2021. COVID-19 cough classification using machine learning and global smartphone recordings. Computers in Biology and Medicine. 135 (104572).
- [18] D. M. Ibrahim, N. M. Elshennawy, and A. M. Sarhan. 2021. Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases. Computers in Biology and Medicine. 132 (104348).
- [19] T. Tuncer, F. Ozyurt, S. Dogan, and A. Subasi. 2021. A novel Covid-19 and pneumonia classification method based on F-transform. Chemometrics and Intelligent Laboratory Systems. 210 (104256).
- [20] Y. Shen, L. Zhang, J. Zhang, M. Yang, B. Tang, Y. Li, and K. Lei. 2018. CBN: Constructing a clinical Bayesian network based on data from the electronic medical record. Journal of Biomedical Informatics. 88. 1-10.
- [21] G. Polancik. 2017. Empirical Research Method Poster. Maribor, Slovenia.
- [22] C. W. Dawson. 2009. Project in Computing and Information System a Student Guide 2nd Edition. Addison-Wesley.