

Penerapan Metode *Radial Basis Function* (RBF) dalam Mengklasifikasikan Penyakit Demam Berdarah

Steffany ^{1*}, Akhmad Rezki Purnajaya ², Rika Jelita ³, Eva Tesvara ⁴, Milson Nestelrody ⁵, Joni Irwansyah ⁶

^{1,3,4,5,6}Sistem Informasi, Universitas Universal

²Teknik Perangkat Lunak, Universitas Universal

*Corresponding author E-mail: steffanyfany0@gmail.com

Article Info

Article history:

Received 30-05-2023

Revised 31-05-2023

Accepted 31-05-2023

Keyword:

Dengue Fever, Classification, Disease Prediction, Radial Basis Function

ABSTRACT

Dengue fever is an arbovirus disease caused by dengue virus infection through the bite of the *Aedes mosquito*, *Aedes aegypti*. Signs or symptoms of dengue usually resemble ordinary viral infections, but can become more severe and cause other symptoms that can paralyze the activities of the patient who suffers from it. The initial symptoms of someone who is infected with the dengue virus will experience a fever within 4–6 days after being infected. Therefore, early diagnosis is needed whether a person is infected or not with dengue fever. Because if someone is late getting treatment by medical personnel, then this can lead to death. Prevention can be done by making the right predictions. Dengue fever disease prediction research is needed to assist medical personnel in making a diagnosis. To get a good level of accuracy, research collected data and used the Radial Basis Function (RBF) approach method. The results of the RBF prediction evaluation obtained are that the RBF model gets 100% accuracy, 100% sensitivity, 100% specificity, and 100% AUC value.

Copyright © 2023 Journal of Digital Ecosystem for Natural Sustainability.
All rights reserved.

I. PENDAHULUAN

Penyakit Demam Berdarah (DBD), biasanya banyak dijumpai terutama di daerah tropis dengan beberapa faktor yang mempengaruhi munculnya DBD yaitu rendahnya status kekebalan masyarakat dan kepadatan populasi nyamuk *Aedes aegypti* karena banyak tempat perindukan nyamuk untuk berkembang yang biasanya terjadi pada musim hujan [1].

World Health Organization (WHO) memperkirakan setidaknya dalam satu tahun, 50 miliar penduduk dunia yang terinfeksi *dengue virus*. Saat ini infeksi *dengue virus* menjadi endemis dilebih dari 100 negara, antara lain Afrika, Amerika, Mediterania Timur, Pasifik Barat, dan Asia Tenggara. Di Asia Tenggara salah satunya adalah Indonesia. Kasus DBD di Indonesia mengalami peningkatan dalam kurun waktu 5 tahun terakhir [2]. Diperkirakan setiap tahun ada 3.000.000 kasus di Indonesia dan 500.000 kasus DBD memerlukan perawatan di rumah sakit dan minimal 12.000 di antaranya meninggal dunia, terutama anak-anak [1].

Pada penelitian ini digunakan metode *Radial Basis Function* (RBF) untuk melakukan klasifikasi dalam mendiagnosa DBD. Metode RBF merupakan struktur jaringan sederhana yang tidak perlu menggunakan perhitungan panjang dan memiliki kemampuan untuk mempelajari sesuatu dengan cepat. Dalam pendekatan klasik metode RBF memiliki *hidden layer* didapat dari input data. Struktur jaringan RBF *Neural Network* terdiri dari tiga lapisan yaitu *input layer*, *hidden layer* dan *output layer* [3]. RBF telah digunakan sebagai salah satu metode dalam menyelesaikan masalah yaitu “Identifikasi Pola Penyakit pada Citra Iris Mata” dengan akurasi sebesar 53,33% [4]. Dan juga menyelesaikan masalah yaitu “Model Prediksi penyakit Ginjal Kronik” dengan akurasi sebesar 93,75% [5].

Dilakukannya penelitian ini, berupaya untuk melakukan pencegahan dan penanganan kasus DBD agar tidak mengalami peningkatan lagi. Penelitian tersebut, menggunakan RBF sebagai metode pengujian untuk mendapatkan akurasi yang lebih baik. Uji ini dilakukan berdasarkan jumlah trombosit, hemoglobin, hematokrit dan leukosit dalam tubuh pasien. Sehingga berdasarkan hasil

klasifikasi tersebut dapat diketahui apa saja penyebabnya dan bagaimana penanganannya [5]. Kesimpulan dari tujuan program kami dibuat yakni, untuk meneliti, memprediksi dan mengetahui pola gejala *dengue* atau DBD dengan menggunakan metode RBF.

II. METODE

Penelitian ini menggunakan metode eksperimen. Dengan tujuan untuk meningkatkan akurasi dalam memprediksi DBD menggunakan RBF untuk menganalisis data pasien yang di diagnosa memiliki penyakit DBD.

2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang didapatkan dari <https://www.kaggle.com/discussions/general/91461>. Data tersebut di *publish* oleh Ashish Chawla. Datasetnya tersusun atas 18 nama variabel dan 108 data pasien [6].

2.2 Data Preprocessing

Pada tahap ini, data yang masih mentah akan dibersihkan dengan mengisi nilai yang kosong, menghilangkan data *noisy* dan menyelesaikan inkonsistensi data yang ditemukan. Kemudian ada 3 variabel yang dihapus yaitu, *dengue.p_i_d*, *dengue.date_of_fever*, dan juga *dengue.residence* karena 3 variabel tersebut tidak mempengaruhi terhadap variabel kelas penyakit DBD. Variabel independen pada data ini ada 17 variabel dengan keterangan ditunjukkan seperti Tabel 1 di bawah:

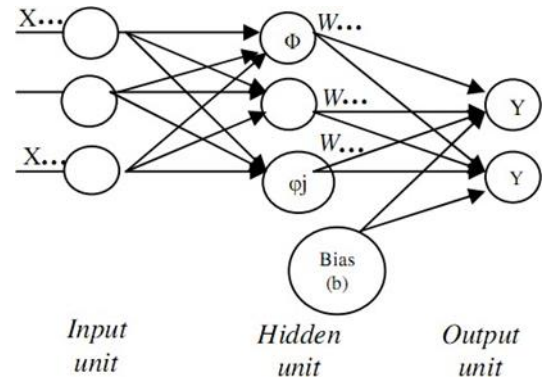
Table 1. Keterangan variabel independen dataset

NAMA VARIABEL	KETERANGAN
<i>dengue.p_i_d</i>	ID pasien
<i>dengue.date_of_fever</i>	tanggal pasien terdeteksi
<i>dengue.residence</i>	kota pasien
<i>dengue.days</i>	berapa hari pasien terdeteksi
<i>dengue.current_temp</i>	suhu tubuh pasien saat ini dalam fahrenheit
<i>dengue.wbc</i>	sel darah putih
<i>dengue.severe_headche</i>	sakit kepala parah
<i>dengue.pain_behind_the_eyes</i>	sakit pada belakang mata
<i>dengue.joint_muscle_aches</i>	nyeri otot sendi
<i>dengue.metallic_taste_in_the_mouth</i>	rasa logam dalam mulut pasien
<i>dengue.appetite_loss</i>	hilangnya nafsu makan
<i>dengue.addominal_pain</i>	sakit perut
<i>dengue.nausea_vomiting</i>	mual muntah
<i>dengue.diarrhoea</i>	diare
<i>dengue.hemoglobin</i>	protein pada sel darah merah pasien
<i>dengue._hematocri</i>	persentase sel darah merah dalam tubuh pasien
<i>dengue.platelet</i>	trombosit

2.3 Metode RBF

Metode analisis yang digunakan yaitu *Radial Basis Function* (RBF). Metode ini dapat digunakan untuk

menangani beberapa kasus, terutama untuk pengklasifikasian data. RBF juga memiliki topologi jaringan, terdiri dari unit lapisan masukan (x), unit lapisan tersembunyi (ϕ), dan unit lapisan keluaran (y). Berikut di bawah ini adalah gambar topologi jaringan RBF:



Gambar 1. Arsitektur jaringan RBF

Penjelasan mengenai struktur jaringan RBF, yakni:

a. Input Layer

Input layer merupakan masukan data dari jaringan RBF yang terletak di lapisan pertama. *Input layer* tidak hanya digunakan sebagai data input saja, namun berfungsi juga untuk menentukan centroid dan standar deviasi dari data yang berfungsi membangkitkan fungsi aktivasi pada *hidden layer*. Terdapat dua cara yang digunakan dalam penentuan *centroid* yaitu dengan metode acak dan menggunakan algoritma *k-means clustering*. Setelah proses pencarian nilai centroid dari data input dilakukan, tahap selanjutnya dilakukan perhitungan nilai Gaussian dengan menggunakan persamaan Gaussian. Nilai *centroid* hasil pencarian dari data input akan digunakan pada *hidden layer* [5].

b. Hidden Layer

Hidden Layer merupakan lapisan yang digunakan fungsi aktivasi untuk memperoleh bobot jaringan yang digunakan pada *output layer*. Penerapan fungsi yang umum digunakan dalam pemecahan kasus dalam jaringan RBF yaitu dengan menggunakan fungsi basis Gaussian yang diformulasikan sebagai berikut:

$$\phi(r) = \exp\left(-\frac{\|x_i - c_i\|^2}{2\sigma_i^2}\right) \quad (1)$$

dengan :

$$\sigma = \frac{\text{jarak maksimum antara 2 pusat}}{\sqrt{\text{banyak pusat}}} = \frac{d_{max}}{\sqrt{n}} \quad (2)$$

Setelah nilai Gaussian diketahui, maka proses selanjutnya yakni menghitung nilai bobot baru (w) dengan mengalikan *pseudoinverse* dari matriks G (Gaussian) dengan vector target (d) dengan persamaan berikut:

$$w = (G^T G)^{-1} G^T d \quad (3)$$

setelah nilai weight diketahui maka tahap selanjutnya adalah menyimpan nilai hasil dari perhitungan centroid dan nilai weight agar dapat digunakan Kembali pada saat proses pengujian pada jaringan RBF [5].

c. Output Layer

Output Layer merupakan langkah terakhir setelah memperoleh hasil perhitungan dari fungsi basis Gaussian untuk melakukan pembaharuan bobot baru. *Output layer* merespon dari jaringan sesuai dengan pola yang di masukkan sebelumnya pada *input layer*. Transformasi dari ruang *input layer* ke ruang *hidden layer* adalah non-linier, sedangkan transformasi dari ruang *hidden layer* ke *output layer* adalah linier [5]. Pada tahap ini akan menghitung nilai output jaringan $y(n)$ ditambah bobot bias (b) dengan persamaan berikut:

$$y(n) = \sum_{i=1} w_i G(\|X - C\| + b) \quad (4)$$

2.4 Listing Model RBF

```

Program Jurnal
install.packages('MASS')
library(MASS)

install.packages('e1071')
library(e1071)

data = read.csv("C:/tugas kelompok dengue
data.csv", header=TRUE)
str(data)
head(data)
data<-data[,4:18]
dim(data)

table(data$dengue.dengue)

set.seed(123)
train<-sample(1:108,54)
test<-data[-train,]

modelrbf<-
svm(dengue.dengue~.,data,kernel='radial',subset=
train)

rbfpred<-predict(modelrbf,test)
rbfpred[rbfpred<0.5] <- 0
rbfpred[rbfpred>=0.5] <- 1

rbftable<-
table(prediksi=rbfpred,aktual=test$dengue.dengue
)
rbftable

(akurasirbf<-sum(diag(rbftable))/sum(rbftable))
(spesifisitas<-
rbftable[1,1]/(rbftable[1,1]+rbftable[2,1]))
(sensitifitas<-
rbftable[2,2]/(rbftable[2,2]+rbftable[1,2]))
(auc<-(spesifisitas+sensitifitas)/2)

paste("akurasi:",akurasirbf*100,"%",sep=" ")
paste("spesifisitas:",spesifisitas*100,"%",sep="
")
paste("sensitifitas:",sensitifitas*100,"%",sep="
")
paste("auc:",auc*100,"%",sep=" ")

```

2.5 Evaluasi Model RBF

Confusion matrix adalah pengukuran performa untuk masalah klasifikasi *machine learning* dimana keluaran dapat berupa dua kelas atau lebih. *Confusion matrix* adalah tabel dengan empat kombinasi berbeda dari nilai prediksi dan nilai actual [7]. Untuk menghitung *Confusion matrix* dari prediksi dan actual seperti yang diberikan, kita dapat menggunakan formula sebagai berikut:

- *True Positive (TP)*: jumlah data yang diprediksi benar positif (1) oleh model dan kenyataannya juga positif (1).
- *False Positive (FP)*: jumlah data yang diprediksi positif (1) oleh model tetapi kenyataannya negatif (0).
- *True Negative (TN)*: jumlah data yang diprediksi benar negatif (0) oleh model dan kenyataannya juga negatif (0).
- *False Negative (FN)*: jumlah data yang diprediksi negatif (0) oleh model tetapi kenyataannya positif (1).

Parameter dalam pengukuran klasifikasi yang dapat diukur berdasarkan *Confusion matrix* adalah akurasi, sensitivitas dan spesifisitas. Akurasi yakni perbandingan antara jumlah diagonal utama *Confusion matrix* dengan jumlah keseluruhan objek. Apabila nilai akurasi semakin mendekati 100% maka menunjukkan performa klasifikator semakin baik. Sensitivitas atau *recall* yakni ukuran nilai kecepatan dari suatu kejadian yang diinginkan (benar positif). Dan spesifisitas yakni persentase dari suatu kejadian yang tidak diinginkan (benar negatif) [7]. Sehingga dapat ditulis formulasi perhitungan akurasi, sensitivitas dan spesifisitas berdasarkan *confusion matrix* dalam persamaan berikut:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (5)$$

$$\text{Sensitifitas} = \frac{TP}{TP+FN} \times 100\% \quad (6)$$

$$\text{Spesifisitas} = \frac{TN}{FP+TN} \times 100\% \quad (7)$$

$$\text{AUC} = \frac{\text{Sensitifitas} + \text{Spesifisitas}}{2} \quad (8)$$

III. HASIL DAN PEMBAHASAN

3.1 Split Data

3.1.1 Train

`train <- sample(1:108,54)` variabel "train" yang berisi 54 angka acak antara 1 sampai dengan 108. Angka-angka ini merepresentasikan baris-baris data yang akan digunakan sebagai set data *train*.

3.1.2 Test

`test <- data[-train,]` membuat variabel baru bernama "test", yang berisi sisa data yang tidak termasuk dalam set data "train" (54 baris data variabel "train"), yang akan digunakan sebagai set data *test*.

3.2 Train Model RBF

Dalam kasus penelitian ini, kami menggunakan model RBF. Model prediksi RBF memang salah satu metode yang populer digunakan dalam *Machine Learning* untuk masalah klasifikasi. Pengujian dilakukan menggunakan bahasa pemrograman R beserta beberapa *library* seperti *library(MASS)* dan *library(e1071)*.

3.3 Evaluasi Prediksi Model RBF

Dengan menggunakan data yang diberikan, maka didapat *confusion matrix* sebagai berikut:

```
rbf$table <- table(prediksi=rbf$pred, aktual=test$dengue.dengue)
rbf$table
      aktual
prediksi 0 1
0         14 0
1         0 40
```

Gambar 2. Hasil *Confusion matrix* model RBF

Dalam *Confusion Matrix* tersebut, nilai 14 pada posisi (1,1) menunjukkan bahwa ada 14 data yang diprediksi benar negatif (0) oleh model dan kenyataannya juga negatif (0). Nilai 40 pada posisi (2,2) menunjukkan bahwa ada 40 data yang diprediksi benar positif (1) oleh model dan kenyataannya juga positif (1). Sedangkan untuk *false positive* dan *false negative* semuanya bernilai 0 karena tidak ada data yang salah diklasifikasikan dalam kasus ini. Selanjutnya, melakukan pengujian yakni melihat nilai persentase akurasi, spesifisitas, sensitifitas dan AUC, dapat dilihat sebagai berikut:

$$\text{Akurasi} = \frac{14 + 40}{14 + 40 + 0 + 0} \times 100\% = 100\%$$

$$\text{Sensitifitas} = \frac{14}{14 + 0} \times 100\% = 100\%$$

$$\text{Spesifisitas} = \frac{40}{0 + 40} \times 100\% = 100\%$$

$$\text{AUC} = \frac{100 + 100}{2} = 100\%$$

Sehingga dari perhitungan tersebut dapat dilihat bahwa terbukti nilai akurasi dan AUC dari model prediksi diagnosa penyakit DBD dengan metode RBF tercapai pada angka 100%.

IV. KESIMPULAN

Strategi deteksi dini wabah *dengue virus* berdampak besar terhadap kesejahteraan masyarakat. Prediksi yang disajikan dengan menggunakan model RBF dalam penelitian ini sangat menjanjikan karena mendapatkan nilai Akurasi 100%, Sensitifitas 100%, Spesifitas 100% dan AUC 100%. Model deteksi dini yang mempertimbangkan langkah-langkah kerentanan pra-wabah, dikombinasikan dengan variabilitas iklim, memberikan penjelasan yang masuk akal untuk epidemi demam berdarah dan dapat berfungsi sebagai alat prediksi yang layak untuk wabah di masa mendatang. Rencana masa depan harus dilaksanakan yang berfokus pada peningkatan pengujian Pemodelan RBF Dengan lebih baik. di daerah non-tropis mungkin lebih mengandalkan model RBF yang disajikan, karena kondisi iklimnya cenderung tidak optimal untuk nyamuk. Fluktuasi kecil dalam iklim dapat meningkatkan perkembangbiakan nyamuk. Seiring dengan pemanasan global, semua negara mungkin perlu segera mengalihkan fokus mereka ke model RBF.

V. DAFTAR PUSTAKA

- [1] F. R. Azinar, "Model Buku Saku Dan Rapor Pemantauan Jentik Dalam Meningkatkan Perilaku Pemberantasan Sarang Nyamuk", *Journal of Health Education*, pp. 110-117, 2018.
- [2] I. D. Manuaba, I. W. P. Sutirtayasa, D. R. Dewi, "Immunopatogenesis Infeksi Virus Dengue", *E-Jurnal Medika Udayana*, pp. 1-11, 2013
- [3] Y. F. Waruwu, M. Zarlis, E. B. Nababan, M. S. Ziliwu, "Seleksi Atribut Pada Algoritma Radial Basis Function Neural Network Menggunakan Information Gain", *Seminar Nasional Royal (SENAR) 2018*, pp. 21-24, 2018.
- [4] D. H. Yohandy, I. M. N. Wiranata, T. Q. Ferbia, "Identifikasi Pola Penyakit Pada Citra Iris Mata Dengan RBF Neural Network", *Jurnal Informatika*, pp. 195-201, 2018.
- [5] A. W. S. Stefanus, (2016), "Model Prediksi Penyakit Ginjal Kronik Menggunakan Radial Basis Function", *Jurnal Pseudocode*, pp. 163-170, 2016.
- [6] C. Ashish, "Typhoid and Dengue Fever Symptoms Dataset", 2021, Retrieved from Kaggle: <https://www.kaggle.com/discussions/general/91461#1188270>.
- [7] W. Vira, "Penerapan Radial Basis Function Neural Network Dalam Pengklasifikasian Daerah Tertinggal Di Indonesia", *Jurnal Aplikasi Statistika & Komputasi Statistik*, pp. 37-54, 2020.