



PREDIKSI KUALITAS UDARA KOTA PONTIANAK MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR (KNN)

Utin Siti Nur Fadillah, Syarifah Putri Agustini A, Barry Ceasar Octariadi Mail

Prodi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer,

Universitas Muhammadiyah Pontianak

Abstrak

Kualitas udara merupakan indikator penting dalam menilai kesehatan lingkungan suatu wilayah. Kota Pontianak, sebagai pusat aktivitas ekonomi dan transportasi di Kalimantan Barat, menghadapi potensi penurunan kualitas udara akibat meningkatnya aktivitas manusia dan kendaraan. Penelitian ini bertujuan membangun sistem prediksi kualitas udara di Kota Pontianak dengan menerapkan algoritma K-Nearest Neighbor (KNN) berdasarkan parameter polusi udara meliputi PM_{2.5}, PM₁₀, CO, O₃, NO₂, dan parameter cuaca (suhu, kelembaban). Data diperoleh dari Dinas Lingkungan Hidup Kota Pontianak dan melalui tahapan preprocessing sebelum digunakan sebagai data pelatihan dan pengujian. Sistem prediksi dirancang dalam bentuk aplikasi web yang memungkinkan pengguna mengunggah data dan memperoleh hasil prediksi kategori kualitas udara, yaitu baik, sedang, atau tidak sehat. Hasil evaluasi menunjukkan bahwa model KNN dengan nilai K=7 menghasilkan akurasi sebesar 90%, presisi 92% untuk kelas "Tidak Sehat", dan recall 88%. Sistem ini diharapkan dapat membantu instansi terkait dalam mengambil langkah preventif terhadap potensi penurunan kualitas udara serta meningkatkan kesadaran masyarakat mengenai pentingnya menjaga kondisi udara yang bersih dan sehat.

Kata Kunci: Kualitas Udara, Prediksi, K-Nearest Neighbor, Pontianak, Aplikasi Web.

PENDAHULUAN

Kualitas udara merupakan salah satu indikator utama dalam menjaga kesehatan masyarakat dan kelestarian lingkungan. Kota Pontianak, sebagai ibu

kota Provinsi Kalimantan Barat, menghadapi tantangan serius terkait pencemaran udara, terutama pada musim kemarau. Data dari Dinas Lingkungan Hidup Kota Pontianak tahun 2023 mencatat bahwa nilai harian ISPU

(Indeks Standar Pencemar Udara) untuk parameter PM2.5 sering kali mencapai kategori "sedang" hingga "tidak sehat", bahkan sempat menembus angka $135 \mu\text{g}/\text{m}^3$ akibat kebakaran hutan dan lahan di sekitar Kalimantan Barat. Peningkatan kendaraan bermotor sebesar 7% per tahun turut menyumbang emisi karbon monoksida (CO), nitrogen dioksida (NO₂), dan partikel halus (Gupta et al., 2023).

Dalam upaya mengantisipasi dampak kesehatan dari polusi udara, pengembangan sistem prediksi kualitas udara menjadi hal yang sangat penting. Teknologi kecerdasan buatan, khususnya *machine learning*, telah digunakan secara luas untuk menganalisis data historis dan memprediksi konsentrasi polutan udara. Salah satu algoritma yang populer adalah *K-Nearest Neighbor (KNN)* karena kemampuannya dalam mengenali pola data dan menghasilkan prediksi yang cukup akurat dengan proses yang relatif cepat (Zhang & Zhang, 2020).

Penelitian ini bertujuan untuk memprediksi kualitas udara di Kota Pontianak menggunakan algoritma *K-Nearest Neighbor (KNN)*, dengan memanfaatkan data parameter polutan (PM2.5, PM10, CO, NO₂, O₃) dan data cuaca (suhu, kelembaban). Novelty penelitian ini terletak pada penerapan model prediksi yang dikhususkan untuk kondisi geografis dan iklim tropis basah Kota Pontianak dengan karakteristik polusi unik, serta integrasi data parameter polutan dan cuaca dalam satu model prediksi tunggal berbasis web

METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif dengan pendekatan eksperimen komputasional. Data yang digunakan terdiri dari 500 data harian periode Januari-Desember 2024 yang diperoleh dari Dinas Lingkungan Hidup Kota Pontianak, mencakup parameter polutan (PM2.5, PM10, CO, NO₂, O₃) dan

parameter cuaca (suhu, kelembaban). Label kategori kualitas udara ditentukan berdasarkan standar ISPU: Baik (0-50), Sedang (51-100), Tidak Sehat (101-200).

Tahapan penelitian mengikuti framework CRISP-DM (*Cross Industry Standard Process for Data Mining*), meliputi: 1) *Business Understanding*: memahami kebutuhan sistem prediksi sebagai peringatan dini; 2) *Data Understanding*: eksplorasi data; 3) *Data Preparation*: pembersihan data, penanganan *missing value* dengan interpolasi linier, penghapusan *outlier* menggunakan IQR, normalisasi data dengan Min-Max Scaling, dan pelabelan; 4) *Modeling*: penerapan algoritma KNN dengan pembagian data 80% pelatihan dan 20% pengujian; 5) *Evaluation*: evaluasi model menggunakan metrik akurasi, presisi, recall, dan F1-score; 6) *Deployment*: implementasi sistem berbasis web menggunakan PHP, MySQL, dan Bootstrap.

Algoritma KNN diimplementasikan dengan menghitung jarak Euclidean antara data uji dan data latih. Nilai K optimal ditentukan melalui eksperimen dengan nilai K=3, 5, 7, 9, dan 11. Sistem web dikembangkan dengan arsitektur client-server, fitur utama meliputi unggah dataset, input data prediksi, proses prediksi, dan tampilan hasil.

HASIL DAN PEMBAHASAN

Hasil eksperimen menunjukkan bahwa nilai K=7 menghasilkan akurasi tertinggi sebesar 90%. Hasil pengujian nilai K disajikan pada Tabel 1.

Tabel 1. Hasil Pengujian Nilai K Optimal

Nilai K	Akurasi (%)
3	86.4%
5	88.2%

7	90.0%
9	89.1%
11	87.5%

Sumber: Hasil Pengolahan Data (2025)

Model KNN dengan K=7 berhasil memprediksi 45 dari 50 data uji dengan benar. Hasil evaluasi model menunjukkan presisi untuk kelas "Tidak Sehat" sebesar 92%, recall 88%, dan F1-score 90%. Kesalahan prediksi terjadi pada 5 data yang berada pada kondisi ambang batas (*borderline*) antara kategori "Baik" dan "Sedang". Hal ini disebabkan oleh kesamaan nilai fitur pada rentang ISPU yang berdekatan.

Sistem berbasis web telah berhasil diimplementasikan dengan semua fitur berfungsi dengan baik, meliputi unggah dataset, proses prediksi, tampilan hasil, dan ekspor data. Antarmuka sistem dirancang responsif menggunakan Bootstrap 5. Gambar 1 menunjukkan halaman form input data untuk prediksi.

Gambar 1. Form Input Data Prediksi

Pembahasan mengacu pada penelitian terdahulu oleh Zhang & Zhang (2020) yang juga menyimpulkan efektivitas KNN dalam prediksi kualitas udara. Temuan penelitian ini sejalan dengan hasil Khaire & Sonawane (2021) bahwa KNN memiliki waktu komputasi cepat dan akurasi kompetitif, cocok untuk prediksi semi *real-time*. Namun, penelitian ini memiliki keunikan pada

spesifikasi lokasi Kota Pontianak dengan integrasi data cuaca lokal yang mempengaruhi dispersi polutan

SIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa algoritma K-Nearest Neighbor (KNN) berhasil diterapkan untuk memprediksi kualitas udara Kota Pontianak dengan akurasi 90%. Sistem prediksi berbasis web yang dikembangkan telah berfungsi dengan baik, memungkinkan pengguna mengunggah data dan mendapatkan hasil prediksi kategori kualitas udara secara cepat. Model menunjukkan kinerja terbaik dengan parameter K=7 dan mampu mendeteksi kondisi "Tidak Sehat" dengan presisi yang tinggi. Untuk pengembangan selanjutnya, disarankan untuk menambah parameter lain seperti SO₂ dan arah angin, memperbanyak volume data, serta mengintegrasikan sistem dengan data *real-time* dari stasiun pemantauan DLH.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada Dinas Lingkungan Hidup Kota Pontianak yang telah menyediakan data penelitian, serta kepada pembimbing skripsi yang telah memberikan bimbingan dan masukan berharga selama pelaksanaan penelitian ini.

DAFTAR PUSTAKA

- Amalia, A. dkk. (2023). Prediksi Kualitas Udara Menggunakan Algoritma K-Nearest Neighbor.
- Gupta, N.S., Mohta, Y., Heda, K., Armaan, R., Valarmathi, B., & Arulkumaran, G. (2023). Prediction of Air Quality Index Using Machine Learning Techniques: A Comparative Analysis. *J Environ Public Health*, 2023, 1-26.
- Khaire, V.M., & Sonawane, S.S. (2021). Comparative Study of Machine Learning Algorithms for Air Quality

Forecasting. *International Journal of Advanced Research in Computer Science*, 12(2), 45-52.

Zhang, Y., & Zhang, L. (2020). Air Quality Prediction Using Machine Learning-Based KNN Model. *Journal of Environmental Management*, 270, 110-123.