

Pembentukan Model Regresi Linier Menggunakan Algoritma Genetika untuk Prediksi Parameter Indeks Standar Pencemar Udara (ISPU)

Fitri Insani¹, Sri Indah Darlianti²

^{1,2}Teknik Informatika, UIN Sultan Syarif Kasim Riau

Jl. H.R. Soebrantas no. 155 KM. 18 Simpang Baru, Pekanbaru 28293

fitri.insani@uin-suska.ac.id¹, sri.indah.darlianti@students.uin-suska.ac.id²

Abstrak – Prediksi merupakan upaya untuk mengetahui suatu peristiwa di masa yang akan datang. Pada kasus prediksi data yang dipelajari merupakan data historis, agar data tersebut dapat menghasilkan informasi berupa hasil prediksi maka digunakan suatu model yaitu regresi linier, dalam proses pembentukan model regresi linier digunakan pendekatan kecerdasan buatan algoritma genetika. Algoritma genetika digunakan untuk mendapatkan nilai koefisien terbaik pada persamaan regresi linier. Penelitian ini menggunakan data parameter Indeks Standar Pencemar Udara (ISPU), terdapat lima parameter ISPU meliputi Sulfur dioksida (SO₂), Partikulat (PM₁₀), Karbondioksida (CO), Ozon (O₃), dan Nitrogen dioksida (NO₂) sehingga dibangun lima model prediksi. Berdasarkan hasil pengujian menggunakan data harian bulan Januari 2016 sampai dengan Desember 2016, algoritma genetika mampu menentukan nilai koefisien yang digunakan pada model regresi linier dalam memprediksi parameter ISPU dengan kesalahan prediksi untuk parameter SO₂ yaitu 2,33958%, kesalahan prediksi parameter PM₁₀ 6,623923%, kesalahan prediksi parameter CO 2,62279%, kesalahan prediksi parameter O₃ 6,34495%, dan kesalahan prediksi parameter NO₂ 2,927575%

Kata Kunci – Algoritma Genetika, ISPU, Kualitas Udara, Prediksi, Regresi linier

PENDAHULUAN

A. Latar Belakang

Meningkatnya jumlah kendaraan bermotor di Indonesia setiap tahun mengakibatkan semakin banyaknya gas buangan dari kendaraan tersebut hal ini mengakibatkan pencemaran di udara hingga 70 sampai 80 persen[1]. Tingginya tingkat pencemaran udara mengakibatkan semakin menurunnya kualitas udara. Terkait dengan masalah kualitas udara, maka pengukuran dan pengamatan kualitas udara harus selalu dilakukan[2]. Di Indonesia pengukuran

standar kualitas udara yang resmi yaitu Indeks Standar Pencemar Udara (ISPU), hal ini sesuai dengan Keputusan Menteri Negara Lingkungan Hidup Nomor : KEP 45 / MENLH / 1997 Tentang Indeks Standar Pencemar Udara (ISPU). Parameter ISPU meliputi Sulfur dioksida (SO₂), Partikulat (PM₁₀), Karbon monoksida (CO), Ozon (O₃), dan Nitrogen dioksida (NO₂). Kumpulan data harian ISPU tersebut dapat dimanfaatkan untuk penelitian seperti melakukan prediksi parameter ISPU pada masa yang akan datang. Agar dapat memprediksi parameter ISPU digunakan suatu model yaitu regresi linier, model ini menggunakan data runtun waktu atau data *time series* untuk melakukan peramalan, dan data yang digunakan berbentuk numerik[3]. Berdasarkan hal tersebut regresi linier dirasa sesuai untuk melakukan prediksi pada kasus yang sedang teliti. Dalam proses pembentukan model regresi linier digunakan pendekatan kecerdasan buatan algoritma genetika. Algoritma genetika pada model regresi linier digunakan untuk mendapatkan koefisien regresi terbaik yang nantinya nilai koefisien hasil algoritma genetika tersebut digunakan dalam model regresi linier untuk pencarian prediksi. Algoritma Genetika merupakan algoritma pencarian heuristik yang terinspirasi dari cara kerja evolusi biologis[3]. Algoritma genetika melakukan proses evolusi yaitu menggunakan proses seleksi alam. Pada saat berevolusi, individu akan mengalami perubahan gen secara terus-menerus agar dapat menyesuaikan diri dengan lingkungannya. Jadi individu yang mampu bertahan dan menyesuaikan diri dengan lingkungannya adalah individu yang muncul sebagai solusi yang optimal.

B. Rumusan Masalah

Rumusan masalah yang didapatkan dari penjabaran permasalahan di latar belakang yaitu bagaimana penerapan model regresi linier yang dibentuk dengan menggunakan algoritma genetika untuk mendapatkan nilai koefisien dalam memprediksi parameter ISPU.

C. Batasan Masalah

Adapun batasan pada penelitian yaitu :

1. Penelitian menggunakan lima model prediksi sesuai dengan lima parameter ISPU meliputi prediksi Sulfur dioksida (SO₂), Partikulat (PM₁₀), Karbon monoksida (CO), Ozon (O₃), dan Nitrogen dioksida (NO₂).
2. Data masukan berupa data time series empat hari sebelumnya untuk masing-masing parameter ISPU.
3. Data keluaran berupa hasil prediksi beserta nilai MSE prediksi dari masing-masing parameter ISPU.
4. Data yang digunakan berjumlah 366 data yang merupakan data periode harian dalam rentang waktu 1 Januari 2016 sampai dengan 31 Desember 2016 yang didapatkan dari Laboratorium Udara Kota Pekanbaru.
5. Algoritma genetika pada tahap inisialisasi menggunakan real coded, proses crossover menggunakan extend intermediate, proses mutasi menggunakan random mutation, dan seleksi menggunakan replacement selection.

D. Tujuan Penelitian

Penelitian ini dibuat untuk menerapkan model regresi linier yang modelnya dibentuk dengan algoritma genetika untuk mendapatkan nilai koefisien dalam memprediksi parameter ISPU.

LANDASAN TEORI

A. Pencemaran Udara

Pencemaran udara terjadi karena adanya zat asing di udara yang menyebabkan komposisi udara berubah dari keadaan normal menjadi tidak normal. Jika zat asing tersebut ada di udara dalam jumlah dan waktu yang lama akan menyebabkan gangguan pada kelangsungan makhluk hidup di muka bumi[4].

Di Indonesia pencemaran udara dibahas dalam Undang-Undang Nomor 23 tahun 1997 pasal 1 ayat 12, yaitu pencemaran udara adalah pencemaran yang disebabkan oleh aktivitas manusia seperti pencemaran yang berasal dari pabrik, kendaraan bermotor, pembakaran sampah, sisa pertanian, dan peristiwa alam seperti kebakaran hutan, letusan gunung api yang mengeluarkan debu, gas, dan awan panas. Pencemaran udara dapat menyebabkan terganggunya aktivitas dan kelangsungan hidup seluruh makhluk hidup yang ada dipermukaan bumi karena dapat mempengaruhi keseimbangan kualitas udara normal.

B. Indeks Standar Pencemar Udara (ISPU)

Sesuai dengan Keputusan Menteri Negara Lingkungan Hidup Nomor : KEP 45 / MENLH / 1997 Tentang Indeks Standar Pencemar Udara, standar kualitas udara yang digunakan secara resmi di Indonesia yaitu Indeks Standar Pencemaran Udara (ISPU). Keputusan ini dibuat untuk menyeragamkan informasi kualitas udara kepada masyarakat di lokasi dan waktu tertentu dan untuk memudahkan upaya pengendalian pencemaran udara.

ISPU merupakan informasi berbentuk angka dan tidak memiliki satuan yang menggambarkan bagaimana kondisi kualitas udara pada lokasi dan waktu tertentu berdasarkan bagaimana dampaknya terhadap kesehatan makhluk hidup. Kriteria dan rentang kualitas udara dapat dilihat pada Tabel 1 yang bersumber dari situs Kementerian Lingkungan Hidup dan Kehutanan Republik Indonesia[5].

Tabel 1. Kriteria Kualitas Udara

Kategori	Rentang	Penjelasan
Baik	1-50	Tingkat kualitas udara yang tidak memberikan efek bagi kesehatan manusia atau hewan dan tidak berpengaruh pada tumbuhan, bangunan ataupun nilai estetika
Sedang	51-100	Tingkat kualitas udara yang tidak memberikan efek bagi kesehatan manusia atau hewan, tetapi berpengaruh pada tumbuhan yang sensitif dan nilai estetika
Tidak Sehat	101-199	Tingkat kualitas udara yang bersifat merugikan pada manusia ataupun kelompok hewan yang sensitif atau bisa menimbulkan kerusakan pada tumbuhan ataupun nilai estetika
Sangat Tidak Sehat	200-299	Tingkat kualitas udara yang dapat merugikan kesehatan pada sejumlah segmen populasi yang terpapar
Berbahaya	300-lebih	Tingkat kualitas udara berbahaya yang secara umum dapat merugikan kesehatan yang serius pada populasi

Berdasarkan Keputusan Kepala Badan Pengendalian Dampak Lingkungan, Parameter-Parameter Dasar yang digunakan dalam perhitungan

ISPU dan periode waktu pengukurannya dapat dilihat pada Tabel 2.

Tabel 2. Parameter dan Periode Pengukuran Waktu ISPU

PARAMETER	WAKTU PENGUKURAN
Partikulat (PM ₁₀)	24 jam (periode pengukuran rata-rata)
Sulfur Dioksida (SO ₂)	24 jam (periode pengukuran rata-rata)
Karbon Monoksida (CO)	8 jam (periode pengukuran rata-rata)
Ozon (O ₃)	1 jam (periode pengukuran rata-rata)
Nitrogen Dioksida (NO ₂)	1 jam (periode pengukuran rata-rata)

Keterangan :

1. Hasil pengukuran untuk pengukuran kontinyu diambil harga rata-rata tertinggi dari waktu pengukuran.
2. Informasi kualitas udara disampaikan kepada masyarakat setiap 24 jam dari data rata-rata sebelumnya.
3. Waktu terakhir pengambilan data dilakukan pukul 15.00 WIB.
4. Informasi ISPU yang disampaikan kepada masyarakat berlaku 24 jam ke depan dari pukul 15.00 tanggal (n) sampai pukul 15.00 tanggal (n+1).

C. Prediksi

Prediksi merupakan suatu proses memperkirakan secara sistematis tentang suatu kemungkinan di masa depan berdasarkan data historis ataupun data-data yang memiliki pengaruh untuk kejadian dimasa yang akan datang[6]. Terdapat dua pendekatan yang sering digunakan pada proses prediksi yaitu:

1. Kualitatif.
Bersifat subjektif, hal ini disebabkan oleh pengaruh faktor-faktor seperti emosi, intuisi, dan pengalaman seseorang.
2. Kuantitatif.
Bersifat numerik, disebabkan oleh perhitungan pada bilangan angka dan perhitungan matematis yang beragam berdasarkan data historis terkait dan sebab akibat.

D. Regresi Linier

Analisis regresi termasuk ilmu statistika yang memanfaatkan hubungan antara dua atau lebih variabel kuantitatif sehingga salah satu variabel dapat diramalkan berdasarkan variabel lainnya. Ketika suatu hasil keluaran atau kelas berupa data numerik, dan semua atribut juga merupakan data numerik maka regresi linier adalah teknik yang tepat

untuk menyelesaikan permasalahan tersebut. Regresi linier terbagi dua sebagai berikut:

1. Regresi Linier Sederhana

Regresi linier sederhana didasarkan pada hubungan fungsional ataupun kausal satu variabel independen(bebas) dengan satu variabel dependen (terikat). Variabel bebas (independent) merupakan variabel yang memengaruhi atau yang menjadi sebab perubahannya atau timbulnya variabel dependent (terikat). Sedangkan variabel terikat (dependent) yang dipengaruhi atau yang menjadi akibat karena adanya variabel bebas.

2. Regresi Linier Berganda

Pada prinsipnya persamaan regresi linier berganda sama dengan persamaan pada regresi linier sederhana, yang menjadi pembeda adalah pada regresi linier berganda jumlah variabel independent (prediktor) lebih dari satu. Berikut adalah persamaan untuk regresi linier berganda:

$$Y' = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n \quad (1)$$

Dimana :

Y' = variabel dependent (terikat)

a = konstanta (nilai Y' apabila X = 0)

b = koefisien regresi

X = variabel independent (bebas)

Penelitian ini menggunakan regresi linier berganda karena data yang digunakan untuk memprediksi merupakan data *time series*. Artinya variabel yang akan diprediksi (dependent) memiliki hubungan sebab akibat dengan lebih dari satu variabel bebas (independent) yaitu data kualitas udara beberapa hari sebelumnya.

Dalam suatu model regresi terdapat koefisiensi yang merupakan nilai duga parameter di dalam model regresi untuk kondisi yang sebenarnya (*true condition*). Koefisiensi untuk model regresi merupakan suatu nilai rata-rata yang berpeluang terjadi pada variabel terikat (Y) bila suatu variabel bebas (X) diberikan. Berdasarkan hal tersebut dilakukan proses optimasi untuk mendapatkan koefisien yang paling baik untuk melakukan prediksi dengan menggunakan regresi linier.

E. Optimasi Algoritma Genetika

Algoritma Genetika merupakan algoritma pencarian heuristik yang terinspirasi dari cara kerja dari evolusi biologis[3]. Algoritma genetika melakukan proses evolusi yaitu menggunakan proses seleksi alam. Pada saat berevolusi, individu akan mengalami perubahan gen secara terus-

menerus agar dapat menyesuaikan diri dengan lingkungannya. Pada proses algoritma genetika untuk mendapatkan hasil yang paling optimal, tiap individu akan berevolusi dari generasi ke generasi. Seleksi alam diibaratkan sebagai tahapan mengoptimasi individu-individu tersebut. Jadi individu yang unggul, yang mampu beradaptasi dengan lingkungannya adalah individu yang muncul sebagai solusi yang optimal. Keunggulan Algoritma genetika yaitu memiliki tahapan yang sederhana dan kemampuan dalam mencari solusi terbaik dengan cepat meskipun pada permasalahan yang rumit dan kompleks[7].

Adapun tahapan dalam proses algoritma genetika yaitu[3]:

1. Representasi Individu

Suatu populasi terdiri dari kumpulan individu, sedangkan individu terdiri dari sekumpulan gen. Gen akan mewakili suatu nilai yang menggambarkan suatu maksud tertentu. Pada tahapan inisialisasi ditentukan berapa jumlah individu dan banyaknya gen dalam satu individu pada suatu populasi yang akan dibangkitkan diawal.

Pada penelitian ini digunakan representasi *real* untuk menyelesaikan permasalahan yang diteliti.

2. *Crossover*/Perkawinan Silang

Crossover atau penyilangan dilakukan untuk mendapatkan individu yang baru yaitu dengan cara mengambil dua individu kemudian dilakukan pertukaran gen. Frekuensi *crossover* tergantung dari nilai *Crossover Rate* (*cr*) sehingga untuk menentukan frekuensi *crossover* digunakan persamaan:

$$\text{Frekuensi } Crossover = cr \times \text{jumlah individu} \quad (2)$$

Penelitian ini *crossover* yang digunakan adalah *extended intermediate crossover*. Caranya dengan menentukan terlebih dahulu *crossover rate* (*cr*) lalu tentukan jumlah individu yang akan melakukan *crossover*, selanjutnya memilih dua individu sebagai *parent* secara random untuk melakukan *crossover*, bangkitkan nilai *a* secara *random* pada interval yang telah ditentukan. Barulah melakukan *crossover* dengan menggunakan persamaan:

$$\begin{aligned} C1 &= P1 + a(P2 - P1) \\ C2 &= P2 + a(P1 - P2) \end{aligned} \quad (3)$$

Dimana :

C1 = *child* 1

C2 = *child* 2

P1 = *parent* 1

P2 = *parent* 2

a = bilangan *random*

3. Mutasi

Proses mutasi merupakan perubahan gen dalam kromosom sehingga karakteristiknya berubah. Hal ini bertujuan untuk meningkatkan variasi populasi. Frekuensi dilakukannya mutasi ditentukan oleh nilai *Mutation Rate* (*mr*). Dalam setiap populasi, banyaknya gen yang akan melakukan mutasi ditentukan dengan persamaan:

$$\text{Frekuensi mutasi} = mr \times \text{jumlah individu} \quad (4)$$

Metode mutasi yang digunakan adalah *random mutation* yaitu memilih salah satu individu/kromosom secara *random* kemudian menambahkan atau mengurangi nilai gen yang terpilih dengan bilangan *random* yang kecil. Berikut persamaan yang digunakan dalam proses *random mutation*:

$$xi = xi + r(max_i - min_i) \quad (5)$$

Dimana:

xi = gen yang akan dimutasi

r = bilangan *random*

max_i = nilai maksimum inisialisasi kromosom

min_i = nilai minimum inisialisasi kromosom

4. Evaluasi Individu

Individu-individu yang terbentuk dari proses algoritma genetika dievaluasi untuk menentukan nilai kualitas individu tersebut, yang dikenal sebagai fungsi *fitness*. Penentuan *fitness* bertujuan untuk mengetahui seberapa baik solusi (individu) tersebut. Pada penelitian ini pembentukan koefisien regresi dengan menggunakan algoritma genetika bertujuan untuk meminimalkan nilai *error*. Maka untuk menghitung nilai *fitness* dimulai dengan menghitung nilai prediksi dengan regresi linier kemudian mencari nilai *error* dengan metode *MSE* (*Mean Square Error*) setelah itu hitung nilai *fitness* dengan persamaan:

$$Fitness = \frac{1}{MSE} \quad (6)$$

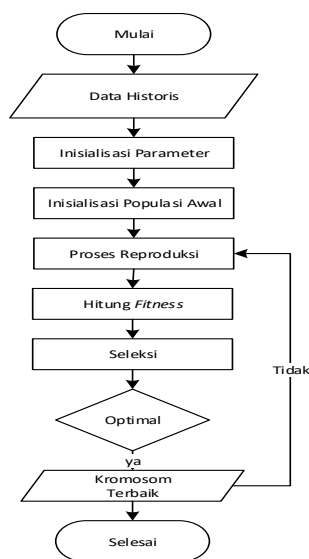
Artinya nilai *fitness* berbanding terbalik dengan nilai *error*, sehingga semakin kecil nilai *error* maka semakin besar nilai *fitness* dan begitu pula sebaliknya.

5. Seleksi
 Tahap seleksi dilakukan dengan *Replacement Selection*. Aturan seleksi dengan menggunakan *Replacement Selection* adalah [8]:
 - Individu yang diproduksi melalui proses *crossover* akan menggantikan induk yang paling lemah jika mempunyai nilai *fitness* yang lebih baik dibandingkan *fitness* induk yang terlemah tersebut.
 - Individu yang diproduksi dari proses mutasi menggantikan induknya jika memiliki nilai *fitness* yang lebih baik.
6. Kriteria Optimasi Tercapai
 Beberapa kriteria berhenti yang sering digunakan antara lain[9]:
 - Berhenti pada generasi tertentu.
 - Berhenti setelah beberapa generasi berturut-turut didapatkan nilai *fitness* tertinggi/terendah (Konvergensi).
 - Berhenti bila dalam n generasi berikutnya tidak diperoleh nilai *fitness* yang lebih tinggi/rendah.

F. Regresi Linier dengan Algoritma Genetika

Algoritma genetika dalam penelitian ini digunakan untuk membangun model regresi linier. Algoritma genetika berfungsi untuk mendapatkan nilai koefisien regresi terbaik untuk nantinya koefisien tersebut digunakan dalam pencarian prediksi dengan menggunakan model regresi linier.

Berikut ini merupakan *flowchart* yang berisi tahapan dalam melakukan prediksi menggunakan model regresi linier yang dibangun dengan algoritma genetika[3].



Gambar 1. *Flowchart* Regresi Linier dengan Algoritma Genetika

G. Mean Squared Error (MSE)

Mean Squared Error (MSE) merupakan rata-rata selisih kuadrat antara nilai prediksi dengan nilai yang diamati. *MSE* dihitung dengan menjumlahkan kuadrat semua kesalahan prediksi kemudian membaginya dengan jumlah periode prediksi. Proses pengujian *MSE* dapat dihitung menggunakan persamaan:

$$MSE = \frac{\sum(Y_i - Y'_i)^2}{n} \quad (7)$$

Dimana :

MSE = nilai *error*

n = jumlah data periode waktu

Y = data aktual

Y' = hasil prediksi

METODOLOGI PENELITIAN

Metodologi penelitian merupakan alur penelitian selama melakukan penelitian tugas akhir. Berikut dijelaskan tahapan-tahapan yang dilakukan dalam penelitian tugas akhir ini.

A. Studi Pustaka

Studi pustaka/studi literatur merupakan teori-teori yang digunakan untuk pedoman dasar dalam pembuatan sistem serta menyelesaikan permasalahan yang ada. Studi literatur bersumber dari buku, jurnal, halaman web, dan literatur yang membahas tentang regresi linier dan algoritma genetika untuk memprediksi parameter ISPU.

B. Pengumpulan Data

Data didapatkan dari Laboratorium Udara Pekanbaru yang berlokasi di kantor walikota Pekanbaru. Data merupakan data Indeks Standar Pencemar Udara (ISPU) yang terdiri dari lima parameter yaitu Sulfur dioksida (SO₂), Partikulat (PM₁₀), Karbon monoksida (CO), Ozon (O₃), dan Nitrogen dioksida (NO₂). Data yang didapatkan dari Laboratorium Udara Pekanbaru berjumlah 366 untuk masing-masing parameter ISPU yang merupakan data periode harian mulai dari tanggal 1 Januari 2016 sampai 31 Desember 2016. Prediksi dilakukan menggunakan data *time-series* dari masing-masing lima parameter ISPU tersebut.

C. Analisa

Tahap analisa dilakukan pada data serta pada metode regresi linier dan algoritma genetika.

Analisa Data

Data yang digunakan sebagai *input-an* merupakan data *time series* kualitas udara empat hari

sebelumnya dari masing-masing parameter ISPU. Data *time series* ini diinisialisasikan sebagai x_1 yaitu data satu hari sebelumnya, x_2 data dua hari sebelumnya, x_3 data tiga hari sebelumnya, dan x_4 data empat hari sebelumnya. Proses prediksi dilakukan satu-persatu terhadap parameter ISPU yang terdiri dari Sulfur dioksida (SO_2), Partikulat (PM_{10}), Karbon monoksida (CO), Ozon (O_3), dan Nitrogen dioksida (NO_2).

Analisa Regresi Linier dan Algoritma Genetika

Tahapan dari proses Regresi Linier dan Algoritma Genetika adalah sebagai berikut:

- Data *input*-an terdiri dari data *time series* salah satu parameter ISPU
- Data *time series* tersebut dikelompokkan menjadi empat periode harian. Misalnya untuk memprediksi data kelima digunakan data *time series* empat hari sebelumnya.
- Inisialisasi dibagi dua yaitu inisialisasi parameter dan inisialisasi koefisien regresi. Pada inisialisasi parameter ditentukan jumlah individu, jumlah generasi, serta nilai *cr* dan *mr*. Sedangkan pada inisialisasi koefisien regresi menggunakan *real-coded genetic algorithm* yaitu dengan membangkitkan bilangan *random* pada interval yang sudah ditentukan. Hasil dari nilai *random* tersebut mewakili proporsi koefisien regresi.
- Lakukan *Crossover* untuk mendapatkan individu baru. dengan menggunakan metode *Extended Intermediate Crossover*.
- Kemudian melakukan mutasi dengan metode *random Mutation*.
- Menghitung nilai *fitness* masing-masing individu, pertama-tama tentukan nilai prediksi parameter ISPU dengan model regresi untuk tiap-tiap individu. Kemudian mencari nilai *error* setiap individu dengan menggunakan metode *MSE (Mean Squared Error)*
- Nilai *error* yang didapatkan digunakan untuk menghitung nilai *fitness*.
- Lakukan seleksi dengan memilih individu yang memiliki nilai *fitness* yang tinggi untuk dapat lolos ke generasi selanjutnya. Seleksi dilakukan dengan menggunakan *Replacement Selection*.
- Individu yang didapatkan dari hasil seleksi akan menjadi populasi baru untuk proses pada generasi selanjutnya. Proses pencarian individu terbaik terus dilakukan hingga kriteria optimasi tercapai.
- Didapatkan *output* berupa individu terbaik yang merupakan nilai koefisien regresi optimum, hasil prediksi menggunakan koefisien tersebut, serta nilai *MSE* dari hasil prediksi.

D. Perancangan

Pada tahap perancangan sistem dilakukan perancangan struktur menu dan perancangan tampilan *interface* pada sistem yang akan dibangun.

Perancangan Struktur Menu

Perancangan struktur menu sistem menggambarkan menu-menu yang ada pada sistem.

Perancangan Interface

Perancangan *interface* digunakan sebagai gambaran komunikasi antara pengguna dengan sistem. *Interface* dirancang untuk memudahkan pengguna ketika menggunakan sistem.

E. Implementasi/Pengkodean

Implementasi merupakan tahapan pembuatan sistem/*software* berdasarkan hasil analisa dan rancangan yang telah dikerjakan sebelumnya. Untuk tahapan implementasi pada penelitian ini perangkat pendukung yang dibutuhkan adalah:

Operating System : Windows 10 64-bit
Program Language : Matlab R2016b

F. Pengujian

Model pengujian yang digunakan adalah model pengujian periode data, pengujian jumlah individu, pengujian jumlah generasi, serta pengujian kombinasi *cr* dan *mr*. Pengujian tersebut dilakukan untuk menguji periode yang paling optimal untuk mendapatkan hasil prediksi parameter ISPU. Pengujian dilakukan pada data periode 4 hari, periode 7 hari, dan periode 10 hari. Pengujian jumlah individu bertujuan untuk mengetahui jumlah individu yang paling optimal. Ukuran jumlah individu yang akan diujikan adalah 200 individu dan 400 individu. Pengujian jumlah generasi bertujuan untuk mengetahui banyaknya generasi yang paling optimal untuk mendapatkan hasil prediksi terbaik. Ukuran generasi yang diujikan yaitu 250, 500, 750 dan 1000 generasi. Sedangkan pengujian kombinasi *crossover rate(cr)* dan *mutation rate(mr)* bertujuan untuk mengetahui kombinasi optimal untuk mendapatkan hasil prediksi. Kombinasi *cr* dan *mr* yang diujikan yaitu *cr* 0,9 dengan *mr* 0,1, *cr* 0,8 dengan *mr* 0,2, *cr* 0,7 dengan *mr* 0,3, *cr* 0,6 dengan *mr* 0,4, *cr* 0,5 dengan *mr* 0,5, *cr* 0,4 dengan *mr* 0,6, *cr* 0,3 dengan *mr* 0,7, *cr* 0,2 dengan *mr* 0,8, dan *cr* 0,1 dengan *mr* 0,9.

G. Kesimpulan dan Saran

Tahapan akhir adalah menyimpulkan hasil akhir dari penelitian pembentukan model regresi linier dengan algoritma genetika untuk prediksi parameter ISPU dan pemberian saran mengenai

HASIL DAN PEMBAHASAN

A. Hasil Analisa Data

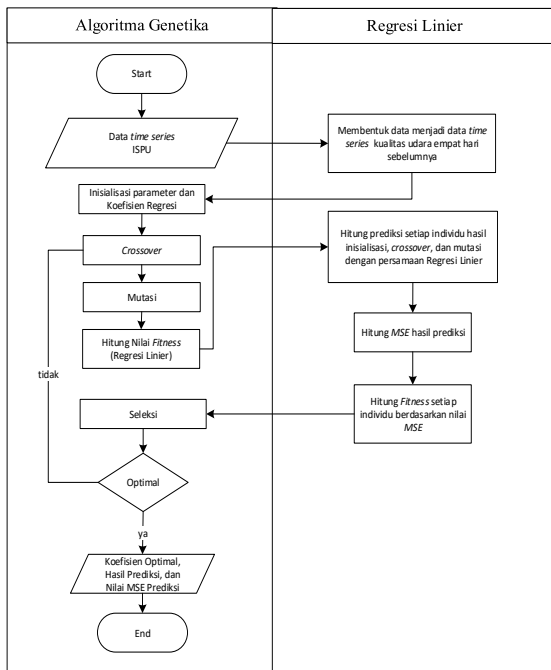
Proses prediksi dilakukan satu-persatu terhadap lima parameter ISPU yang terdiri dari Sulfur dioksida (SO₂), Partikulat (PM₁₀), Karbon monoksida (CO), Ozon (O₃), dan Nitrogen dioksida (NO₂). Data ISPU yang digunakan merupakan data harian mulai dari 1 Januari 2016 sampai dengan 31 Desember 2016. Data dapat dilihat pada Tabel 3.

Tabel 3. Data Lima Parameter ISPU

TANGGAL	PARAMETER				
	SO ₂	PM ₁₀	CO	O ₃	NO ₂
01-Jan-16	4	8	1	15	5
02-Jan-16	4	9	1	14	3
03-Jan-16	4	12	0	29	2
04-Jan-16	4	21	2	21	3
05-Jan-16	4	22	22	22	5
06-Jan-16	4	21	2	23	4
07-Jan-16	4	19	2	19	4
...
26-Dec-16	9	30	7	18	5
27-Dec-16	11	33	7	24	6
28-Dec-16	12	32	6	21	6
29-Dec-16	13	40	9	32	8
30-Dec-16	10	33	7	21	7
31-Dec-16	11	37	7	23	7

B. Analisa Regresi Linier dan Algoritma Genetika

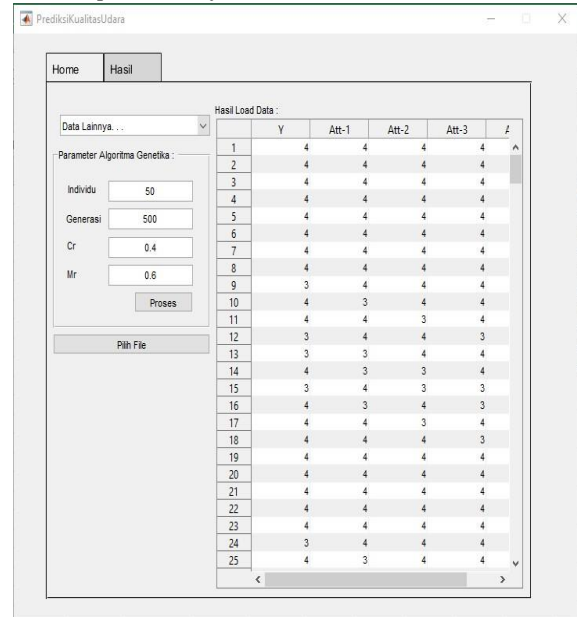
Tahapan-tahapan analisa regresi linier yang dibentuk dengan algoritma genetika dalam memprediksi parameter ISPU dapat dilihat pada Gambar 2.



Gambar 2. Alur Proses Regresi Linier & Algoritma Genetika

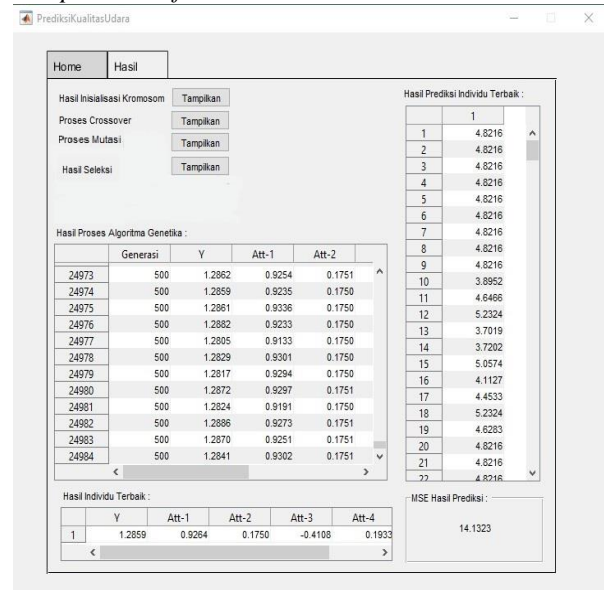
C. Tampilan Sistem

Tampilan Interface Home



Gambar 3. Tampilan Interface Home

Tampilan Interface Hasil



Gambar 4. Tampilan Interface Hasil

D. Pengujian Sistem

Berdasarkan hasil pengujian didapatkan parameter algoritma optimal dan periode waktu optimal untuk prediksi data ISPU yaitu:

Tabel 4. Parameter Optimal

PARAMETER ISPU	IND	GEN	CR & MR	PERIODE WAKTU
SO ₂	200	500	0,1 & 0,9	4 HARI
PM ₁₀	200	750	0,5 & 0,5	4 HARI
CO	400	1000	0,3 & 0,7	7 HARI

PARAMETER ISPU	IND	GEN	CR & MR	PERIODE WAKTU
O ₃	400	1000	0,1 & 0,9	4 HARI
NO ₂	400	1000	0,5 & 0,5	7 HARI

Menggunakan parameter algoritma genetika dan periode waktu optimal tersebut dilakukan prediksi terhadap data ISPU sehingga didapatkan nilai *MSE* prediksi dan persentase *error* prediksi sebagai berikut:

Tabel 5. Nilai *MSE* dan Persentase *Error*

PARAMETER ISPU	<i>MSE</i>	PERSENTASE <i>ERROR</i>
SO ₂	11,7357	2,33958%
PM ₁₀	84,1519	6,623923%
CO	19,7001	2,62279%
O ₃	74,7607	6,34495%
NO ₂	21,4577	2,927575%

KESIMPULAN DAN SARAN

A. Kesimpulan

Kesimpulan dari penelitian ini adalah:

1. Algoritma genetika pada penelitian ini mampu menentukan nilai koefisien yang digunakan pada model regresi linier dalam memprediksi parameter ISPU.
2. Jumlah individu pada algoritma genetika akan mempengaruhi nilai *fitness*, jika ukuran individu semakin kecil maka area dalam pencarian solusi akan semakin sempit sehingga solusi yang ditemukan kurang optimal. Namun jika jumlah individu terlalu besar, kenaikan nilai *fitness* tidak terlalu terlihat sehingga terjadi konvergensi. Hal yang sama juga berlaku pada jumlah generasi.
3. Dari hasil pengujian sistem dapat disimpulkan parameter algoritma genetika dan periode data yang paling optimal:
 - SO₂ adalah 200 individu, generasi 500, kombinasi *cr* dan *mr* adalah 0,1 dan 0,9, dan periode data 4 hari. *MSE* prediksi yaitu 11,7357 dan persentase kesalahan prediksi yaitu 2,33958%.
 - PM₁₀ adalah 200 individu, jumlah generasi 750, kombinasi *cr* dan *mr* adalah 0,5 dan 0,5, dan periode data 4 hari. *MSE* prediksi yaitu 84,1519 dan persentase kesalahan prediksi yaitu 6,623923%.
 - CO adalah 400 individu, jumlah generasi 1000, kombinasi *cr* dan *mr* adalah 0,3 dan 0,7, dan periode data 7 hari. *MSE* prediksi

yaitu 19,7001 dan persentase kesalahan prediksi yaitu 2,62279%.

- O₃ adalah 400 individu, jumlah generasi 1000, kombinasi *cr* dan *mr* adalah 0,1 dan 0,9, dan periode data 4 hari. *MSE* prediksi yaitu 74,7607 dan persentase kesalahan prediksi yaitu 6,34495%.
- NO₂ adalah 400 individu, jumlah generasi 1000, kombinasi *cr* dan *mr* adalah 0,5 dan 0,5, dan periode data 7 hari. *MSE* prediksi yaitu 21,4577 dan persentase kesalahan prediksi yaitu 2,927575%.

B. Saran

Saran yang diberikan untuk pengembangan sistem di masa yang akan datang yaitu rentang nilai *random* yang digunakan untuk inialisasi individu awal disesuaikan dengan kondisi rentang data pada kasus yang diteliti, sehingga untuk mendapatkan hasil prediksi yang baik tidak membutuhkan banyak individu dan generasi hal ini akan mempercepat waktu pencarian solusi

REFERENSI

- [1] Yudhana, A. (2016). Alat Ukur Kualitas Udara di Yogyakarta, 2(1), 17–26.
- [2] Pradana, radyan putra, & Heriyanto, E. (2009). Analisis Pemantauan Kualitas Udara pada saat Arus Mudik dan Balik Lebaran di Gerbang Tol Cikampek Tahun 2009, (2), 261–269.
- [3] Rahmi, A., Mahmudy, W. F., & Darma Setiawan, B. (2015). Prediksi Harga Saham Berdasarkan Data Historis Menggunakan Model Regresi yang Dibangun Dengan Algoritma Genetika.
- [4] Sebayang, M. A. (2017). Stasiun Pemantau Kualitas Udara Berbasis Web/1(1).
- [5] Kementerian Lingkungan Hidup dan Kehutanan. (2019). Kriteria Kualitas Udara. Diakses pada 6 Januari 2019, dari iku.menlhk.go.id
- [6] Herdianto. (2013). Prediksi Kerusakan Motor Induksi Menggunakan Jaringan Syaraf Tiruan Backpropogation.
- [7] Wicaksana, A. S., Setiawan, B. D., & Rahayudi, B. (2018). Algoritma Genetika Untuk Optimasi Fuzzy Time Series Dalam Memprediksi Kepadatan Lalu Lintas di Jalan Tol, 2(3), 1063–1071.
- [8] Rahmi, A., & Mahmudy, W. F. (2016). Pembentukan Model Regresi Harga Saham Menggunakan Algoritma Genetika, 2016(Sentika), 18–19.
- [9] Laksono, A. T., Utami, M. C., & Sugiarti, Y. (2016). Sistem Penjadwalan Kuliah Menggunakan Algoritma Genetika (Studi Kasus: Fakultas Kedokteran dan Kesehatan Universitas Muhammadiyah Jakarta), 177–188