

# Pemanfaatan Informasi Iklim Terhadap Kasus Demam Berdarah *Dengue* Di Kota Bandung Menggunakan Algoritma *Partial Least Square*

1<sup>st</sup> M. Fikri Andika Kurniawan

Fakultas Teknik Elektro  
Universitas Telkom  
Bandung, Indonesia

coldzaera@student.telkomuniversity.ac.id

2<sup>nd</sup> Meta Kallista

Fakultas Teknik Elektro  
Universitas Telkom  
Bandung, Indonesia

metakallista@telkomuniversity.ac.id

3<sup>rd</sup> Ashri Dinimaharawati

Fakultas Teknik Elektro  
Universitas Telkom  
Bandung, Indonesia

ashridini@telkomuniversity.ac.id

**Abstrak** — Demam berdarah Dengue (DBD) adalah salah satu penyakit endemik di seluruh wilayah tropis dan sebagian wilayah subtropics. Penyakit ini ditularkan oleh nyamuk *Aedes aegypti*. Penyakit dapat menyebarkan infeksi dalam suatu wilayah dengan cepat. Pada tahun 2016, terdapat ratusan ribu penderita DBD di seluruh wilayah Indonesia akibat penularan DBD yang disebabkan oleh nyamuk *Aedes aegypti*. Perubahan iklim bisa berpengaruh terhadap pola penyakit infeksi akan mempertinggi risiko penularan. Penelitian ini dilakukan untuk menemukan pengaruh dari iklim dengan tingkat kasus DBD di daerah Kabupaten Bandung menggunakan algoritma *Partial Least Square* (PLS) sebagai metode untuk mendapatkan korelasi antara variabel independent dengan variabel dependen. Sedangkan, algoritma *Support Vector Machine* (SVM) digunakan untuk melakukan prediksi terhadap kasus DBD. Hasil korelasi menggunakan PLS menunjukkan bahwa kelembapan memiliki pengaruh lebih tinggi terhadap jumlah kasus DBD. nilai *cross-loadings* yang didapatkan adalah 0.315865 dengan nilai akurasi koefisien determinasi ( $R^2$ ) adalah 0.099771. pada algoritma SVR, kombinasi antara partisi data, parameter dan kernel yang paling baik adalah partisi data yang memiliki perbandingan 80:20 antara *data training* dan *data testing* dengan nilai parameter masing-masing  $C = 10$ ,  $\gamma = 1$ , dan  $\epsilon = 0.05$ . Kernel yang digunakan adalah kernel RBF. Hasil regresi dari data kasus DBD memiliki nilai *error* dan koefisien determinasi masing-masing RMSE = 0.12314 dan  $R^2 = 0.58873$ .

**Kata Kunci**— Demam Berdarah Dengue, *Partial Least Square*, Python, *Support Vector Machine*, Website

## I. PENDAHULUAN

Demam berdarah Dengue (DBD) adalah salah satu penyakit endemik di seluruh wilayah tropis dan sebagian wilayah subtropics. Penyakit ini ditularkan oleh nyamuk *Aedes aegypti*. Pada tahun 2016, terdapat ratusan ribu penderita DBD di seluruh wilayah Indonesia dan ribuan penderita diantaranya meninggal dunia akibat penularan DBD yang disebabkan oleh nyamuk *Aedes aegypti*. Pada beberapa provinsi, kasus DBD terjadi peningkatan dan bersifat fluktuatif namun tetap pada jumlah yang cukup tinggi [1].

Perubahan iklim bisa berpengaruh terhadap pola penyakit infeksi akan mempertinggi risiko penularan. Hal yang berpengaruh adalah suhu, kelembapan, dan curah hujan. Suhu ideal untuk transmisi DBD adalah 21.6–32.9 °C dengan kelembapan berkisar 79%. Perubahan iklim pula mengakibatkan beberapa virus diperkirakan mengalami

peningkatan dalam peralihan musim yang ditandai oleh curah hujan dan suhu udara yang tinggi [2].

Sebuah metode yang membuat sebuah model untuk melakukan perhitungan *Incident Rate* dapat membantu pengambilan tindakan pencegahan terhadap kasus penyakit DBD. Penelitian dilakukan untuk menemukan pengaruh iklim dengan tingkat kasus DBD pada sebuah daerah. Metode yang digunakan adalah *Partial Least Square* (PLS) dan *Support Vector Machine* (SVM). Kedua metode ini digunakan untuk mendapatkan korelasi antara variabel iklim dengan kasus DBD dan melakukan regresi terhadap kasus DBD. Sedangkan data yang digunakan adalah data DBD dan data iklim harian di Kota Bandung dalam 10 tahun terakhir. Pada penelitian ini, permasalahan yang dibahas adalah bagaimana pengimplementasian metode PLS dan SVR dalam menemukan korelasi data dan meregresikan data kasus DBD dan data iklim. Hasil dari penelitian ini bertujuan untuk dapat melihat pengaruh dari iklim terhadap tingkat kasus DBD di kota Bandung

## II. KAJIAN TEORI

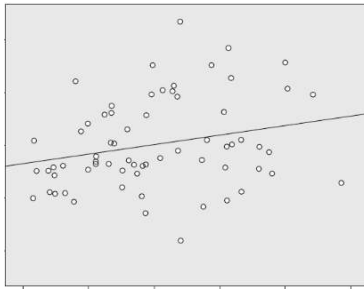
### A. Demam Berdarah Dengue

Demam berdarah dengue (DBD) adalah suatu penyakit yang disebabkan oleh infeksi virus dengue. DBD adalah penyakit akut dengan manifestasi klinis perdarahan yang menimbulkan syok yang berujung kematian. DBD disebabkan oleh salah satu dari empat serotipe virus dari genus *Flavivirus*, famili *Flaviviridae*. Setiap serotipe cukup berbeda sehingga tidak ada proteksi silang dan wabah yang disebabkan empat serotipe dapat terjadi. Virus ini bisa masuk ke dalam tubuh manusia dengan perantara nyamuk *Aedes aegypti* dan *Aedes albopictus* [4]. *Aedes aegypti* dan *Aedes albopictus* hidup pada habitat yang berbeda. *Aedes aegypti* memiliki habitat yang dekat dengan manusia, sedangkan *Aedes albopictus* memiliki habitat di kebun, rawa, dan hutan [3].

### B. Regresi

Regresi adalah metode analisis statistik untuk memperkirakan hubungan antara variabel dependen dengan satu atau lebih variabel independent. Analisis regresi dapat memahami variasi pada variabel dependen menggunakan variasi variabel independent dengan variabel pengganggu lain

yang dikendalikan. Analisis regresi banyak digunakan dalam melakukan prediksi dan estimasi kondisi variabel dependen dan variabel independen [4]. Hubungan antara variabel dependen dengan independen dapat dideskripsikan sebagai berikut.



GAMBAR  
ILUSTRASI DARI REGRESI.

$$y = a + bx + \varepsilon$$

Keterangan:

Y = Variabel dependen

a = Konstanta

b = Koefisien variabel x

$\varepsilon$  = kesalahan

Pengukur dari kinerja regresi bisa dilihat dari nilai *error* dan koefisien determinasi. Beberapa perhitungan dapat digunakan untuk menilai kesalahan dan akurasi dari permodelan regresi, di antaranya *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan koefisien determinasi ( $R^2$ ).

#### 1. Mean Squared Error

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Keterangan:

$Y_i$  = Nilai Asli

$\hat{Y}_i$  = Nilai prediksi

n = Jumlah data

#### 2. Root Mean Squared Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}$$

Keterangan:

$Y_i$  = Nilai Asli

$\hat{Y}_i$  = Nilai prediksi

n = Jumlah data

#### 3. Koefisien Determinasi

$$R^2 = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y}_i)^2}$$

Keterangan:

$Y_i$  = Observasi respon ke-i

$\hat{Y}_i$  = Ramalan respon ke-i

$\bar{Y}_i$  = Rata-rata

#### C. Partial Least Square

*Partial Least Square* (PLS) adalah alternative untuk regresi *Ordinary Least Square* (OLS), Korelasi Kanonik, atau *Covariance-Based Structural Equation Modeling* (SEM) dari variabel independen dan respons. PLS menghubungkan variabel independen dan beberapa variabel dependen. PLS juga dapat banyak variabel independen, bahkan pada saat *predictor* menunjukkan multikolinearitas. PLS biasanya

diimplementasi sebagai model regresi untuk memprediksi satu atau banyak variabel dependen dari satu set variabel independen. Selain itu PLS juga dapat diimplementasikan sebagai model jalur yang menghubungkan *predictor* dan jalur penghubung antara *predictor* dengan variabel dependen [5].

*Partial Least Square* (PLS) memodelkan hubungan variabel Y dengan variabel X berdasarkan variabel internal. Variabel X dibagi ke dalam skor  $t_h$  dan *loading*  $P_h$ , dinyatakan sebagai berikut:

$$X = t_1 p_1 + t_2 p_2 + t_3 p_3 + \dots + t_h p_h + E_h$$

Keterangan:

X = Variabel Bebas

$t_h$  = Vektor skor (*score vector*) variabel X

$P_h$  = Vektor muatan (*loading vector*) variabel X

$E_h$  = Matriks sisaan variabel X

Variabel Y juga dibagi dalam skor  $u_h$  dan *loading*  $q_h$  yang dinyatakan sebagai berikut:

$$Y = u_1 q_1 + u_2 q_2 + u_3 q_3 + \dots + u_h q_h + F_h$$

Keterangan:

Y = Variabel Bebas

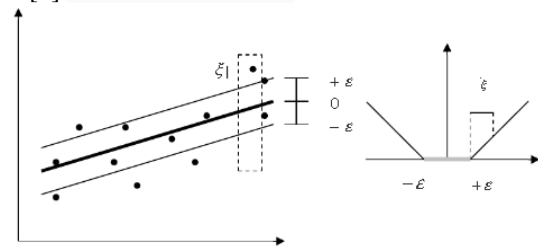
$u_h$  = Vektor skor (*score vector*) variabel X

$q_h$  = Vektor muatan (*loading vector*) variabel X

$F_h$  = Matriks sisaan variabel X

#### D. Support Vector Regression

*Support Vector Regression* adalah bentuk pendekatan dari SVM. SVR merupakan metode yang digunakan untuk mengatasi *overfitting*, yaitu mengatasi kondisi saat model dibangun dengan memperhitungkan semua ciri-ciri, termasuk *noise* [6].



GAMBAR

ILUSTRASI DARI FUNGSI SUPPORT VECTOR MACHINE.

Karakteristik SVR adalah mengecilkan batas kesalahan yang digeneralisasi. Batas kesalahan generalisasi adalah kombinasi dari kesalahan *training* dan mengontrol kompleksitas hipotesis [12]. Berikut ini adalah fungsi dari algoritma SVR.

$$f(x) = \omega \cdot x + b$$

Keterangan:

$\omega$  = Vektor bobot

x = fungsi yang memetakan x dalam suatu dimensi

b = bias

Performansi dari algoritma SVR ditentukan oleh beberapa parameter, yaitu parameter C, parameter epsilon ( $\epsilon$ ), dan gamma ( $\gamma$ ). Selain parameter, kernel yang digunakan juga mempengaruhi performansi algoritma SVR. Kernel yang digunakan antaralain, kernel Linear, Kernel Polinomial, Kernel RBF, dan Kernel Sigmoid.

##### 1. Kernel Linear

Kernel Linear adalah kernel yang paling sederhana pada algoritma SVR. Berikut adalah persamaan kernel linear [7].

$$y(x, w) = w^T \phi(x)$$

Keterangan:

$x$  = Variabel *input*

$w$  = Parameter

$\phi(x)$  = Fungsi Basis

## 2. Kernel Polinomial

Kernel polinomial melihat fitur sampel input yang digunakan untuk penentuan kesamaan dan kombinasinya. Dalam analisa regresi, kombinasi juga disebut sebagai fitur interaksi. Berikut adalah persamaan kernel polinomial [8].

$$K(x, x_i) = (\gamma x^T x_i + r)^d$$

Keterangan:

$\gamma, r, d$  = Parameter kernel

$x$  = Variabel *input*

$x_i$  = fungsi basis

## 3. Kernel RBF

Kernel *Radial Basis Function* (RBF) merupakan kernel yang paling umum digunakan. RBF bisa diaplikasikan ke dalam kasus non-linear pemetaan data *training* menjadi berhingga pada dimensi ruang. Berikut adalah bentuk persamaan kernel RBF [9].

Keterangan:

$x$  = Variabel *input*

$x_i$  = fungsi basis

$\sigma$  = *spread*

## 4. Kernel Sigmoid

Kernel sigmoid berasal dari algoritma *Neural Network*. Fungsi sigmoid sering digunakan sebagai aktivasi untuk neuron buatan [10].

$$K(x_i, x) = \tanh(\gamma x_i^T x + r)^d$$

Keterangan:

$\gamma, r, d$  = Parameter kernel

$x$  = Variabel *input*

$x_i$  = fungsi basis

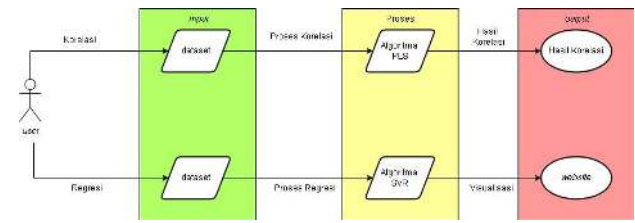
## E. Website

*Website* merupakan media informasi yang ada di internet. *Website* dapat digunakan untuk penyebaran informasi dan berbagai kegunaan lainnya. *Website* ini adalah kumpulan dari halaman situs yang biasanya terdapat dalam sebuah domain atau subdomain *World Wide Web* (WWW) di internet. Halaman dari sebuah *website* dapat diakses melalui sebuah *Uniform Resource Locator* (URL). URL mengatur halaman-halaman situs untuk menjadi sebuah hirarki [10].

## III. METODE

### A. Gambaran Umum Sistem

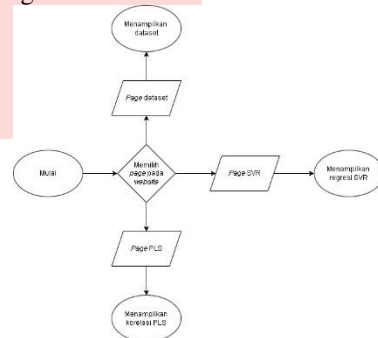
Fungsi dari sistem yang dibuat adalah untuk menemukan korelasi antara kasus Demam Berdarah *Dengue* (DBD) dengan Iklim menggunakan algoritma *Partial Least Square* (PLS) dan melakukan regresi terhadap kasus DBD menggunakan algoritma *Support Vector Machine* (SVM). Hasil dari sistem ini ditampilkan pada *website* berupa grafik dan tabel angka.



GAMBAR  
GAMBARAN UMUM SISTEM.

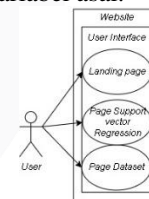
Pengguna dapat melihat hasil regresi yang disajikan dalam tabel dan grafik dengan menekan *page Support Vector Machine* pada situs. Hasil regresi yang dihasilkan adalah hasil dari permodelan SVR pada kerangka *Flask*. Pengguna juga dapat melihat dataset yang digunakan dalam melakukan permodelan SVR dan PLS pada *page dataset*. Pada algoritma PLS, hasil korelasi ditampilkan pada kerangka *Flask*.

### B. Perancangan Website



GAMBAR  
DIAGRAM ALUR WEBSITE.

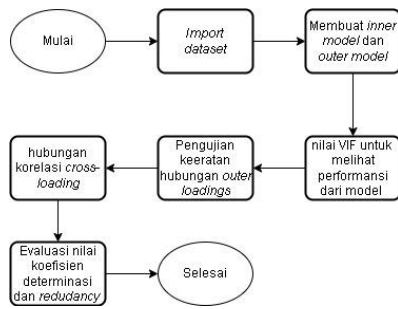
Pada halaman awal *website*, terdapat tiga *page* yang dapat dipilih oleh pengguna. Pada *page dataset*, pengguna akan melihat tampilan dari dataset yang digunakan dalam permodelan algoritma SVR. Pada *page SVR*, terdapat visualisasi dari hasil permodelan menggunakan algoritma SVR, berupa tabel angka dan grafik. Pada *page PLS*, terdapat tabel hasil korelasi *cross-loadings* yaitu hubungan antara variabel target dengan variabel asal.



GAMBAR  
USE CASE DIAGRAM DARI WEBSITE.

Pengguna dapat mengakses *landing page* tanpa harus melakukan login terlebih dahulu, *page Support Vector Regression* untuk melihat visualisasi regresi kasus DBD dan penjelasan SVR, dan *page dataset* untuk melihat dataset yang digunakan untuk melakukan permodelan menggunakan algoritma PLS dan SVR.

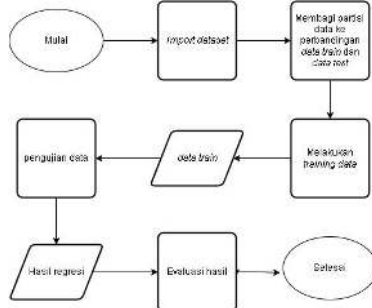
### C. Perancangan Algoritma *Partial Least Square*



GAMBAR  
DIAGRAM ALUR RANCANGAN ALGORITMA PLS.

Hal pertama yang dilakukan adalah melakukan *import dataset* lalu membuat *inner model* dan *outer model*. Setelah model terbentuk, dilakukan perhitungan nilai VIF untuk melihat performansi dari model, semakin tinggi nilai VIF maka performansi dari model akan semakin baik. Selanjutnya, melihat keerratan hubungan *outer loadings*. Hubungan antara variabel dependen dengan variabel independent akan semakin bagus apabila nilai dari *outer loading* lebih besar dari 0.7. setelah mendapatkan hasil *outer loadings* yang bagus, dari model tersebut dapat dilihat korelasi hubungan antara variabel dependen dan independent dengan melihat nilai pengujian *cross-loadings*. Selanjutnya, model dievaluasi nilai koefisien determinasi dan *mean redundancy*. Semakin tinggi nilai koefisien determinasi dan *mean redundancy*, maka model menunjukkan kemampuan variabel independent semakin mampu mengukur variasi variabel endogen dan semakin tinggi nilai akurasi.

#### D. Perancangan Algoritma *Support Vector Regression*



GAMBAR  
DIAGRAM ALUR DARI RANCANGAN ALGORITMA SVR

Hal pertama yang dilakukan adalah *import data*, kemudian partisi data dibagi ke dalam perbandingan antara *training data* dan *testing data*. Selanjutnya, *training data* berguna agar metode SVR dapat mempelajari model himpunan data yang ada. Setelah data terlatih, pengujian data dilakukan dan hasil prediksi keluar. Angka *error* dan koefisien determinasi menunjukkan tingkat bagus dan akurasi dari permodelan yang telah dibuat. Semakin kecil angka *error* dan semakin tinggi nilai koefisien determinasi menandakan model yang dihasilkan semakin bagus dan semakin akurat.

#### E. Pengujian Partisi Data dan Parameter

Langkah pertama dalam pengujian partisi adalah menentukan perbandingan antara *training data* dan *testing data* yang ingin diuji. Setelah itu nilai *error* dan koefisien

determinasi akan keluar. Pengujian ini dilakukan sebanyak jumlah perbandingan partisi data yang dijelaskan dalam tabel pembagian partisi data yang diuji.

TABEL  
PEMBAGIAN PARTISI DATA YANG DIUJI.

| Training Data | Testing Data |
|---------------|--------------|
| 60 %          | 40 %         |
| 70 %          | 30 %         |
| 80 %          | 20 %         |
| 90 %          | 10 %         |

Pencarian parameter terbaik dilakukan setelah melakukan proses pengujian partisi data. Penetapan angka parameter yang dicoba berdasarkan kumpulan data yang dijelaskan pada tabel kelompok angka parameter.

TABEL  
KELOMPOK ANGKA PARAMETER.

|         |   |
|---------|---|
| C       | 1, 10, 100, 1000, ...   |
| Epsilon | 0.01, 0.025, 0.05, 0.075, 0.1, 0.125, 0.15, 0.175, ...  |
| Gamma   | 0.01, 0.075, 0.15, 0.225, 0.3, 0.375, 0.45, 0.525, 0.6, 0.675, 0.75, 0.825, 0.9, 0.975, 1, 1.125, 1.2, 1.275, ... |

#### IV. HASIL DAN PEMBAHASAN

Pada penelitian ini, terdapat beberapa pengujian yang dilakukan, yaitu pengujian *alpha*, pengujian metode *Partial Least Square*, dan pengujian metode *Support Vector Regression*.

##### A. Pengujian Alpha

Pengujian *alpha* dilakukan dengan mengamati hasil yang dijalankan melalui data uji dan memverifikasi fungsionalitas *website*. Pengujian *alpha* adalah jenis pengujian yang dilakukan untuk mengidentifikasi semua kemungkinan kesalahan sebelum mempublikasikan produk ke pengguna. Pengujian ini dilakukan mulai dari proses masuk *website* hingga pemilihan *page* yang ada pada *website*. Perhitungan akurasi pengujian *alpha* menggunakan persamaan:

$$\text{Akurasi} = \frac{\text{Pengujian berhasil}}{\text{Total Pengujian}} \times 100 \%$$

$$\text{Akurasi} = \frac{3}{3} \times 100 \% = 100 \%$$

Perhitungan di atas dapat disimpulkan tidak terdapat masalah atau kesalahan pada web dan semua fungsionalitas berjalan dengan baik dengan hasil akurasi 100%.

##### B. Pengujian metode *Partial Least Square*

Terdapat beberapa tahapan dalam melakukan pengujian algoritma *Partial Least Square*. Pertama, melakukan pengujian model, selanjutnya yaitu pengujian terhadap *Outer Loadings* dan *Cross-Loadings*, dan yang terakhir mendapatkan nilai Koefisien Determinasi dan *Redudancy*.



Model dari PLS ini dibagi menjadi 2 yaitu *inner model* dan *outer model*. Variabel asal yang dibentuk pada *inner model*, yaitu *Cases* dan Variabel target yang dibentuk adalah *Humidity*, *Temperature*, dan *Sunlight*. Pada *outer model*, variabel target dibentuk sesuai dengan kolom yang ada pada dataset yaitu, *Cases* berisikan kolom "KasusDBD", *Humidity* berisikan kolom "RHavg", *Temperature* berisikan kolom "Tavg", dan *Sunlight* berisikan kolom "ss". Setelah terbentuk model, model tersebut diukur performansinya dengan mendeteksi nilai multikolinearitas pada setiap variabelnya. Nilai VIF yang semakin besar dan kurang dari 10 menunjukkan bahwa performansi dari model semakin bagus dan tidak adanya pengaruh nilai multikolinearitas.

TABEL  
PENGUJIAN MODEL PLS

| Variabel  | VIF       |
|-----------|-----------|
| Kasus DBD | 3.587076  |
| RHavg     | 8.142478  |
| Tavg      | 12.004641 |
| ss        | 5.493723  |

Pada pengujian *outer Loadings* dan *cross Loadings*, bertujuan untuk melakukan pengujian keeratan hubungan antar *inner model* dan *outer model* melalui nilai *outer loadings* dan melihat korelasi hubungan antara variabel asal dengan variabel target melalui nilai *cross-loadings*.

TABEL  
PENGUJIAN OUTER LOADINGS

| Variabel  | Outer Loadings |
|-----------|----------------|
| Kasus DBD | 1.0            |
| RHavg     | 1.0            |
| Tavg      | 1.0            |
| ss        | 1.0            |

nilai *outer loadings* memiliki nilai 1 yang berarti semua variabel pada *inner model* memiliki keeratan hubungan yang sangat bagus terhadap *outer model*. Sehingga, model ini dapat dilihat hubungan korelasi antar variabel asal dengan variabel target pada *outer model*.

TABEL  
PENGUJIAN CROSS LOADINGS

|              | Cases    | Sunlight  | Temperature | Humidity |
|--------------|----------|-----------|-------------|----------|
| KasusD<br>BD | 1.000000 | 0.117148  | 0.162691    | 0.315865 |
| ss           | 0.117148 | 1.000000  | -0.077002   | 0.725719 |
| Tavg         | 0.162691 | -0.077002 | 1.000000    | 0.085357 |
| RHavg        | 0.315865 | 0.725719  | 0.085357    | 1.000000 |

Nilai korelasi yang paling tinggi yaitu hubungan antara kelembapan rata-rata (RHavg) dengan lamanya penyinaran matahari (ss) dengan nilai *cross-loadings* 0.725719. Apabila hasil pengujian ini dilihat dari variabel target yang berhubungan dengan kasus DBD, Kelembapan rata-rata (Tavg) memiliki nilai korelasi yang paling tinggi dengan nilai *cross-loadings* 0.315865.

Koefisien determinasi bertujuan untuk melihat akurasi dari permodelan. Nilai koefisien determinasi yang semakin tinggi berarti nilai akurasi juga semakin tinggi. Pada pengujian ini dilihat nilai akurasi dari setiap variabel target untuk melihat seberapa tinggi akurasi dari variabel target.

TABEL  
KOEFSIEN DETERMINASI SETIAP VARIABEL TARGET

| Variabel                         | R <sup>2</sup> |
|----------------------------------|----------------|
| Kelembapan rata-rata (RHavg)     | 0.099771       |
| Temperatur rata-rata (Tavg)      | 0.026468       |
| Lamanya penyinaran matahari (ss) | 0.013724       |

Nilai koefisien determinasi (R<sup>2</sup>) yang paling tinggi yaitu kelembapan rata-rata (RHavg), yaitu 0.099771. Hal ini menunjukkan bahwa kelembapan rata-rata memiliki tingkat akurasi yang paling bagus dibandingkan dengan variabel lainnya. Namun, nilai akurasi ini bisa dibilang tidak terlalu bagus karena masih jauh dari nilai 1.

### C. Pengujian metode Support Vector Regression

Pada pengujian metode SVR, hal pertama yang diuji adalah partisi data. Partisi data dibagi menjadi beberapa pengujian perbandingan antara *data training* dan *data testing*. Selanjutnya, pengujian parameter dilakukan dengan memasukkan nilai C, gamma, dan epsilon yang diuji beberapa kali. Terakhir, melakukan pengujian kernel. Pada pengujian

kernel ini, kernel yang akan diuji, yaitu kernel *Radial Basis Function* (RBF), kernel Linear, kernel Polinomial, dan kernel Sigmoid. Pengujian ini dilakukan beberapa kali sehingga mendapatkan hasil regresi yang paling baik.

Pada pengujian SVR, pengujian dilakukan dengan mencari nilai *error* dan koefisien determinasi dari setiap pengujian partisi data yang dibagi, kernel yang digunakan, dan nilai parameter yang dimasukkan. Setelah itu, didapatkan kesimpulan dari nilai *error* dan koefisien determinasi dari pengujian SVR ini.

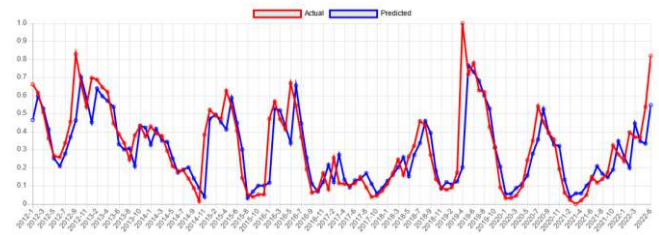
TABEL  
KESIMPULAN DARI PENGUJIAN METODE SVR

| Partisi data | Kernel     | Parameter                              | RMSE    | MSE     |
|--------------|------------|--|---------|---------|
| 90:10        | RBF        | C = 10,<br>Epsilon = 0.05, Gamma = 1   | 0.12177 | 0.57791 |
|              | Linear     | C = 1,<br>Epsilon = 0.05, Gamma = 0.5  | 0.13312 | 0.45775 |
|              | Polinomial | C = 10,<br>Epsilon = 0.05, Gamma = 0.5 | 0.16211 | 0.07723 |
|              | Sigmoid    | C = 1,<br>Epsilon = 0.1, Gamma = 0.3   | 0.13982 | 0.27374 |
| 80:20        | RBF        | C = 10,<br>Epsilon = 0.05, Gamma = 1   | 0.12314 | 0.58873 |
|              | Linear     | C = 1,<br>Epsilon = 0.01, Gamma = 0.01 | 0.13285 | 0.45708 |
|              | Polinomial | C = 10,<br>Epsilon = 0.05, Gamma = 1   | 0.15823 | 0.09483 |
|              | Sigmoid    | C = 1,<br>Epsilon = 0.1, Gamma = 0.3   | 0.14325 | 0.20033 |
| 70:30        | RBF        | C = 1,<br>Epsilon = 0.01, Gamma = 1    | 0.12834 | 0.52656 |
|              | Linear     | C = 1,<br>Epsilon = 0.01, Gamma = 1    | 0.13305 | 0.51751 |
|              | Polinomial | C = 1,<br>Epsilon = 0.01, Gamma = 1    | 0.18621 | 0.24119 |
|              | Sigmoid    | C = 1,<br>Epsilon = 0.1, Gamma = 0.3   | 0.14813 | 0.27334 |
| 60:40        | RBF        | C = 10,<br>Epsilon = 0.05, Gamma = 0.5 | 0.13452 | 0.50131 |
|              | Linear     | C = 10,<br>Epsilon = 0.05, Gamma = 0.5 | 0.14069 | 0.48724 |
|              | Polinomial | C = 10,<br>Epsilon = 0.05, Gamma = 0.5 | 0.17747 | 0.13023 |
|              | Sigmoid    | C = 1,<br>Epsilon = 0.1, Gamma = 0.3   | 0.14168 | 0.28434 |

Dari kesimpulan yang didapatkan dari tabel kesimpulan di atas, nilai *error* dan koefisien determinasi yang paling bagus dan akurat adalah partisi data dengan perbandingan 80:20. Partisi data ini menggunakan kernel *Radial Basis Function* (RBF) dan nilai parameter masing-masing C = 10, gamma = 1, dan epsilon = 0.05. Nilai RMSE dan R<sup>2</sup> terbaik

masing-masing dari pengujian yang didapat adalah 0.12314 dan 0.58873.

Setelah mendapatkan hasil yang paling baik dari pengujian partisi data, parameter, dan kernel, maka hasil regresi dari metode SVR dapat divisualisasikan dengan grafik *actual* dan *predict*.



Gambar hasil regresi metode SVR

Gambar diatas merupakan grafik regresi yang dihasilkan dari pengujian SVR terbaik yang dilakukan.

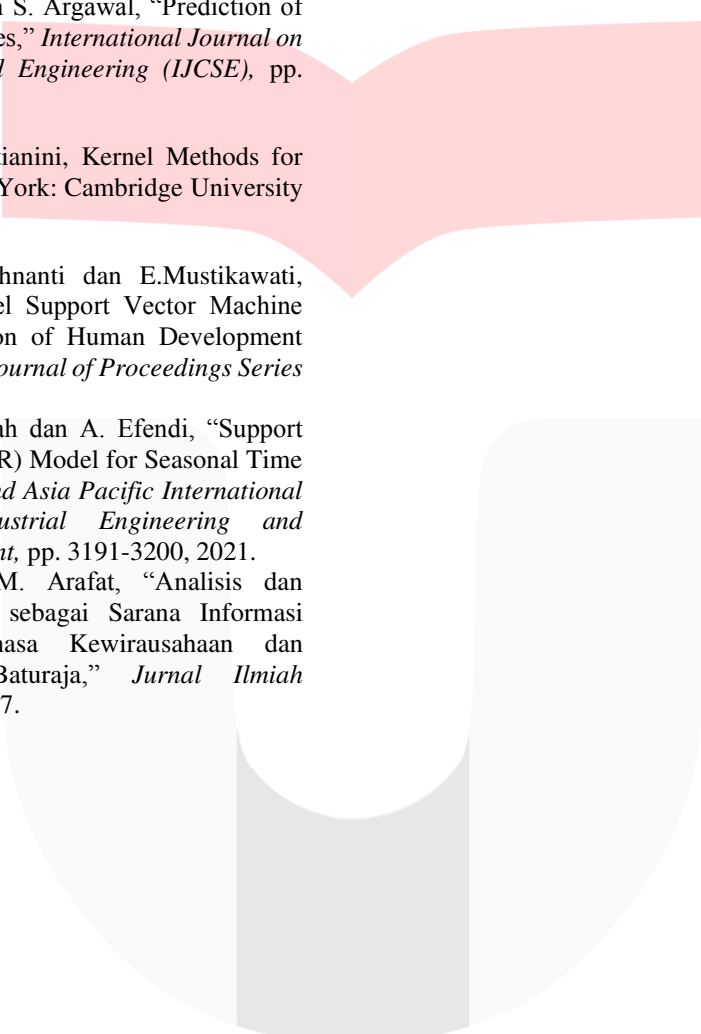
## V. KESIMPULAN

Dari hasil pengujian yang telah dilakukan, dari tugas akhir ini dapat disimpulkan bahwa algoritma *Partial Least Square* (PLS) dan *Support Vector Regression* (SVR) dapat termodelkan dengan baik. Hasil korelasi yang didapat menunjukkan bahwa kelembapan memiliki pengaruh lebih tinggi terhadap jumlah kasus DBD. nilai *cross-loadings* yang didapatkan adalah 0.315865 dengan nilai akurasi koefisien determinasi (R<sup>2</sup>) 0.099771. Pada data kasus DBD ini, kombinasi antara partisi data, parameter dan kernel yang paling baik adalah partisi data yang memiliki perbandingan 80:20 antara *data training* dan *data testing* dengan nilai parameter masing-masing C = 10, gamma = 1, dan epsilon = 0.05. Kernel yang digunakan adalah kernel *Radial Basis Function* (RBF). Hasil regresi dari data kasus DBD, memiliki nilai *error* dan koefisien determinasi masing-masing *Root Mean Score Error* (RMSE)= 0.12314 dan R<sup>2</sup> = 0.58873. Nilai *error* dan koefisien determinasi menunjukkan performansi dari model yang dibuat. Semakin tinggi nilai *error* dan koefisien determinasi, maka hasil yang didapatkan dari permodelan akan semakin baik.

## REFERENSI

- [1] Syamsir dan A. Daramusseng, “ANALISIS SPASIAL EFEKTIVITAS FOGGING DI

WILAYAH KERJA PUSKESMAS MAKROMAN, KOTA SAMARINDA,” *Jurnal Nasional Ilmu Kesehatan*, pp. 1-7, 2018.

- 
- [2] M. R. Ridha, L. Indriyati, A. Tomia dan Juhairiyah, "PENGARUH IKLIM TERHADAP KEJADIAN DEMAM BERDARAH DENGUE," *Spirakel*, pp. 53-62, 2019.
- [3] A. Kusairi dan R. yulia, "Mapping of Dengue Fever Distribution Based on Indonesian National Standard Cartography Rules as an Prevention Indicator of Outbreaks," *Jurnal pendidikan IPA Indonesia*, pp. 91-96, 2020.
- [4] Q. Yang, "Regression," *Encyclopedia of Big Data*, pp. 1-4, 2017.
- [5] G. D. Garson, "Partial Least Square : Regression & Structural Equation Models," *Statistical Associates Publishing*, pp. 51-56, 2016.
- [6] D. Tomar, R. Arya dan S. Argawal, "Prediction of Profitability of Industries," *International Journal on Computer Science and Engineering (IJCSE)*, pp. 1938-1944, 2011.
- [7] J. Shawe dan N. Cristianini, *Kernel Methods for Pattern Analysis*, New York: Cambridge University Press, 2004.
- [8] H. A. Azies, D. trishnanti dan E. Mustikawati, "Comparison of Kernel Support Vector Machine (SVM) in Classification of Human Development Index (HDI)," *IPTEK Journal of Proceedings Series*, pp. 53-57, 2019.
- [9] H. Muthiah, U. Sa'adah dan A. Efendi, "Support Vector Regression (SVR) Model for Seasonal Time Series Data," *the Second Asia Pacific International Conference on Industrial Engineering and Operations Management*, pp. 3191-3200, 2021.
- [10] Y. Trimarsiah dan M. Arafat, "Analisis dan Perancangan Website sebagai Sarana Informasi Pada Lembaga Bahasa Kewirausahaan dan Komputer AKMI Baturaja," *Jurnal Ilmiah MATRIK*, pp. 1-10, 2017.