

Klasifikasi prediksi kualitas udara Menggunakan *metode Support Vector Machine (SVM)*

1st Renaldy Eka Putra
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

renaldyekaputra@student.telkomuniversit
y.ac.id

2nd Meta Kallista
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

metakallista@telkomuniversity.ac.id

3rd Casi Setianingsih
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

setiacasie@telkomuniversity.ac.id

Abstrak— Penelitian pada kualitas udara sangatlah penting, karena kualitas udara sangat berpengaruh terhadap kesehatan makhluk hidup. Oksigen adalah kebutuhan utama manusia untuk kelangsungan kehidupan, baik buruknya kualitas udara pasti akan sangat berpengaruh terhadap kesehatan. Sehingga kualitas udara sangat penting untuk diteliti lebih lanjut. Pada penelitian ini akan dilakukan dengan metode Support Vector Machine (SVM) untuk mengklasifikasikan kualitas udara, kemudian sebelum data di klasifikasikan akan dilakukan dahulu metode smote untuk membuat data menjadi seimbang. Setelah melakukan pengujian ditemukan kernel terbaik untuk sistem ini yaitu kernel polinomial dengan nilai G-Mean 0.98, Specificity 0.99, dan sensitivity 0.98. Presisi 0.98, Recall 0.98, F1-Score 0.98, serta akurasi 98%, Setelah melakukan pengujian akan mendapatkan hasil klasifikasi terbaik yang akan divisualisasikan pada web yang telah dibuat.

Kata kunci— kualitas udara, oksigen, *support vector machine*, klasifikasi, polinomial

I. PENDAHULUAN

Oksigen adalah salah satu kebutuhan primer manusia untuk keberlangsungan hidup. Polusi udara sangat mempengaruhi kesehatan manusia, seperti polusi yang ada pada konsentrasi PM₁₀ dan PM_{2.5} [1]. Menurut data dari IQAir, pada bulan Oktober 2021, Ibukota Indonesia, Jakarta, menduduki peringkat ke-9 untuk kualitas udara dan polusi kota. Dan juga Indonesia menempati peringkat 9 dari 106 negara untuk negara paling berpolusi di dunia pada tahun 2020 dari konsentrasi PM_{2.5}. Maka dari itu, dibuatlah penelitian tentang klasifikasi prediksi kualitas udara [2].

Pada penelitian sebelumnya, sudah dilakukan prediksi kualitas udara menggunakan algoritma Long Short Term Memory (LSTM) namun pada penelitian sebelumnya belum ada melakukan klasifikasi dengan menggunakan algoritma Support Vector Machine (SVM). Algoritma Long Short Term Memory merupakan jenis arsitektur dari Recurrent Neural Network (RNN) yang biasa digunakan untuk menyelesaikan permasalahan dengan menggunakan *deep learning* [3]. *Support Vector Machine* merupakan proses pembelajaran terbimbing (*supervised learning*), *support vector machine* dikembangkan agar mampu bekerja pada masalah non linear dengan memasukkan kernel [4].

Walaupun penelitian ini berfokus pada klasifikasi prediksi kualitas udara juga bisa menggunakan solusi

alternatif, seperti deep learning. Deep learning adalah jenis metode pembelajaran mesin yang sangat banyak peminatnya dan juga banyak dipakai pada industri. Deep learning digunakan untuk membangun kerangka prediksi spatio temporal, yang mempertimbangkan hubungan temporal dan spasial dari dataset dalam proses pemodelan [1].

II. KAJIAN TEORI

A. Polusi Udara

Salah satu kebutuhan primer manusia adalah udara. Seiring perkembangan teknologi dan industri yang menghasilkan beberapa pencemaran udara, kualitas udara juga menurun [1]. Menurut Peraturan Pemerintah Nomor 41 Tahun 1999, pencemaran udara dapat didefinisikan sebagai masuknya zat, energi, atau komponen lain ke udara yang disebabkan oleh kegiatan manusia, sehingga kualitas udara turun sampai pada tingkat tertentu yang menyebabkan udara tidak dapat berfungsi lagi. untuk memenuhi fungsinya [5]. Faktor penyebab peningkatan pencemaran udara disebabkan oleh beberapa polutan. Polutan tersebut antara lain partikulat (PM₁₀), sulfur dioksida (SO₂), karbon monoksida (CO), ozon (O₃), dan nitrogen dioksida (NO₂) [6].

B. Indeks Kualitas Udara

Dalam menentukan parameter dari kualitas udara maka instansi pemerintah membuat Indeks Kualitas Udara (AQI) hal ini dibuat pemerintah untuk memberitahukan kepada masyarakat tentang kualitas udara serta memberitahu pencemaran yang akan terjadi. Pencemaran menurut AQI dibagi menjadi enam kategori dan dari enam kategori ini menunjukkan peningkatan bahaya serta dampaknya bagi Kesehatan. Kemudian AQI juga memiliki lima kriteria polutan yang diatur dalam clean air act yaitu tingkat ozon permukaan (O₃), partikel debu (PM_{2.5}, dan PM₁₀), karbon monoksida (CO), sulfur dioksida (SO₂), dan nitrogen dioksida (NO₂). Nilai yang melebihi angka

300 tergolong sebagai kualitas udara yang berbahaya dan kualitas udara yang memiliki nilai dibawah 20 tergolong kualitas udara yang baik [2]

TABEL 1
(KATEGORI DAN RENTANG ISPU)

| Nilai AQI | Level kesehatan | Warna |
|-----------|--------------------|--------|
| 0 - 50 | Baik | Hijau |
| 51 - 100 | Sedang | Kuning |
| 151 - 200 | Tidak sehat | Merah |
| 201 - 300 | Sangat tidak sehat | Ungu |

TABEL 2
(TABEL KONVERSI NILAI KONSENTRASI PARAMETER ISPU)

| ISPU | 24 Jam partikulat (PM ₁₀) µg/m ₃ | 24 Jam sulfur dioksida (SO ₂) µg/m ₃ | 24 Jam karbon monoksida (CO) µg/m ₃ | 24 Jam ozon (O ₃) µg/m ₃ | 24 jam nitrogen dioksida (NO ₂) µg/m ₃ |
|-----------|--|--|---|--|--|
| 0 – 50 | 50 | 52 | 4000 | 120 | 80 |
| 51-100 | 150 | 180 | 8000 | 235 | 200 |
| 151 - 200 | 350 | 400 | 15000 | 400 | 1130 |
| 201 - 300 | 420 | 800 | 30000 | 800 | 2260 |

Keterangan:
 • Data pengukuran selama 24 jam secara terus-menerus.
 • Hasil perhitungan ISPU parameter partikulat (PM₁₀) disampaikan tiap jam selama 24 jam
 • Hasil perhitungan ISPU parameter partikulat (PM₁₀), sulfur dioksida (SO₂), karbon monoksida (CO), ozon (O₃), nitrogen dioksida (NO₂), diambil nilai ISPU parameter tertinggi dan paling sedikit disampaikan setiap jam 09.00 dan jam 15.00.

Perhitungan ISPU dilakukan berdasarkan nilai ISPU batas atas, ISPU batas bawah, ambien batas atas, ambien batas bawah, dan konsentrasi ambien hasil pengukuran. Berikut adalah rumus yang digunakan untuk mendapatkan nilai ISPU:

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

Keterangan:

- I = ISPU terhitung
 I_a = ISPU batas atas
 I_b = ISPU batas bawah
 X_a = Konsentrasi ambien batas atas (µg/m³)
 X_b = Konsentrasi ambien batas bawah (µg/m³)
 X_x = Konsentrasi ambien hasil pengukuran (µg/m³)

C. Klasifikasi

Klasifikasi merupakan sebuah pengujian yang kemudian dieksekusi dari sebuah Training lalu hasil dari pengujian akan didapatkan model terbaik dan dapat mengklasifikasikan dengan akurat [7]. . Kelayakan penerapan SVM dipelajari dengan membandingkan kinerja tiga kernel. Algoritma SVM tidak mempertimbangkan apakah data seimbang atau tidak. Namun, masalah ketidakseimbangan data merupakan masalah yang sulit dalam membangun model klasifikasi yang akurat. Parameter kernel memiliki dampak signifikan terhadap kompleksitas dan efisiensi model klasifikasi. Oleh karena itu, pemilihan model dalam SVM menyertakan parameter kernel. Oleh karena itu, algoritma pembelajaran mesin klasifikasi sangat penting dalam pengambilan keputusan untuk mencapai hasil klasifikasi yang baik [8]

D. Support Vector Machine

Boser, Guyon, dan Vapnik adalah orang-orang pertama yang memperkenalkan Support Vector Machine (SVM) pada tahun 1992 di *annual workshop on computational learning theory*. Konsep yang ditawarkan pada metode SVM

merupakan gabungan atau kombinasi dari teori-teori komputasi yang telah ada pada tahun sebelumnya [6].

Berikut merupakan fungsi kernel polinomial [9]

E. Synthetic Minority Oversampling Technique

Synthetic Minority Oversampling Technique (SMOTE) merupakan metode oversampling kelas membuat data buatan (sintetis) untuk memecahkan masalah ketidakseimbangan kelas data [10]. . SMOTE adalah teknik oversampling yang digunakan untuk menghindari penurunan kinerja pengklasifikasi yang disebabkan oleh ketidakseimbangan kelas dalam kumpulan data [11].

F. Aplikasi Website

Aplikasi web adalah salah satu media dengan tingkat penyebaran yang sangat mudah, dikarenakan sangat mudah untuk diakses. Meyebarkan aplikasi web diatur dari komponen sisi server pada jaringan. Pengguna aplikasi web dapat dengan mudah menyebarkan informasi tanpa perlu menggunakan aplikasi lain yang terpenting pengguna mengakses internet. Arsitektur dasar aplikasi web diantaranya adalah browser, koneksi internet, dan web server. Browser meminta "halaman web" dari server. Kombinasi konten dan intruksi pemformatan, dinyatakan dengan HTML. Bahasa adalah Bahasa markup populer yang digunakan oleh pengembang. Untuk kemudahan penggunaannya, HTML memiliki fungsi memutuskan format teks, membuat daftar, membuat tautan, menyisipkan dokumen, gambar, dan lainnya [12]

G. Pengujian Confusion Matrix

Metode confusion matrix digunakan untuk mengevaluasi hasil klasifikasi. Metode confusion matrix digunakan untuk memperkirakan target yang benar dan tidak benar. Dalam penerapan metode confusion matrix dilakukan perhitungan terhadap tiga parameter yaitu recall, precision dan f1-score. Precision merupakan klasifikasi true positive atau positif yang benar serta data keseluruhan yang diperkirakan kelas positif [9]. Berikut persamaan untuk memperoleh precision.

$$P = \frac{\text{true positive}}{\text{true positive} + \text{false positive}} = \frac{A}{A + C}$$

Recall adalah jumlah data yang memiliki klasifikasi *true positive* atau positif yang benar dari seluruh data yang positif yang dapat diperoleh dari persamaan berikut.

$$R = \frac{\text{true positive}}{\text{true positive} + \text{false negative}} = \frac{A}{A + B}$$

F1-score merupakan rata-rata dari nilai precision dan recall. *F1-score* dapat dihasilkan dari persamaan berikut..

$$F1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

G-mean dapat dihasilkan dari persamaan berikut.

$$G - \text{mean} = \sqrt{TP - TN}$$

Akurasi dapat dihasilkan dari persamaan berikut.

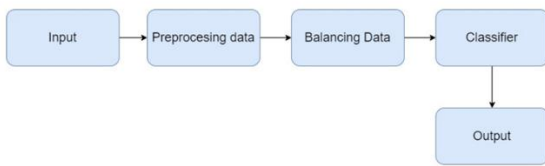
$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

III. METODE

A. Gambaran Umum sistem

Sistem klasifikasi prediksi Indeks Standar Pencemaran Udara (ISPU) dengan menggunakan metode *Support Vector Machine* (SVM) memiliki gambaran umum dimana terdapat

Langkah-langkah untuk mendapatkan hasil akhir. Untuk lebih jelasnya dapat dilihat diagram proses di bawah ini.



GAMBAR 1
(Diagram proses sistem)

B. Perancangan Sistem

terdapat metode dan Pustaka yang digunakan. Metode dan Pustaka ini membangun sistem klasifikasi yang memiliki komponen sebagai berikut. Data kualitas udara terdiri dari 5 daerah DKI Jakarta yaitu DKI1, DKI2, DKI3, DKI4, dan DKI5. Setiap daerah himpunan data memiliki 1826 data yang terhitung selama 4 tahun dari 1 Januari 2017 sampai dengan 31 Desember 2021. Zat yang terkandung dalam ke lima stasiun terdiri dari PM₁₀, SO₂, CO, O₃, dan NO₂. Berikut adalah tabel data dari ke lima stasiun.

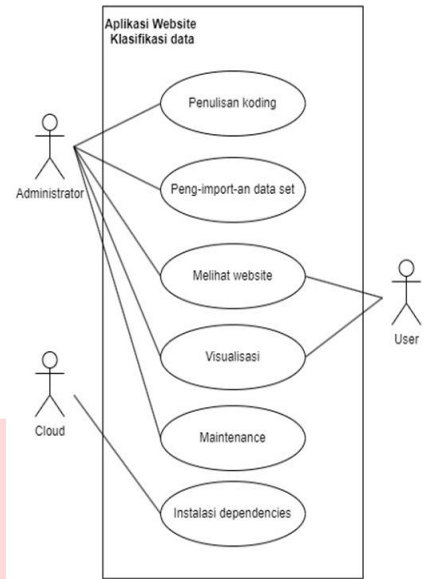
TABEL 3

| no | Nama stasiun | Kategori Kelas | | | |
|----|---------------|----------------|--------|-------------|--------------------|
| | | Baik | Sedang | Tidak Sehat | Sangat Tidak Sehat |
| 1 | Bundaran HI | 523 | 1142 | 65 | 0 |
| 2 | Kelapa Gading | 282 | 1192 | 266 | 11 |
| 3 | Jagakarsa | 226 | 1193 | 313 | 0 |
| 4 | Lubang Buaya | 226 | 1232 | 229 | 0 |
| 5 | Kebon Jeruk | 282 | 1015 | 443 | 30 |

(DETAIL DATA KELAS DARI KELIMA STASIUN).

C. Perancangan Perangkat Lunak Website

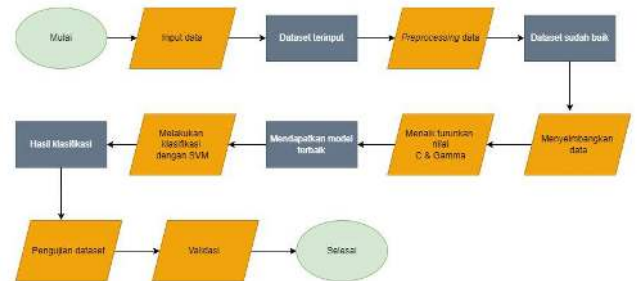
Perancangan website menggunakan aplikasi open source framework yang menggunakan bahasa Python, yaitu streamlit. Streamlit berguna untuk pembuatan aplikasi web untuk data science dan *machine learning* yang mudah dipelajari. Streamlit juga kompatibel dengan library-library yang ada pada Python. Dalam aplikasi web ini terdapat halaman utama dan halaman klasifikasi, prediksi, dan forecasting data. Berikut adalah diagram use case untuk aplikasi situs web ini.



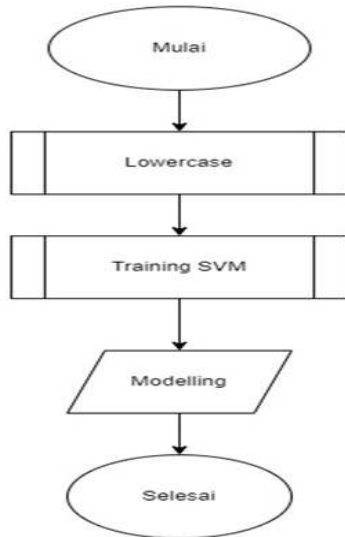
GAMBAR 2
(DIAGRAM USE CASE SITUS WEB)

D. Perancangan Metode Support Vector Machine

Perancangan program klasifikasi prediksi ISPU dimulai dari peng-inputan data, preprocessing data, membuat data seimbang, melakukan pengklasifikasian, hingga dilakukan pengujian.



GAMBAR 3
(PERANCANGAN METODE SUPPORT VECTOR MACHINE)



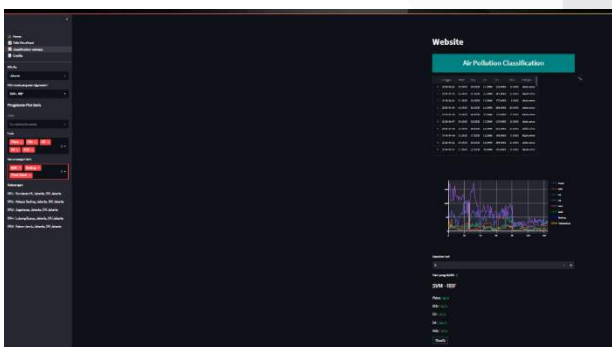
GAMBAR 4 (FLOWCHART SVM)

IV. HASIL DAN PEMBAHASAN

Implementasi klasifikasi prediksi kualitas udara ini meliputi implementasi pengujian kemudian juga implementasi klasifikasi dari kinerja algoritma yang dipakai. Untuk melakukan klasifikasi diperlukan beberapa tools, Tools yang dipakai menggunakan bahasa python lalu versi python yang digunakan yaitu anaconda. Python anaconda merupakan sebuah bahasa pemrograman untuk mesin pembelajaran yang bisa melakukan penulisan koding dengan menggunakan jupyter notebook.

Algoritma yang digunakan dalam proses klasifikasi yaitu *support vector machine* yang diperkenalkan oleh Vapnik. Kemudian untuk antarmuka web menggunakan framework streamlit.

Untuk implementasi aplikasi *website* tampilan halaman klasifikasi udara, terdapat fitur untuk memilih hari yang akan di klasifikasikan kualitas udaranya.



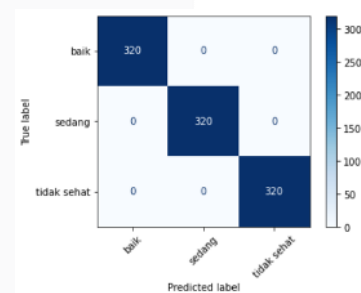
GAMBAR 5 (HALAMAN KLASIFIKASI DATA)

Dalam pengujian *confussion matrix*, akan dilakukan beberapa pengujian terhadap hasil eksekusi data uji dari proses klasifikasi untuk DKI1 sampai DKI5 yang masing-masing memiliki 5 zat, yaitu PM10, SO2, CO, O3, dan NO2. Hal-hal yang diperhatikan dalam pengujian ini yaitu pengujian dari 3 kernel dan pembagian partisi data tes dan train. 3 kernel yang digunakan yaitu polinomial, Radial Basis function (RBF), dan sigmoid. Untuk pembagian partisi data dibagi 80% untuk training dan 20% untuk tes. Setelah melakukan proses klasifikasi, sistem akan melakukan proses evaluasi. Evaluasi yang dilakukan yaitu G-mean, Specificity, dan Sensitivity serta akurasi. Proses klasifikasi yang memiliki hasil evaluasi yang paling bagus akan dipilih

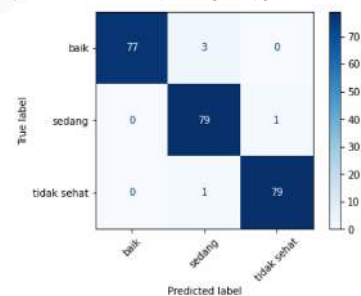
| Train/Test | G-mean | Specificity | Sensitivity | Kernel |
|------------|--------|-------------|-------------|------------|
| 60/40 | 0.97 | 0.98 | 0.97 | Polinomial |
| 70/30 | 0.98 | 0.99 | 0.98 | |
| 80/20 | 0.97 | 0.98 | 0.97 | |
| 90/100 | 0.98 | 0.99 | 0.98 | |
| Train/Test | G-mean | Specificity | Sensitivity | RBF |
| 60/40 | 0.0 | 0.66 | 0.33 | |
| 70/30 | 0.0 | 0.66 | 0.33 | |
| 80/20 | 0.11 | 0.68 | 0.36 | |
| 90/100 | 0.07 | 0.67 | 0.34 | |
| Train/Test | G-mean | Specificity | Sensitivity | Sigmoid |
| 60/40 | 0.0 | 0.66 | 0.33 | |
| 70/30 | 0.0 | 0.66 | 0.33 | |
| 80/20 | 0.0 | 0.66 | 0.33 | |
| 90/100 | 0.0 | 0.66 | 0.33 | |

sebagai model terbaik untuk diimplementasikan ke website.

TABEL 4 (HASIL EVALUASI IMBALANCED DATA(TANPA SMOTE))



GAMBAR 6 (CONFUSION MATRIX KERNEL POLINOMIAL TRAINING 80%)

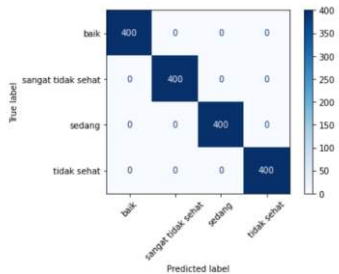


GAMBAR 7 (CONFUSION MATRIX KERNEL POLINOMIAL TESTING 20%)

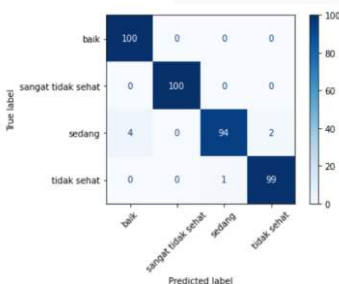
A. Pengujian

TABEL 5
(HASIL EVALUASI BALANCED DATA(DENGAN SMOTE))

| Train/Test | G-mean | Specificicity | Sensitivity | kernel |
|------------|--------|---------------|-------------|------------|
| 60/40 | 0.96 | 0.98 | 0.96 | Polinomial |
| 70/30 | 0.96 | 0.99 | 0.97 | |
| 80/20 | 0.98 | 0.99 | 0.98 | |
| 90/100 | 0.96 | 0.99 | 0.97 | |
| Train/Test | G-mean | Specificicity | Sensitivity | RBF |
| 60/40 | 0.0 | 0.76 | 0.30 | |
| 70/30 | 0.06 | 0.76 | 0.28 | |
| 80/20 | 0.05 | 0.76 | 0.28 | |
| 90/100 | 0.0 | 0.76 | 0.30 | |
| Train/Test | G-mean | Specificicity | Sensitivity | Sigmoid |
| 60/40 | 0.0 | 0.75 | 0.25 | |
| 70/30 | 0.0 | 0.75 | 0.25 | |
| 80/20 | 0.0 | 0.75 | 0.25 | |
| 90/100 | 0.0 | 0.75 | 0.25 | |



GAMBAR 8
(CONFUSION MATRIX KERNEL POLINOMIAL TRAINING 80%)

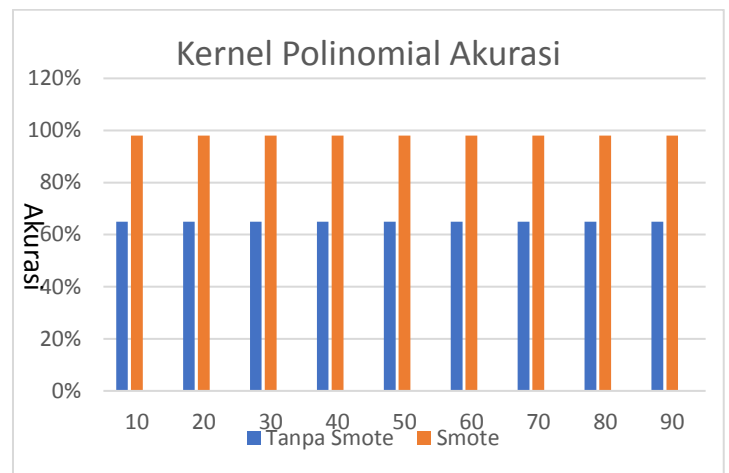


GAMBAR 9
(CONFUSION MATRIX KERNEL POLINOMIAL TESTING 20%)

B. Pengujian Model

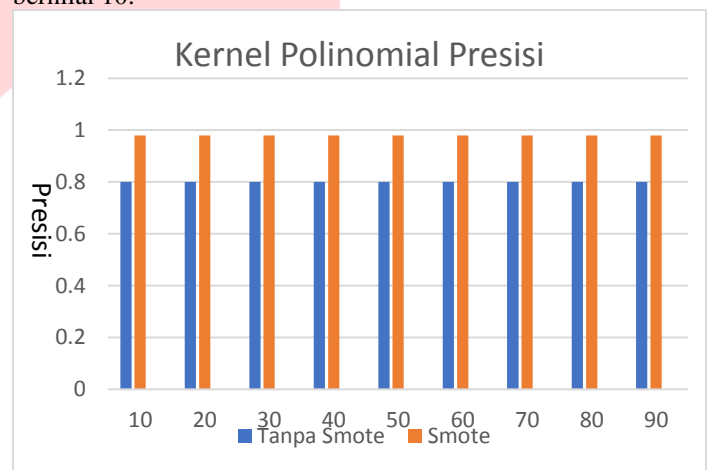
Pada pengujian ini nilai parameter yang akan dilihat yaitu presisi, recall, f1-score dan akurasi, Untuk nilai gamma sebesar 10 hingga 90 dan C bernilai 10.

A. Kernel polinomial



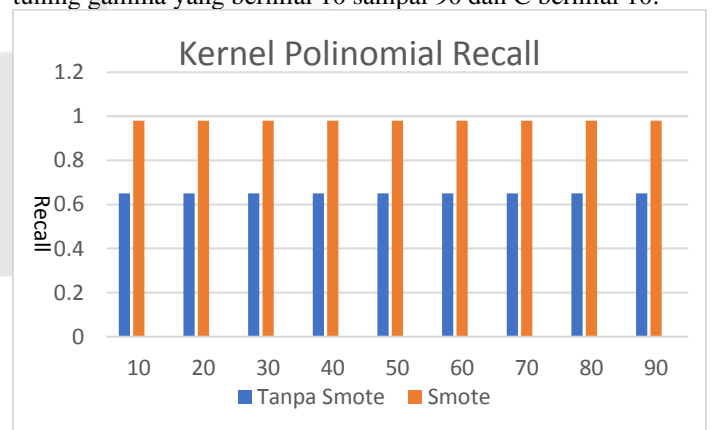
GAMBAR 10
(TEHADAP GAMMA,C1)

Gambar 10, menjabarkan pengujian akurasi dengan parameter tuning gamma yang bernilai 10 sampai 90 dan C bernilai 10.



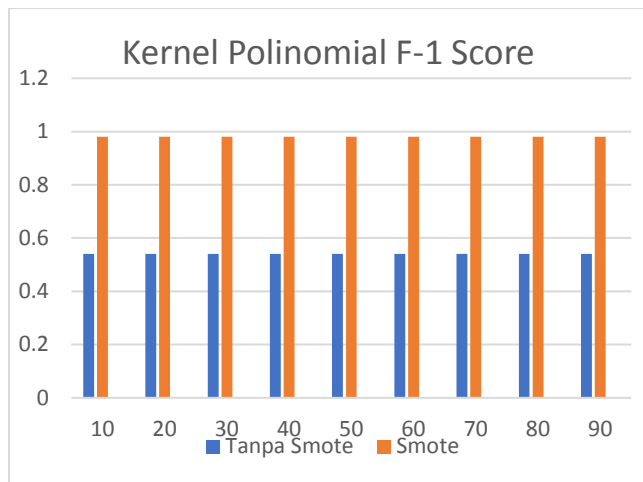
GAMBAR 11
(TERHADAP GAMMA,C1)

Gambar 11, menjabarkan pengujian presisi dengan parameter tuning gamma yang bernilai 10 sampai 90 dan C bernilai 10.



GAMBAR 12
(TERHADAP GAMMA, C1)

Gambar 12, menjabarkan pengujian recall dengan parameter tuning gamma yang bernilai 10 sampai 90 dan C bernilai 10.



GAMBAR 13 (TERHADAP GAMMA, C1)

C. Kuesioner Aplikasi

Kuesioner aplikasi dimaksudkan untuk mengetahui apakah *user* mendapatkan informasi yang sesuai tentang kualitas udara. Untuk hasil kuesioner aplikasi dapat dilihat pada TABEL 6

TABEL 6 (TABEL KUESIONER EDUKASI)

| No | Daftar Pertanyaan | Penilaian | | | | |
|----|---|---------------------|--------------|--------|--------|---------------|
| | | Sangat Tidak Setuju | Tidak Setuju | Netral | Setuju | Sangat Setuju |
| 1 | Situs web memiliki tampilan yang menarik. | 0% | 2.5% | 22.5% | 47.5% | 27.5% |
| 2 | Situs web ini memiliki tampilan yang mudah digunakan. | 0% | 2.5% | 15% | 52.5% | 30% |
| 3 | Semua halaman situs web berfungsi dengan baik. | 0% | 2.5% | 10% | 37.5% | 50% |
| 4 | Apakah situs web ini sudah memberikan informasi dengan baik. Tentang Fasilitas dan Gedung yang ada di Telkom University | 0% | 0% | 20% | 40% | 40% |
| 5 | kinerja situs web ini dalam mengklasifikasikan kualitas udara sudah baik. | 0% | 2.5% | 7.5% | 37.5% | 52.5% |
| 6 | Situs web mudah dibaca dan dipahami. | 0% | 2.5% | 17.5% | 35% | 45% |
| 7 | Fitur yang ditawarkan pada situs web memuaskan | 0% | 2.5% | 17.5% | 40% | 40% |
| 8 | Fitur di situs web berfungsi dengan baik. | 0% | 2.5% | 7.5% | 45% | 45% |

| No | Daftar Pertanyaan | Penilaian | | | | |
|-----------|--|---------------------|--------------|--------|--------|---------------|
| | | Sangat Tidak Setuju | Tidak Setuju | Netral | Setuju | Sangat Setuju |
| 9 | Akses situs web berfungsi dengan lancar. | 0% | 0% | 7.5% | 45% | 47.5% |
| Rata-rata | | 0% | 2% | 14% | 42% | 42% |

Pada tabel hasil kuesioner di atas, 82% responden mendapatkan informasi yang dibutuhkan tentang kualitas udara.

V. KESIMPULAN

Berdasarkan hasil penelitian, pengujian, serta analisa yang telah dilakukan, maka dapat disimpulkan bahwa:

- B. Data yang tidak seimbang akan sangat berpengaruh terhadap klasifikasi seperti data ISPU DKI. Jakarta, sehingga klasifikasi tidak akurat. Dengan metode SMOTE membuat data seimbang sehingga dapat melakukan klasifikasi data dengan tingkat akurat yang tinggi.
- C. Sistem website untuk melakukan klasifikasi memiliki uji alfa atau uji fungsionalitas sebesar 100% dan uji beta sebesar 98.2%.
- D. Pengujian kinerja algoritma pada klasifikasi memiliki akurasi pada kernel polinomial yang terbaik sebesar 98% dan untuk nilai G-mean 0.98, Specificity 0.99, Sensitivity 0.98, Presisi 0.98, Recall 0.98, F1 Score 0.98.
- E. Parameter zat yang digunakan untuk klasifikasi yaitu 5 zat, PM10, SO2, CO, O3, dan NO2. Parameter zat tersebut sangat menentukan untuk hasil klasifikasi yang maksimal.

REFERENSI

- [1] H. L. THANONGSAK XAYASOUK, "Air pollution prediction system using deep learning," *WIT Transactions on Ecology and the Environment*, vol. 230, no. -, p. 9, 2018.
- [2] M. L. H. RI, "peraturan menteri lingkungan hidup dan kehutanan," - June 2020. [Online]. Available: https://ditppu.menlhk.go.id/portal/uploads/laporan/1601040067_P_14_2020_ISPU_menlhk.pdf. [Accessed - - 2021].
- [3] J. Kim, "An Air Pollution Prediction Scheme Using Long Short Term Memory Neural Network Model," in *E3S Web of Conferences*, Shenyang, 2021.
- [4] T. F. Abidin, "Metode Klasifikasi Metode Klasifikasi ((SVM Light dan K SVM Light dan K--NN NN))," Taufik Fuadi Abidin, Banda Aceh, 2013.
- [5] J. Government, "Indeks Kualitas Udara (AQI) Jakarta dan Polusi Udara Indonesia," Government, Jakarta, - - 2021. [Online]. Available: <https://www.iqair.com/id/indonesia/jakarta>. [Accessed 18 December 2022].

- [6] S. M. H. Syamsul Bahri, "View of MODEL WAVELET NEURAL NETWORK UNTUK PEMODELAN POLUSI UDARA DI KOTA MATARAM LOMBOK NUSA TENGGARA BARAT INDONESIA," in *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya*, Samarinda, 2019.
- [7] L. A. Demidova, "Two-stage hybrid data classifiers based on svm and knn algorithms," *symmetry*, vol. 13, no. 4, p. 32, 2021.
- [8] S. S. T. E. A. N. L. H. Ade Silvia Handayani, "Air Quality Classification Using Support Vector Machine," *Computer Engineering and Applications*, vol. 10, no. 1, p. 16, 2021.
- [9] R. N. CHORY, "ANALISIS SENTIMEN PADA TINGKAT KEPUASAN," *Telkom University*, Vols. -, no. -, p. 8, 2019.
- [10] B. L. H. P. K. N V. Chawla, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, no. -, p. 37, 2002.
- [11] N. G. Ramadhan, "Comparative Analysis of ADASYN-SVM and SMOTE-SVM Methods on the Detection of Type 2 Diabetes Mellitus," *Scientific Journal of Informatics*, Vols. -, no. -, p. 8, 2021.
- [12] U. N.A, "Prediksi Jumlah Penumpang dan Penambahan Gerbong Kereta Api Menggunakan Metode Support Vector Regression Prediction of Total Passengers and Addition of Railway Train Using Support Vector Regression (SVR) Method," *Telkom University*, Vols. -, no. -, p. 8, 2020.