

BAB III

METODOLOGI PENELITIAN

3.1 Objek Penelitian

3.1.1 Subjek penelitian

Penelitian dilakukan dari data Dinas Lingkungan Hidup yaitu berupa data Indeks Standart Pencemaran Udara (ISPU) yang beralamat di Jalan Mandala V No.67, Cililitan, Kramat Jati, Jakarta Timur, Daerah Khusus Ibukota Jakarta 13640. Dinas lingkungan hidup bertugas melayani masyarakat dalam bidang kesehatan lingkungan di kota Jakarta.

3.1.2 Pengambilan Data

Penelitian menggunakan data Indeks standart pencemaran udara (ISPU) kota Jakarta pada bulan Januari s.d Oktober 2018. Data hasil pengamatan yang dilakukan oleh Dinas Lingkungan Hidup kemudian akan diklasifikasi menggunakan data mining dengan Algoritma *K-Nearest Neighbor* (KNN). Kemudian data akan dikelompokan menjadi beberapa kategori sesuai ketentuan yang sudah ada dan hasil klasifikasi akan menemukan data kategori yang terdekat sehingga dapat memprediksi kategori data pencemaran udara dalam bulan November. Berikut adalah data ISPU pada bulan Januari dan data lengkapnya dapat dilihat pada Lampiran.

Tabel 3.1 Data ISPU

Tanggal	lokasi ispu	pm10	SO ₂	CO	O ₃	NO ₂	max	critical	kategori
1/1/2018	DKI4	76	31	38	35	9	76	PM10	SEDANG
1/2/2018	DKI3	23	31	24	39	14	39	O3	BAIK
1/3/2018	DKI3	53	35	35	101	23	101	O3	TIDAK SEHAT
1/4/2018	DKI3	53	49	34	57	15	57	O3	SEDANG
1/5/2018	DKI4	44	32	26	28	10	44	PM10	BAIK
1/6/2018	DKI5	20	30	22	31	5	31	O3	BAIK
1/7/2018	DKI5	26	30	14	46	5	46	O3	BAIK
1/8/2018	DKI5	62	33	19	73	17	73	O3	SEDANG
1/9/2018	DKI3	45	32	16	63	14	63	O3	SEDANG
1/10/2018	DKI5	42	34	16	95	17	95	O3	SEDANG
1/11/2018	DKI5	40	38	18	81	19	81	O3	SEDANG

Tanggal	lokasi ispu	pm10	SO ₂	CO	O ₃	NO ₂	max	critical	kategori
1/12/2018	DKI5	37	33	18	83	17	83	O3	SEDANG
1/13/2018	DKI5	51	30	17	93	21	93	O3	SEDANG
1/14/2018	DKI5	33	72	17	73	14	73	O3	SEDANG
1/15/2018	DKI4	30	72	19	70	12	72	SO2	SEDANG
1/16/2018	DKI4	25	70	19	61	10	70	SO2	SEDANG
1/17/2018	DKI5	35	52	19	60	13	60	O3	SEDANG
1/18/2018	DKI5	41	52	22	56	15	56	O3	SEDANG
1/19/2018	DKI4	46	51	29	48	19	51	SO2	SEDANG
1/20/2018	DKI4	27	51	18	38	7	51	SO2	SEDANG
1/21/2018	DKI4	25	51	18	40	8	51	SO2	SEDANG
1/22/2018	DKI4	26	52	17	32	6	52	SO2	SEDANG
1/23/2018	DKI5	28	52	17	57	9	57	O3	SEDANG
1/24/2018	DKI5	33	53	20	86	19	86	O3	SEDANG
1/25/2018	DKI4	26	53	22	51	15	53	SO2	SEDANG
1/26/2018	DKI4	21	53	19	41	7	53	SO2	SEDANG
1/27/2018	DKI5	27	54	19	63	6	63	O3	SEDANG
1/28/2018	DKI4	34	55	20	52	6	55	SO2	SEDANG
1/29/2018	DKI4	22	54	18	39	7	54	SO2	SEDANG
1/30/2018	DKI5	27	41	22	41	7	41	SO2	BAIK
1/31/2018	DKI5	28	54	21	58	12	58	O3	SEDANG

Keterangan:

1. tanggal : Tanggal pengukuran kualitas udara
2. stasiun : Lokasi pengukuran di stasiun
3. pm10 : Partikulat salah satu parameter yang diukur
4. so2 : Sulfida (dalam bentuk SO₂) salah satu parameter yang diukur
5. co : Carbon Monoksida salah satu parameter yang diukur
6. o3 : Ozon salah satu parameter yang diukur
7. no2 : Nitrogen dioksida salah satu parameter yang diukur
8. max : Nilai ukur paling tinggi dari seluruh parameter yang diukur dalam waktu yang sama
9. critical : Parameter yang hasil pengukurannya paling tinggi
10. kategori : Kategori hasil perhitungan indeks standar pencemaran udara

ISPU dikategorikan kedalam 4 tingkatan pengukuran yaitu :

1. Kategori Baik rentang nilai pengukuran (0-50)
2. Kategori sedang rentang nilai pengukuran (51-100)
3. Kategori Tidak Sehat rentang nilai pengukuran (101-199)
4. Kategori sangat tidak sehat rentan nilai pengukuran (200-299)

Keterangan Kategori Rentang Nilai yaitu :

1. Baik 0 – 50 Hijau, Tingkat kualitas udara yang tidak memberikan efek bagi kesehatan manusia atau hewan dan tidak berpengaruh pada tumbuhan, bangunan ataupun nilai estetika.
2. Sedang 51 – 100 Biru, Tingkat kualitas udara yang tidak berpengaruh pada kesehatan manusia ataupun hewan tetapi berpengaruh pada tumbuhan yang sensitive dan nilai estetika
3. Tidak Sehat 101 – 199 Merah, Tingkat kualitas udara yang bersifat merugikan pada manusia ataupun kelompok hewan yang sensitive atau bisa menimbulkan kerusakan pada tumbuhan ataupun nilai estetika.
2. Sangat Tidak Sehat 200 – 299 Kuning, Tingkat kualitas udara yang dapat merugikan kesehatan pada sejumlah segmen populasi yang terpapar.

3.2 Metode Penelitian

Metode Penelitian yaitu tahapan penelitian yang akan dilakukan oleh peneliti. dalam tahapan ini juga akan menggunakan suatu alat uji penelitian berupa sebuah aplikasi pengujian data mining dengan bahasa pemrograman PHP.

Tahapan penelitian yang dilakukan ditujukan pada gambar 3.2



Gambar 3.1 Tahapan Penelitian

3.2.1 Pengumpulan Data

Data yang digunakan adalah data pengamatan Indeks standart pencemaran udara (ISPU) pada bulan Januari sampai Oktober tahun 2018 yang diperoleh dari *website* Dinas Lingkungan Hidup kota Jakarta.

3.2.2 Dataset

Data hasil pengamatan Indek Standart Pencemaran Udara (ISPU) harian dikota Jakarta pada bulan Januari sampai Oktober tahun 2018. Dataset ini berisi 304 data dengan 10 atribut yaitu tanggal, stasiun. Pm10, SO₂, CO, O₃, NO₂, max, *critical*, *category* dan 4 label kategori yaitu Baik, Sedang, Tidak Sehat, dan Sangat Tidak Sehat

3.2.3 Pengujian dengan Algoritma KNN

Pada penelitian ini pengujian akan mengklasifikasi data hasil pengamatan Indeks standar pencemaran Udara (ISPU) harian dengan menggunakan Algoritma *K-Nearest Neighbor* (KNN) sehingga nanti akan dihasilkan sebuah kategori yang nilai kesamaanya paling banyak



3.2.4 Performance atau Hasil perhitungan

Pada pengujian data set sebanyak 304 data akan dijadikan data pelatihan dan data pengujian. Dengan 4 kategori pencemaran udara yang sudah ditentukan maka Hasilnya akan diketahui kategori tingkat pencemaran udara di Kota Jakarta pada bulan November .

Evaluasi dan validasi hasil hitung smenggunakan rumus akurasi, *precision*, *recall* dan *f-measure* berikut ini:

a. Akurasi

Perhitungan akurasi dilakukan dengan cara membagi jumlah data yang diklasifikasi secara benar dengan total sampel data *testing* yang diuji.

$$\text{Akurasi} = \frac{(TP+TN)}{Tp+TN++FB+FN}$$

b. Precision

menghitung nilai *precision* dengan cara membagi jumlah data bernilai positif (*True Positive*) dibagi dengan jumlah data benar yang bernilai positif (*True Positive*) dan salah bernilai positif (*False Negative*).

$$\text{Precision} = \frac{TP}{TP + FB}$$

c. Recall

Sedangkan *recall* dihitung dengan cara membagi data benar yang bernilai positive (*True Positive*) dengan hasil penjumlahan dari data benar yang bernilai positif (*True Positive*) dan data salah yang bernilai negative (*False Negative*).

$$\text{Recall} = \frac{TP}{TP+FN}$$

d. F-Measure

Nilai *F-Measure* didapat dari perhitungan pembagian hasil dari perkalian *precision* dan *recall* dengan hasil penjumlahan *precision* dan *recall*,kemudian dikalian dua.

$$\text{F - Measure} = 2 * \frac{\text{Precision} * \text{recall}}{\text{Precision} + \text{recall}}$$