

# PREDIKSI INDEKS STANDAR PENCEMAR UDARA MENGGUNAKAN METODE *SUPPORT VECTOR REGRESSION* BERBASIS WEB (STUDI KASUS PADA DKI JAKARTA)

## *PREDICTION OF AIR POLLUTANT STANDARD INDEX USING THE WEB-BASED SUPPORT VECTOR REGRESSION METHOD (CASE STUDY IN THE JAKARTA CAPITAL SPECIAL REGION)*

Irgihansyah Izzul Haque<sup>1</sup>, Dr. Meta Kallista<sup>2</sup>, Casi Setianingsih<sup>3</sup>

<sup>1,2,3</sup>Prodi S1 Teknik Komputer, Fakultas Teknik Elektro,  
Universitas Telkom

<sup>1</sup>[irghiansyah@student.telkomuniversity.ac.id](mailto:irghiansyah@student.telkomuniversity.ac.id),

<sup>2</sup>[metakallista@telkomuniversity.ac.id](mailto:metakallista@telkomuniversity.ac.id), <sup>3</sup>[setiacasie@telkomuniversity.ac.id](mailto:setiacasie@telkomuniversity.ac.id)

---

### Abstrak

Udara adalah salah satu kebutuhan utama bagi semua makhluk hidup di muka bumi. DKI Jakarta adalah ibu kota Negara Indonesia yang padat penduduk karena perannya sebagai kota metropolitan. Mobilitas yang padat setiap harinya di DKI Jakarta membuat polusi udara kurang baik.

Prediksi bisa disinonimkan dengan peramalan. Peramalan didefinisikan sebagai prosedur di mana data historis dimasukkan ke dalam model dan keluaran atau hasil dari model tersebut berupa prakiraan. Prediksi dilakukan menggunakan metode *support vector regression* berdasarkan kumpulan data berupa indeks standar pencemar udara (ISPU) yang didapatkan dari web Jakarta Open Data Dinas Lingkungan Hidup DKI Jakarta. Ada 5 parameter di 5 stasiun yang diprediksi. Dengan adanya prediksi ISPU membuat masyarakat dapat mengantisipasi lebih awal terkait pencemaran udara.

Prediksi yang dilakukan diimplementasikan ke dalam web yang menggunakan kerangka flask. Kernel yang digunakan pada penelitian ini adalah *radial basis function*. Hasil dari pengujian partisi data tiap zat adalah PM10 80% data training 20% data testing, SO2 50% data training 50% data testing, CO 80% data training 20% data testing, O3 70% data training 30% data testing, dan NO2 50% data training 50% data testing. Sementara itu, hasil pengujian parameter terbaik dari semua zat menghasilkan kisaran nilai RMSE dari 0.052748 sampai 0.117431 dan  $R^2$  dari 0.393280 sampai 0.91676.

**Kata Kunci:** ISPU, Prediksi, *Support Vector Regression*, Web.

---

### Abstract

Air is one of the main needs of all living things on earth. DKI Jakarta is the capital of Indonesia, which is densely populated because of its role as a metropolitan city. Dense mobility every day in DKI Jakarta makes the air less good.

Predictions can be synonymous with forecasting. Forecasting is defined as the procedure by which historical data is fed into the model and the output or outcome of that model is an estimate. The prediction was made based on a dataset in the form of air pollution standard index (APSI) obtained from the Jakarta Open Data website of DKI Jakarta Environment Agency using the support vector regression method. There are 5 parameters in 5 predicted stations. With the prediction of APSI makes the public can anticipate early related to air pollution.

The predictions are implemented into the web using the Flask framework. The results of data partition testing of each substance are PM10 80% training data 20% data testing, SO2 50% data training 50% data testing, CO 80% data training 20% data testing, O3 70% data training 30% data testing, and NO2 50% data training 50% data testing. Meanwhile, the best parameter testing results of all substances resulted in a range of RMSE values from 0.052748 to 0.117431 and  $R^2$  from 0.393280 to 0.91676.

**Keywords:** APSI, Prediction, *Support Vector Regression*, Web.

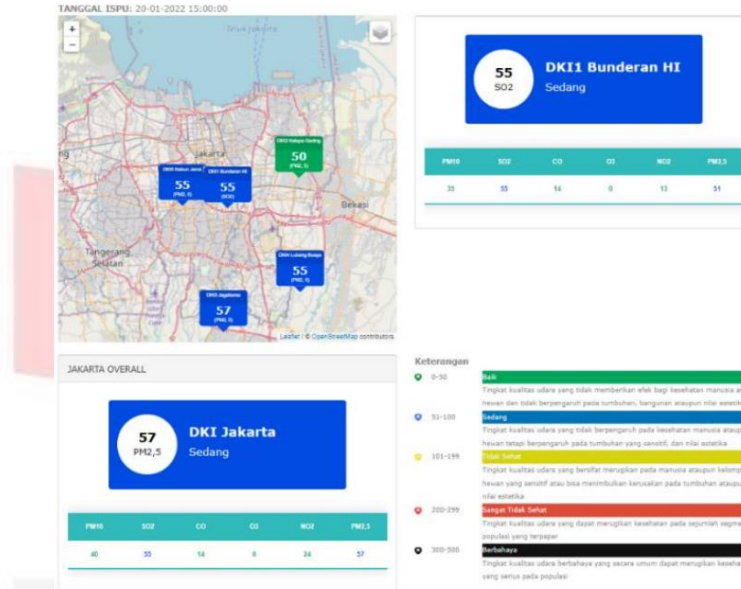
---

### 1. Pendahuluan

Udara adalah salah satu kebutuhan utama bagi semua makhluk hidup di muka bumi. Kualitas udara sangat mempengaruhi kesehatan dan kualitas hidup makhluk hidup di bumi. Kualitas udara yang baik akan berdampak positif bagi kehidupan begitupun sebaliknya, kualitas udara yang

buruk akan berdampak buruk bagi kehidupan. Bumi dan seisinya harus pintar dalam menjaga kualitas udara dengan mengenal dampak dan pengaruh udara di berbagai sektor. Indonesia merupakan negara yang di dalamnya masih banyak kota-kota dengan kualitas udara yang kurang baik.

Pencemaran udara menjadi gangguan untuk kesehatan bagi kehidupan di darat maupun kehidupan akuatik [1]. Banyak aktivitas bahkan kualitas hidup makhluk hidup di bumi yang terganggu dan menurun akibat adanya pencemaran udara ini mengingat udara yang tidak baik dapat menyebabkan masalah kesehatan.



**Gambar 1.1** Status dan pemetaan kualitas udara di DKI Jakarta.

Sumber: Web SILIKA DKI Jakarta diakses pada tanggal 21 Januari 2022 pukul 14:41 WIB.

Gambar 1.1 menunjukkan dari 5 stasiun yang 4 di antaranya memiliki status kualitas udara sedang dan 1 stasiun berstatus baik. Indeks stasiun DKI 1 berada di Gambar 1.1 Status dan pemetaan kualitas udara di DKI Jakarta. 18 angka 55, stasiun DKI 2 berada di angka 50, stasiun DKI 3 berada di angka 57, stasiun DKI 4 berada di angka 55, stasiun DKI 5 berada di angka 55, dan rata-rata kualitas udara di DKI Jakarta berada di angka 57. Pada web SILIKA DKI Jakarta belum ada fitur untuk melihat prediksi ISPU padahal fitur tersebut bermanfaat bagi masyarakat. Keuntungan dari memprediksi ISPU adalah masyarakat dapat mengantisipasi lebih awal terkait kondisi kualitas udara yang akan terjadi termasuk pencemaran udara.

ISPU adalah angka yang menunjukkan kualitas udara. Zat-zat yang menjadi parameter dalam ISPU sebagai kontrol kualitas udara, yaitu partikulat ( $PM_{10}$ ), sulfur dioksida ( $SO_2$ ), ozon ( $O_3$ ), karbon monoksida ( $CO$ ), nitrogen dioksida ( $NO_2$ ). Kelima parameter ini yang akan diprediksi menggunakan metode *support vector regression*.

*Support vector regression* atau SVR adalah bentuk aplikasi *support vector machine* atau SVM yang paling umum dengan cara kerja mencoba mengecilkan batas kesalahan yang digeneralisasi [2]. Metode SVR dapat dipakai untuk memecahkan masalah regresi.

Hasil dari pengerjaan Tugas Akhir ini, yaitu berupa *website* yang menampilkan prediksi indeks standar pencemar udara menggunakan *support vector regression* berdasarkan 5 stasiun di DKI Jakarta. Dengan adanya sistem ini dapat memudahkan masyarakat dalam memprediksi kualitas udara di DKI Jakarta.

## 2. Dasar Teori

### 2.1 Indeks Standar Pencemar Udara

Menurut Peraturan Menteri Lingkungan Hidup dan Kehutanan Republik Indonesia Nomor P.14/MENLHK/SETJEN/KUM.1/7/2020 tentang Indeks Standar Pencemar Udara, Indeks Standar Pencemar Udara atau yang biasa disingkat ISPU adalah angka yang menunjukkan kualitas udara berdasarkan efek bagi kesehatan dan nilai estetika, angka ini tidak mempunyai satuan.

**Tabel 2.1** Kategori, status warna, dan rentang angka indeks standar pencemar udara.

Kategori	Warna	Angka
Baik	Hijau	1 – 50
Sedang	Biru	51 – 100
Tidak Sehat	Kuning	101 – 200
Sangat Tidak Sehat	Merah	201 - 300
Berbahaya	Hitam	≥ 301

Sumber: Peraturan Menteri Lingkungan Hidup dan Kehutanan RI Nomor

P.14/MENLHK/SETJEN/KUM.1/7/2020 Tentang Indeks Standar Pencemar Udara.

Tabel 2.1 menunjukkan bahwa semakin tinggi angka rentang kategori ISPU, maka semakin berbahaya bagi kesehatan makhluk hidup yang mengonsumsinya. Angka-angka tersebut didapat dari hasil kalkulasi zat pencemar yang telah diambil.

Pada Peraturan Menteri Lingkungan Hidup dan Kehutanan Republik Indonesia Nomor P.14/MENLHK/SETJEN/KUM.1/7/2020 tentang Indeks Standar Pencemar Udara pasal 10 menyatakan bahwa pemerintah harus menyebarkan informasi hasil pengkategorian ISPU dengan parameter partikulat (PM<sub>10</sub>), sulfur dioksida (SO<sub>2</sub>), karbon monoksida (CO), ozon (O<sub>3</sub>), dan nitrogen dioksida (NO<sub>2</sub>) diinformasikan paling sedikit dua kali dalam sehari.

ISPU sendiri dihitung dengan persamaan sebagai berikut:

$$I = \frac{I_a I_b}{X_a X_b} (X_x - X_b) + I_b \quad (2.1)$$

Keterangan:

I = ISPU terhitung

I<sub>a</sub> = ISPU batas atas

I<sub>b</sub> = ISPU batas bawah

X<sub>a</sub> = Konsentrasi ambien batas atas (µg/m<sup>3</sup>)

X<sub>b</sub> = Konsentrasi ambien batas bawah (µg/m<sup>3</sup>)

X<sub>x</sub> = Konsentrasi ambien nyata hasil pengukuran (µg/m<sup>3</sup>)

## 2.2 Machine Learning

Pembelajaran mesin adalah bagian dari pemrosesan data dan menggunakan teknik statistik untuk memungkinkan komputer memahami data tanpa diprogram secara eksplisit [3]. Ada 3 jenis machine learning, yaitu [4]:

### 1. Supervised Learning

*Supervised learning* adalah tipe *machine learning* yang memiliki algoritma yang menggunakan data berlabel. Input dengan output yang mudah dibaca. Selain itu, pembelajaran terawasi memiliki algoritma yang memungkinkan untuk memodifikasi modelnya sendiri untuk menyesuaikan model dengan hasil yang diinginkan.

### 2. Unsupervised Learning

Tipe ini merupakan kebalikan dari *supervised learning*. Dengan metode ini, tidak ada spesifikasi khusus untuk data yang digunakan atau dikelola. Oleh karena itu, sistem tidak mengetahui jawaban atau keluaran yang benar. Tujuannya adalah untuk memeriksa data dan mengetahui struktur di dalamnya yang biasa digunakan untuk data transaksional.

### 3. Semi-supervised Learning

*Semi-supervised learning* biasanya menggunakan data berlabel dan tidak melatih algoritma belajar mandiri. Tipe ini biasa digunakan untuk data berlabel dengan jumlah presentase lebih sedikit daripada data tanpa label.

## 2.3 Regresi

Regresi adalah metode analisis statistik untuk mengidentifikasi hubungan antara variabel. Ketergantungan variabel dapat berupa regresi univariat atau multivariat. Regresi univariat mengidentifikasi ketergantungan di antara variabel tunggal. Hubungan bisa diidentifikasi antara variabel dependen dan independen dengan deskripsi sebagai berikut [5].

$$y = a + bx + \varepsilon \quad (2.2)$$

Keterangan:

y= variabel dependen

a= konstanta

b= koefisien variabel x

$\varepsilon$ = kesalahan

Setelah pemodelan regresi telah berhasil dilakukan selanjutnya adalah melakukan perhitungan untuk membuktikan bagus atau tidaknya kinerja dari regresi. Nilai *error* dan koefisien determinasi bisa menjadi pengukur dari kinerja regresi. Terdapat beberapa perhitungan yang digunakan untuk menilai kesalahan dan akurasi, di antaranya *mean absolute error* (MAE), *mean squared error* (MSE), *root mean squared error* (RMSE), dan koefisien determinasi ( $R^2$ ). *Root mean squared error* (RMSE) adalah akar kuadrat dari akar kuadrat rata-rata dari semua nilai *error*. Penggunaan RMSE sangat umum, dan dianggap sebagai metrik kesalahan tujuan umum yang sangat baik untuk prediksi numerik [6].

#### 1. Root Mean Squared Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}} \quad (2.3)$$

Keterangan:

$Y_i$  = Nilai Asli

$\hat{Y}_i$  = Prediksi

n = jumlah data

#### 2. Koefisien Determinasi

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \quad (2.4)$$

$y_i$  = Observasi respon ke-i

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

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

Semakin kecil nilai MAE, MSE, dan RMSE menandakan pemodelan semakin baik. Sedangkan, semakin besar koefisien determinasi dan mendekati 1 menandakan pemodelan semakin baik.

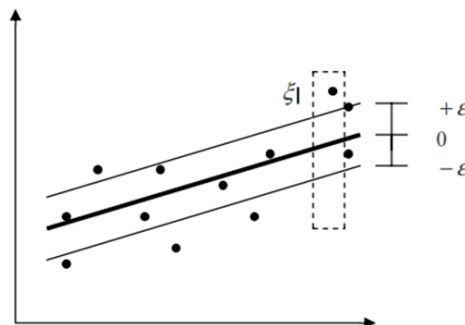
## 2.4 Support Vector Machine

*Support vector machines* atau yang biasa dikenal SVM adalah salah satu algoritma *machine learning* yang paling efektif dari sisi praktis maupun teoretis. Konsep dasar dari SVM sebenarnya kombinasi harmonis dari teori-teori komputasi yang telah ada, seperti batas atau margin hyperplane kernel yang dilansir oleh Aronszajn pada 1950 dan juga konsep-konsep pendukung yang lain tetapi hingga tahun diperkenalkannya algoritma ini, belum pernah ada percobaan untuk merangkai konsep-konsep tersebut sebagai komponen. Support vector machine sendiri dikembangkan oleh Boser, Guyon, Vapnik, dan dipublikasikan pertama kali pada tahun 1992 di *Annual Workshop on Computational Learning Theory* [7].

Perbedaan SVM dan *neural network* adalah SVM berupaya mencari dan menemukan *hyperplane* yang terbaik pada ruang input, sedangkan *neural network* berusaha mencari *hyperplane* pemisah antar kelas. Prinsip dasar SVM adalah *linear classifier*, selanjutnya dikembangkan agar dapat bekerja pada masalah non-linier. dengan memasukkan konsep trik kernel pada ruang kerja dengan dimensi tinggi.

## 2.5 Support Vector Regression

*Support vector regression* atau SVR adalah bentuk aplikasi *support vector machine* atau SVM yang paling umum. SVR juga metode yang dapat mengatasi overfitting artinya adalah suatu kondisi dimana model dibangun dengan memperhitungkan seluruh ciri yang ada, termasuk *noise* [8].



Gambar 2.2 Ilustrasi fungsi support vector regression [2].

Gambar 2.2 menunjukkan ilustrasi dari fungsi SVR,  $\epsilon$  adalah epsilon,  $\xi$  adalah data *slack*, 2 garis tipis adalah batas *margin* atau zona epsilon, dan garis tebal di tengah adalah *hyperplane*. Salah satu karakteristik utama dari SVR adalah bahwa alih-alih meminimalkan kesalahan pelatihan yang diamati, SVR mencoba mengecilkan batas kesalahan yang digeneralisasi sehingga mencapai kinerja yang digeneralisasi. Batas kesalahan generalisasi ini adalah kombinasi dari kesalahan pelatihan dan istilah regularisasi adalah yang mengontrol kompleksitas ruang hipotesis [2]. Berikut ini adalah fungsi regresi dari metode SVR.

$$f(x) = \omega \cdot x + b \quad (2.5)$$

Keterangan:

$\omega$ = Vektor bobot

$x$ = Fungsi yang memetakan  $x$  dalam suatu dimensi

$b$ = bias

### 2.5.1. Parameter-parameter SVR

Adapun pemengaruh performansi SVR, yaitu parameter  $C$ , epsilon ( $\epsilon$ ), dan gamma ( $\gamma$ ), dengan penjelasan sebagai berikut:

#### 1. Parameter $C$

Parameter ini menentukan *trade off* antara kompleksitas model dan jumlah, sehingga deviasi yang lebih besar dari epsilon dapat ditolerir [9]. Untuk nilai  $C$  yang besar, pengoptimal memilih *hyperplane* dengan *margin* kecil untuk mengklasifikasikan semua titik pelatihan dengan tepat. Sebaliknya, jika nilai  $C$  sangat kecil, pengoptimal akan mencari *margin* yang lebih besar untuk memisahkan *hyperplane*. Hal itu bisa menyebabkan salah pengklasifikasian di banyak titik.

#### 2. Parameter Epsilon ( $\epsilon$ )

Parameter ini menandakan lebar dari zona regresi [9]. Semakin besar nilai epsilon, maka garis regresi akan semakin datar.

#### 3. Parameter Gamma ( $\gamma$ )

Gamma adalah parameter fungsi kernel [9]. Parameter ini dapat menentukan seberapa besar pengaruh sampel set pelatihan. Nilai rendah berarti "jauh" dan nilai tinggi berarti "dekat".

### 2.5.2. Kernel dan Fungsinya

Adapun macam-macam kernel yang dipakai dalam sistem algoritma SVR dengan fungsinya masing-masing. Analisis *support vector machines* (SVM) adalah alat pembelajaran mesin yang populer untuk klasifikasi dan regresi, SVM mendukung regresi linier dan nonlinier yang dapat kita sebut sebagai SVR [10]. Salah satu fungsi kernel yang umum digunakan adalah *radial basis function* (RBF). RBF sudah terbukti bisa diaplikasikan ke dalam kasus non-linier pemetaan data pelatihan menjadi berhingga pada dimensi ruang, RBF paling sering dalam bentuk [11].

$$K(x, x_i) = \exp \left\{ -\frac{\|x - x_i\|^2}{2\sigma^2} \right\}, \sigma > 0 \quad (2.6)$$

Keterangan:

$x$ = variabel input

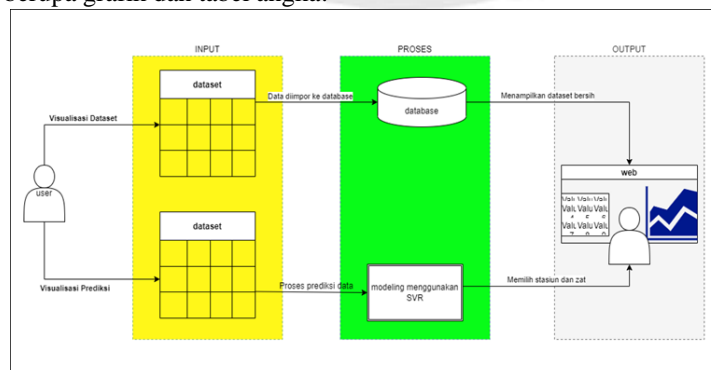
$x_i$ = inti-inti yang dipilih dari pelatihan

$\sigma$ = spread

## 3. Pembahasan

### 3.1. Gambaran Umum Sistem

Fungsi dari sistem yang dibuat adalah untuk memprediksi ISPU yang akan datang menggunakan algoritma *support vector regression* (SVR) dan hasilnya ditampilkan di situs jaringan berupa grafik dan tabel angka:



Gambar 3.1 Gambaran umum sistem yang dibuat.

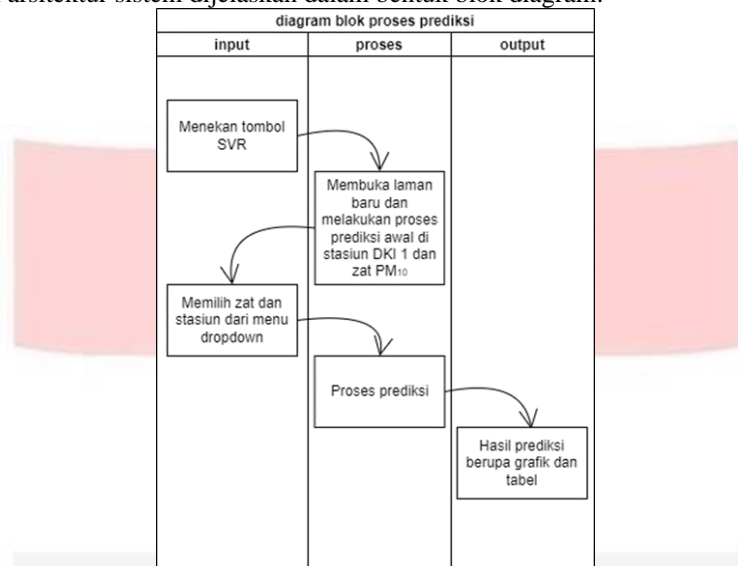


### 3.2. Pengumpulan Data

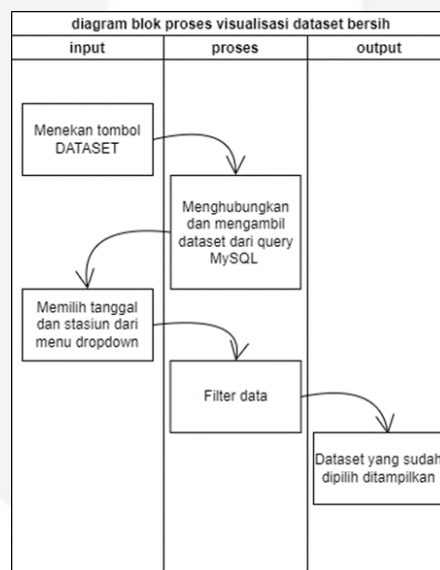
Tahapan pertama pada penelitian adalah mengumpulkan data. Data merupakan unsur paling penting pada penelitian karena objek utama yang akan diolah adalah data. Himpunan data yang diambil berasal dari web Open Data DKI Jakarta dengan rentang 1 Januari 2015 sampai 31 Oktober 2021. Data yang diambil dengan format judul “Indeks Standar Pencemar Udara di SPKU-Bulan-Tahun” dan berbentuk CSV.

### 3.3. Perancangan Sistem

Ada 2 perancangan sistem yang dibuat berdasarkan 2 proses yang ada di dalam sistem. Perancangan arsitektur sistem dijelaskan dalam bentuk blok diagram.



**Gambar 3.2** Diagram blok arsitektur proses prediksi dalam web.



**Gambar 3.3** Diagram blok proses visualisasi dataset bersih.

### 3.4. Preprocessing

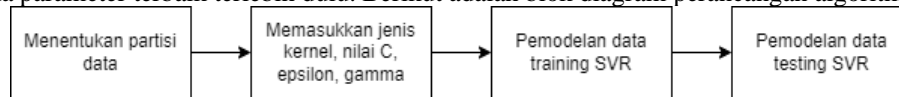
Himpunan data perlu diproses dengan cara pemodelan agar data bisa diolah oleh algoritma, sehingga didapatkan hasil prediksi. Penyiapan data tersebut dinamakan preprocessing.

1. Setelah mengambil data, total hanya 7 kolom parameter saja yang akan diambil untuk diproses, yaitu tanggal, stasiun, pm10, so2, co, o3, dan no2
2. Mengubah format penulisan tanggal menjadi “Year – Month – Day”.
3. Mengubah data kosong ‘---’ menjadi NaN.
4. Mengganti format nama stasiun yang salah.
5. Mengubah angka-angka ISPU dari *string* menjadi *float*.
6. Data-data kosong yang sudah diubah menjadi NaN tadi akan diisi. Banyak cara untuk mengisi data yang kosong satu di antaranya adalah interpolasi linier.

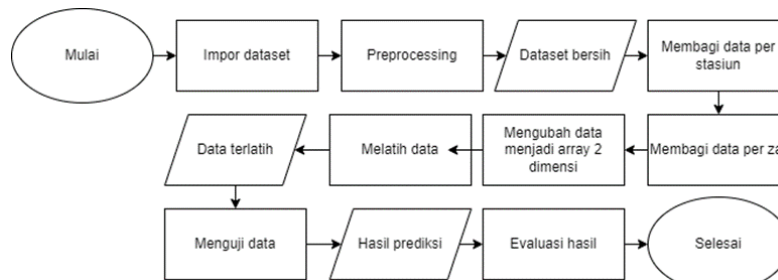
7. Menyaring data yang akan diolah oleh SVR.

### 3.5. Perancangan Algoritma SVR

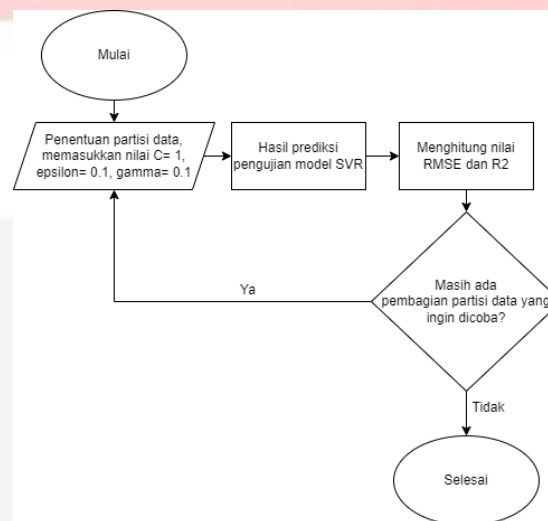
Perancangan metode *support vector regression* dimulai dengan menguji partisi data dan angka-angka parameter terbaik terlebih dulu. Berikut adalah blok diagram perancangan algoritma.



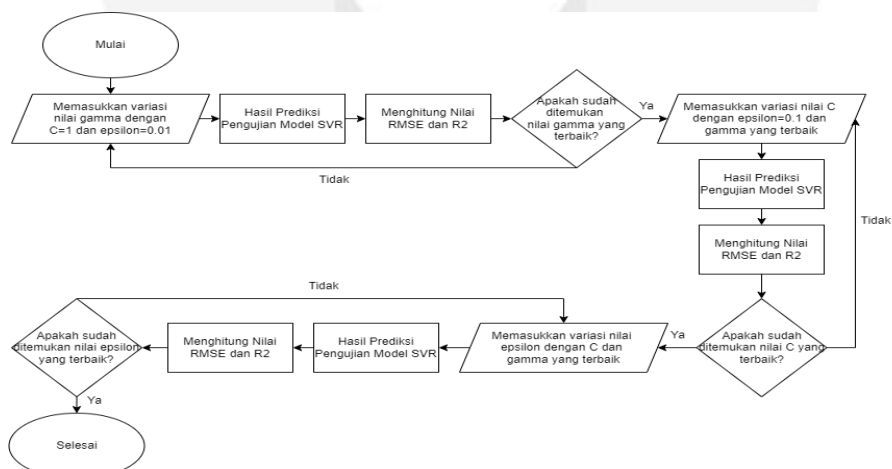
Gambar 3.4 Blok diagram yang menjelaskan rancangan algoritma.



Gambar 3.5 Diagram alir dari rancangan model sistem SVR.



Gambar 3.6 Diagram alir yang menjelaskan alur pengujian partisi data.



Gambar 3.7 Diagram alir yang menjelaskan alur pencarian angka parameter terbaik.

### 3.6. Pengujian Partisi Data

Pengujian partisi data ini menggunakan parameter *default* SVR, yaitu  $C=1$ ,  $\gamma=1$ , dan  $\epsilon=0.1$ . Sedangkan untuk kernel yang digunakan adalah RBF karena selain data terdistribusi secara radial kernel ini juga mudah diimplementasikan. Hasil pengujian dapat dilihat di bawah ini.

**Tabel 3.1** Rangkuman hasil pengujian partisi data terbaik.

Zat	Training	Testing	RMSE	R2
PM10	80%	20%	0.088652	0.524170
SO2	50%	50%	0.144172	0.660124
CO	80%	20%	0.074962	0.742130
O3	70%	30%	0.080973	0.544199
NO2	50%	50%	0.076048	0.619669

### 3.7. Pengujian Parameter SVR

Pengujian nilai parameter SVR dilakukan secara berurut dari pencarian nilai gamma, C, lalu epsilon. Pengujian ini memakai pengaturan awal  $C = 1$  dan  $\epsilon = 0.01$ . Pengujian ini bertujuan untuk mencari parameter-parameter terbaik yang akan digunakan dan memakai kernel RBF. Berikut adalah zat-zat di setiap stasiun beserta parameter-parameter terbaik yang sudah ditemukan:

**Tabel 3.2** Rangkuman hasil pengujian parameter terbaik.

Stasiun	Zat	Gamma	C	Epsilon	RMSE	R2
DKI 1	PM10	0.675	1	0.05	0.087608	0.535320
	SO2	0.375	10	0.01	0.075499	0.906794
	CO	1.125	100	0.01	0.052885	0.871652
	O3	0.675	10	0.01	0.060898	0.742188
	NO2	0.225	1	0.025	0.065257	0.719946
DKI 2	PM10	0.6	1	0.05	0.111026	0.569661
	SO2	0.6	1	0.01	0.074666	0.899641
	CO	0.075	1	0.01	0.104020	0.837899
	O3	1.2	1	0.01	0.062280	0.789782
	NO2	0.225	1	0.025	0.103659	0.494437
DKI 3	PM10	0.15	10	0.525	0.086667	0.393280
	SO2	0.225	1	0.01	0.087938	0.856040
	CO	0.075	1	0.01	0.110165	0.703662
	O3	0.375	10	0.01	0.072332	0.824990
	NO2	0.375	1	0.01	0.064641	0.549751
DKI 4	PM10	0.15	1	0.01	0.087323	0.559398
	SO2	0.3	100	0.01	0.052748	0.916760
	CO	0.075	1	0.01	0.117431	0.486597
	O3	0.225	1	0.01	0.089726	0.767316
	NO2	0.075	10	0.01	0.087245	0.594546
DKI 5	PM10	0.225	1	0.1	0.099929	0.495046
	SO2	0.225	10	0.01	0.068149	0.892325
	CO	0.075	1	0.01	0.077919	0.739426
	O3	1.5	1	0.01	0.095394	0.736527
	NO2	0.15	1	0.01	0.061959	0.732144

## 4. Kesimpulan

Berdasarkan hasil penelitian, pengujian, dan analisis yang dilakukan dalam tugas akhir ini



dapat diambil kesimpulan, sebagai berikut:

1. hasil dari pengujian partisi data tiap zat adalah PM<sub>10</sub> 80% data training 20% data testing, SO<sub>2</sub> 50% data training 50% data testing, CO 80% data training 20% data testing, O<sub>3</sub> 70% data training 30% data testing, dan NO<sub>2</sub> 50% data training 50% data testing.
2. Hasil pengujian parameter terbaik dari semua zat menghasilkan kisaran nilai RMSE dari 0.052748 sampai 0.117431 dan  $R^2$  dari 0.393280 sampai 0.91676.

#### Referensi

- [1] P. D. Landge and R. R. Harne, "Air Quality Monitoring System for City : A Review," *Int. Res. J. Eng. Technol.*, vol. 5, no. 1, pp. 5–6, 2018, [Online]. Available: <https://www.irjet.net/archives/V5/i1/IRJET-V5I1110.pdf>.
- [2] D. Basak, S. Pal, and D. C. Patranabis, "Support vector regression," *Neural Inf. Process. – Lett. Rev.*, vol. Vol. 11, no. 10, pp. 203–224, 2007, doi: 10.1145/2768566.2768568.
- [3] A. L. Samuel, "Eight-move opening utilizing generalization learning. (See Appendix B, Game G-43.1 Some Studies in Machine Learning Using the Game of Checkers)," *IBM J.*, pp. 210–229, 1959.
- [4] W. Apt, "Introduction," *Demogr. Res. Monogr.*, pp. 1–13, 2014, doi: 10.1007/978-94-007-6964-9\_1.
- [5] S. Kavitha, S. Varuna, and R. Ramya, "A comparative analysis on linear regression and support vector regression," *Proc. 2016 Online Int. Conf. Green Eng. Technol. IC-GET 2016*, 2017, doi: 10.1109/GET.2016.7916627.
- [6] S. P. Neill and M. R. Hashemi, *Ocean Modelling for Resource Characterization*. 2018.
- [7] V. N. Vapnik, "The Nature of Statistical Learning," *Theory*. p. 334, 1995, [Online]. Available: <https://ci.nii.ac.jp/naid/10020951890>.
- [8] D. Tomar, "Prediction of Profitability of Industries using Weighted SVR," vol. 3, no. 5, pp. 1938–1944, 1938.
- [9] P. Cortez and A. Morais, "A Data Mining Approach to Predict Forest Fires using Meteorological Data," *Proc. 13th Port. Conf. Artif. Intell.*, pp. 512–523, 2007, [Online]. Available: <http://www.dsi.uminho.pt/~pcortez/fires.pdf>.
- [10] D. Nguyen, "Learning Data Science — Predict Stock Price with Support Vector Regression (SVR)," *ITNEXT*, 2019. <https://itnext.io/learning-data-science-predict-stock-price-with-support-vector-regression-svr-2c4fdc36662> (accessed Jan. 19, 2022).
- [11] Q. Huang, J. Mao, and Y. Liu, "An improved grid search algorithm of SVR parameters optimization," *Int. Conf. Commun. Technol. Proceedings, ICCT*, no. 2, pp. 1022–1026, 2012, doi: 10.1109/ICCT.2012.6511415.