

BAB I

PENDAHULUAN

A. Latar Belakang

Kesehatan masyarakat merupakan bidang yang berfokus pada berbagai usaha untuk memelihara dan meningkatkan kesehatan seluruh masyarakat. Usaha ini dilakukan secara terencana dan kolaboratif oleh individu, komunitas, serta lembaga pemerintah dan swasta [1]. Peran kesehatan masyarakat sangat krusial bagi kemajuan suatu negara, karena masyarakat yang sehat menjadi fondasi dalam pembangunan yang berkelanjutan [2]. Oleh sebab itu, kondisi kesehatan masyarakat harus dipantau secara berkala untuk mencegah dan mengendalikan penyakit.

Salah satu penyakit menular yang masih menjadi permasalahan kesehatan utama di Indonesia adalah Infeksi Saluran Pernapasan Akut (ISPA). Penyakit ini menyerang saluran pernapasan bagian atas maupun bawah yang disebabkan oleh infeksi virus atau bakteri. Gejalanya dapat berupa batuk, demam, pilek, dan sesak napas, yang dalam kondisi tertentu dapat berkembang menjadi penyakit yang lebih serius bahkan dapat berakibat fatal jika tidak ditangani dengan cepat dan tepat [3].

Menurut laporan World Health Organization (WHO), setiap tahun terdapat sekitar 18,8 miliar kasus ISPA di seluruh dunia dengan angka kematian mencapai sekitar 4,25 juta jiwa. Hal ini menunjukkan bahwa ISPA masih menjadi salah satu penyebab utama penyakit menular serta penyebab kematian yang signifikan di dunia [4].

ISPA merupakan salah satu penyebab kematian terbesar di dunia, dengan kelompok paling rentan yaitu bayi, anak-anak, dan lansia [5]. Pada anak balita, penyakit ini berkontribusi sekitar 20–30% terhadap angka kematian, bahkan lebih dari 2 juta balita meninggal setiap tahun akibat penyakit ini [5] [6]. Tingginya risiko pada kelompok ini disebabkan oleh sistem kekebalan tubuh yang belum berkembang secara optimal [7]. Sementara itu, pada kelompok lansia, risiko ISPA meningkat seiring dengan menurunnya fungsi fisiologis tubuh akibat proses penuaan. WHO memperkirakan insiden ISPA pada negara berkembang mencapai 15–20% per tahun pada lansia, dengan angka kematian sekitar 2 juta jiwa per tahun. Di Indonesia, prevalensi ISPA tercatat sebesar 9,3%, dengan prevalensi tertinggi pada kelompok lansia yaitu 13,7% [8].

Kasus ISPA di Provinsi Kepulauan Riau menunjukkan adanya tren peningkatan yang memerlukan perhatian serius. Berdasarkan data dari Dinas Kesehatan Provinsi Kepri, pada minggu ke-5 tahun 2025 tercatat sebanyak 1.991 kunjungan ke fasilitas kesehatan akibat penyakit ISPA, menjadikannya sebagai penyakit dengan jumlah kasus tertinggi dibandingkan dengan penyakit lainnya. Menurut Kepala Dinas Kesehatan Provinsi Kepri, peningkatan ini dipengaruhi oleh faktor cuaca yang tidak menentu. Perubahan suhu ekstrem, seperti cuaca panas yang tiba-tiba disertai hujan, menyebabkan daya tahan tubuh masyarakat menurun sehingga lebih rentan terhadap serangan ISPA [9].

Kondisi serupa juga terjadi di Kota Tanjungpinang, Provinsi Kepulauan Riau. Dinas Kesehatan Kota Tanjungpinang melaporkan bahwa jumlah kasus ISPA telah mencapai lebih dari 20 ribu kasus sejak Januari hingga September 2025. Kepala Dinkes Tanjungpinang, menjelaskan bahwa lonjakan kasus terjadi pada bulan Februari dan kembali meningkat menjelang bulan September. Pola ini diduga kuat berkaitan dengan perubahan cuaca atau musim pancaroba [10]. Kondisi ini menunjukkan bahwa ISPA masih menjadi salah satu masalah kesehatan utama di Kota Tanjungpinang, sehingga memerlukan upaya penanganan yang lebih terukur dan berbasis data.

Beberapa penelitian terdahulu menunjukkan bahwa faktor iklim memiliki pengaruh yang signifikan terhadap kasus ISPA. Salah satu penelitian menganalisis hubungan antara suhu, curah hujan, kelembapan, dan kecepatan angin terhadap kasus ISPA menggunakan data hasil wawancara dan observasi yang diuji dengan metode *Chi-Square* dengan bantuan perangkat lunak SPSS untuk menguji hubungan antara variabel independen dan variabel dependen. Hasil analisis menunjukkan bahwa suhu, kelembapan, dan kecepatan angin memiliki hubungan yang sangat signifikan dengan nilai $p = 0,000$, sementara curah hujan juga terbukti berpengaruh secara signifikan dengan nilai $p = 0,016$ ($p < 0,05$) [11].

Selain itu, penelitian yang menelaah hubungan perubahan iklim terhadap ISPA melalui metode studi literatur menunjukkan bahwa suhu ekstrem, kelembapan tinggi, dan variasi curah hujan dapat menciptakan kondisi lingkungan yang mendukung pertumbuhan mikroorganisme penyebab ISPA serta meningkatkan risiko ISPA [12]. Penelitian lain dengan desain observasional

analitik pendekatan *cross-sectional* yang dianalisis menggunakan uji *Chi-Square* juga menemukan adanya hubungan signifikan antara suhu dan kelembapan dengan kejadian ISPA [13]. Hasil tersebut diperkuat oleh studi literatur lainnya yang menyimpulkan bahwa faktor kualitas fisik lingkungan, termasuk suhu dan kelembapan, berhubungan dengan kejadian ISPA [14].

Berdasarkan kontribusi faktor iklim terhadap peningkatan kasus ISPA, maka diperlukan suatu metode analisis yang dapat menghasilkan prediksi jumlah kasus ISPA. Hal ini sangat penting bagi Kota Tanjungpinang yang memiliki kondisi iklim yang tidak stabil serta angka kasus ISPA yang tergolong tinggi. Prediksi yang tepat dapat membantu fasilitas pelayanan kesehatan, terutama puskesmas, dalam mengantisipasi lonjakan kasus melalui kesiapan obat-obatan, logistik medis, dan sumber daya kesehatan lainnya [15]. Hubungan antara faktor iklim dan kasus ISPA bersifat kompleks dan non-linear, sehingga sulit dimodelkan secara optimal menggunakan metode statistik konvensional. Oleh karena itu, penerapan *machine learning* dinilai relevan karena mampu mengolah data dalam jumlah besar dan mengenali pola yang kompleks sehingga dapat menghasilkan prediksi yang baik terutama di bidang kesehatan [16].

Salah satu algoritma *machine learning* yang terbukti unggul dalam menangani data kompleks dan non-linear antarvariabel yaitu *Extreme Gradient Boosting* (XGBoost). Penelitian sebelumnya memprediksi fluktuasi berat badan berdasarkan variabel fisiologis dan gaya hidup. Hasil menunjukkan bahwa XGBoost mampu mengungguli model *Decision Tree* dan *Deep Learning* dengan nilai RMSE terendah yaitu 5,65. Keunggulan ini didasarkan pada kemampuan XGBoost dalam memproses data kompleks, non-linear, serta efektivitas mekanisme regularisasinya dalam menangani masalah *overfitting*. Hal tersebut menghasilkan stabilitas prediksi yang lebih baik dibandingkan model pohon keputusan konvensional, sekaligus menyediakan kerangka inovatif bagi pengembang teknologi kesehatan berbasis data [17].

Namun, performa optimal XGBoost sangat bergantung pada pemilihan *hyperparameter* yang tepat [18]. Oleh karena itu, diperlukan metode optimasi untuk mencari kombinasi parameter terbaik. Salah satu metode optimasi yang umum digunakan yaitu *Grid Search*, yaitu metode optimasi yang melakukan pencarian

secara sistematis dengan mengevaluasi seluruh kombinasi *hyperparameter* pada ruang pencarian yang telah ditentukan, sehingga berpotensi menemukan kombinasi terbaik, namun dengan konsekuensi waktu komputasi yang lebih tinggi [19]. Selain itu, terdapat metode *Random Search* yang melakukan pencarian dengan memilih sejumlah kombinasi *hyperparameter* secara acak dari ruang pencarian, sehingga proses optimasi dapat dilakukan dengan waktu komputasi yang lebih efisien tanpa menurunkan performa model secara signifikan [20].

Adapun penelitian sebelumnya telah melakukan prediksi jumlah kunjungan pasien ISPA di puskesmas Air Dingin untuk prediksi tahun 2025 menggunakan metode regresi linier sederhana berdasarkan data bulanan periode Januari–Desember 2024 sebanyak 12 data. Hasil penelitian menunjukkan tren peningkatan hingga 8.381 kunjungan, namun evaluasi model menghasilkan performa yang rendah dengan nilai RMSE sebesar 314,27 dan R^2 sebesar -9,684 [21]. Hal ini mengindikasikan bahwa metode statistik sederhana belum mampu menangani pola data yang fluktuatif dan non-linear, sehingga diperlukan pengembangan model yang lebih kompleks dan adaptif seperti *machine learning*.

Berdasarkan uraian tersebut, meskipun pemanfaatan *machine learning* dalam bidang kesehatan terus berkembang, penerapan model prediksi kasus ISPA dengan mempertimbangkan faktor iklim menggunakan data harian yang dioptimasi masih sangat terbatas, khususnya di Kota Tanjungpinang. Sebagian besar penelitian sebelumnya menggunakan data agregat tahunan atau bulanan sehingga belum mampu menangkap fluktuasi harian kasus ISPA yang dipengaruhi oleh dinamika iklim. Oleh karena itu, penelitian ini dilakukan untuk membangun model prediksi jumlah kasus ISPA berdasarkan faktor iklim harian di Kota Tanjungpinang menggunakan algoritma XGBoost dengan optimasi *hyperparameter* menggunakan metode *Grid Search* dan *Random Search*, yang selanjutnya dibandingkan untuk memperoleh model dengan performa terbaik.

Hasil penelitian ini diharapkan dapat memberikan model prediksi yang akurat dan efisien serta menjadi dasar pengambilan keputusan bagi pemerintah dan instansi kesehatan di Kota Tanjungpinang dalam melakukan upaya mitigasi, kesiapsiagaan logistik medis, dan pencegahan ISPA secara lebih efektif dan berbasis data.

B. Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah bagaimana performa algoritma XGBoost dengan optimasi *hyperparameter* menggunakan *Grid Search* dan *Random Search* dalam memprediksi jumlah kasus ISPA di Kota Tanjungpinang berdasarkan faktor iklim, serta metode optimasi manakah yang menghasilkan performa prediksi yang lebih baik?

C. Batasan Penelitian

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Dataset yang digunakan merupakan data harian yang diperoleh dari Dinas Kesehatan dan Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Tanjungpinang periode 1 Januari 2022–29 April 2026 dengan total sebanyak 1.580 data.
2. Variabel yang digunakan dalam penelitian ini meliputi faktor iklim, yaitu suhu, kelembapan, curah hujan, dan kecepatan angin, serta fitur lag jumlah kasus ISPA sebagai variabel prediktor, dengan jumlah kasus ISPA sebagai variabel target.
3. Pembagian dataset dilakukan dengan proporsi 80% sebagai data pelatihan (*training*), 10% sebagai data validasi (*validation*), dan 10% sebagai data pengujian (*testing*).
4. Output penelitian berupa prediksi jumlah kasus ISPA harian, yaitu satu hari ke depan di Kota Tanjungpinang dan tidak membahas mengenai aspek medis klinis penderita.

D. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, maka tujuan penelitian ini adalah untuk menganalisis performa algoritma XGBoost dengan optimasi *hyperparameter* menggunakan *Grid Search* dan *Random Search* dalam memprediksi jumlah kasus ISPA di Kota Tanjungpinang berdasarkan faktor iklim,

serta menentukan metode optimasi yang menghasilkan performa prediksi yang lebih baik.

E. Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat bagi pengembangan ilmu pengetahuan, pemerintah daerah, dan masyarakat. Adapun manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Bagi Pengembangan Ilmu Pengetahuan dan Teknologi (IPTEK)

Penelitian ini diharapkan dapat memperkaya kajian tentang pemodelan prediksi penyakit yang dipengaruhi oleh faktor iklim, khususnya melalui penerapan algoritma XGBoost dengan optimasi *hyperparameter*. Hasil dari penelitian ini dapat dijadikan referensi dan bahan pembanding bagi penelitian selanjutnya di bidang epidemiologi, kesehatan masyarakat, dan penerapan *machine learning* di bidang kesehatan.

2. Bagi Pemerintah

Hasil prediksi yang dihasilkan dalam penelitian ini diharapkan dapat menjadi bahan pendukung pengambilan keputusan bagi pemerintah daerah, khususnya Dinas Kesehatan Kota Tanjungpinang dalam perencanaan dan kesiapsiagaan penanganan ISPA. Informasi prediktif ini dapat dimanfaatkan sebagai indikator awal potensi peningkatan kasus ISPA, sehingga membantu perencanaan kebutuhan sumber daya kesehatan secara lebih optimal.

3. Bagi Masyarakat

Penelitian ini diharapkan memberikan manfaat bagi masyarakat melalui peningkatan kualitas pelayanan kesehatan. Dengan adanya hasil prediksi yang dimanfaatkan oleh pemerintah atau instansi kesehatan dalam perencanaan dan kesiapsiagaan sumber daya, masyarakat dapat memperoleh penanganan medis yang lebih cepat, tepat, dan optimal, terutama dalam menghadapi potensi peningkatan kasus ISPA.

F. Keaslian Penelitian

Penelitian ini memiliki perbedaan dengan penelitian sebelumnya, baik dari segi tujuan, metode, variabel, maupun data yang digunakan. Perbandingan penelitian terdahulu dengan penelitian yang diusulkan dapat dilihat pada Tabel 1.



Tabel 1. Keaslian penelitian

Aspek	Penelitian Terdahulu	Penelitian yang Diusulkan
Nama	Aulia et al. (2025)	Novitasari, Wulan (2026)
Judul	Model Prediksi Risiko Kesehatan Perkotaan Berbasis Lingkungan dengan XGBoost.	Prediksi Jumlah Kasus ISPA di Tanjungpinang Menggunakan Algoritma XGBoost dengan Optimasi <i>Hyperparameter</i>
Fokus Penelitian	Memprediksi risiko kesehatan masyarakat akibat kondisi lingkungan dan kualitas udara	Memprediksi jumlah kasus ISPA di Kota Tanjungpinang berdasarkan faktor iklim.
Metode	XGBoost yang dioptimasi dengan <i>RandomizedSearchCV</i> .	XGBoost dengan optimasi <i>Grid Search</i> dan <i>Random Search</i> .
Variabel	Menggunakan variabel lingkungan (suhu, kelembapan, <i>dew point</i> , <i>heat index</i> , UV, dll) dengan target <i>Health Risk Score</i> .	Variabel iklim (suhu, kelembapan, curah hujan, dan kecepatan angin), serta fitur rekayasa berupa lag, dengan target jumlah kasus ISPA harian.
Dataset	Dataset <i>Urban Air Quality and Health Impact</i> dataset publik (Kaggle) dari kota Amerika Serikat (1.000 data, 46 variabel).	Data harian ISPA dan iklim dari Dinas Kesehatan dan BMKG Kota Tanjungpinang.
Tujuan Penelitian	Mengembangkan model prediksi risiko kesehatan berbasis faktor lingkungan untuk mendukung sistem peringatan dini kesehatan masyarakat.	Menganalisis performa XGBoost dengan membandingkan dua metode optimasi untuk memprediksi jumlah kasus penyakit ISPA berdasarkan faktor iklim.