

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan era modern saat ini menuntut masyarakat untuk memiliki kemampuan literasi yang memadai. Kemampuan ini berperan penting dalam membantu seseorang mengakses informasi, memperoleh pekerjaan yang layak, maupun meningkatkan taraf hidup. Tingkat literasi penduduk dapat tercermin melalui angka buta huruf yang menunjukkan sejauh mana penduduk memiliki kemampuan membaca dan menulis. Menurut Badan Pusat Statistik (BPS) (2025a), angka buta huruf di Indonesia terus menurun dari waktu ke waktu. Meskipun demikian, pada tahun 2024 masih terdapat 3,33% atau sekitar 7.157,6 ribu jiwa penduduk berusia 15 tahun ke atas yang termasuk dalam kategori buta huruf. Kondisi tersebut menunjukkan bahwa masih perlunya upaya penurunan angka buta huruf guna mencapai salah satu tujuan keempat *Sustainable Development Goals* (SDGs), yakni menjamin bahwa pada tahun 2030 seluruh remaja dan proporsi kelompok dewasa tertentu, baik laki-laki maupun perempuan, memiliki kemampuan literasi dan numerisasi (Bappenas, 2025).

Angka buta huruf di Indonesia tidak terlepas dari pengaruh berbagai faktor, di antaranya faktor pendidikan dan kemiskinan. Faktor pendidikan dapat tercermin melalui rata-rata lama sekolah (RLS). BPS (2025b) mencatat bahwa angka RLS di Indonesia pada tahun 2024 mencapai 8,85 tahun, yang menunjukkan bahwa secara rata-rata penduduk berusia 25 tahun ke atas telah menyelesaikan pendidikan hingga setara kelas VIII SMP. Angka ini mengalami kenaikan sebesar 0,08 tahun jika dibandingkan

dengan capaian pada tahun 2023. Peningkatan RLS dapat berpotensi menurunkan angka buta huruf karena semakin lama seseorang menempuh pendidikan, semakin besar peluang mereka untuk menguasai kemampuan dasar membaca dan menulis. Maka dari itu, RLS berpotensi menjadi faktor yang dapat memengaruhi angka buta huruf.

Persentase penduduk miskin (P0) juga dapat memengaruhi angka buta huruf di Indonesia. Penduduk yang hidup dalam kemiskinan cenderung memiliki akses terbatas terhadap pendidikan, sehingga berisiko lebih tinggi mengalami buta huruf. BPS (2025b) mencatat bahwa pada Maret 2024 persentase penduduk miskin di Indonesia sebesar 9,03% atau sekitar 25.219 ribu jiwa, menurun sebesar 0,33% dibandingkan dengan Maret 2023. Meskipun mengalami penurunan, angka tersebut masih menjadi permasalahan mendesak karena berpotensi dalam meningkatkan angka buta huruf di Indonesia.

Sejumlah studi terdahulu telah meneliti mengenai faktor-faktor yang memengaruhi angka buta huruf. Astuti *et al.* (2017) melakukan penelitian dengan menggunakan *Geographically Weighted t Regression* (GWtR) untuk menganalisis faktor-faktor yang memengaruhi angka buta huruf di Provinsi Jawa Timur tahun 2015. Hasil penelitian tersebut menunjukkan bahwa angka partisipasi murni tingkat SD, rasio murid terhadap guru, tingkat pengangguran terbuka, serta persentase penduduk miskin memiliki pengaruh signifikan dengan nilai koefisien determinasi (R^2) sebesar 76,42%. Sementara itu, penelitian yang dilakukan oleh Sanusi *et al.* (2023) dengan pendekatan *Geographically Weighted Regression* (GWR) di Provinsi Sulawesi Selatan tahun 2021 menunjukkan bahwa persentase penduduk miskin, angka partisipasi sekolah, rata-rata

lama sekolah, jumlah tenaga pendidik, dan persentase penduduk yang memiliki telepon seluler berpengaruh signifikan terhadap angka buta huruf dengan nilai *adjusted R²* sebesar 86,1%. Perbedaan metode dan variabel yang digunakan dalam penelitian-penelitian sebelumnya menunjukkan perlunya penerapan pendekatan yang lebih adaptif dalam mengidentifikasi faktor-faktor yang memengaruhi angka buta huruf.

Pendekatan statistik yang dapat diterapkan untuk mengetahui faktor-faktor yang memengaruhi angka buta huruf adalah analisis regresi. Menurut Dani *et al.* (2021), estimasi kurva regresi dapat dilakukan melalui tiga pendekatan, yaitu regresi parametrik, semiparametrik, dan nonparametrik. Pendekatan parametrik digunakan ketika bentuk kurva regresi telah diketahui, semiparametrik digunakan ketika sebagian bentuk kurva regresi diketahui dan sebagian lainnya tidak diketahui, sementara itu nonparametrik digunakan ketika pola kurva regresi tidak diketahui. Pola hubungan antara rata-rata lama sekolah dan persentase penduduk miskin terhadap angka buta huruf menunjukkan sebaran data yang acak dan mengelompok di suatu titik tertentu, sehingga lebih tepat dianalisis menggunakan regresi nonparametrik.

Pendekatan regresi nonparametrik memiliki beragam metode, salah satunya adalah regresi polinomial lokal yang efektif untuk menangani data dengan pola data yang mengelompok di suatu titik tertentu (Suparti & Santoso, 2023). Menurut Fan dan Gijbels (1996), regresi polinomial lokal merupakan pemodelan regresi nonparametrik yang didasarkan pada pengembangan deret Taylor dan menggunakan fungsi kernel sebagai pembobotnya. Aspek penting dalam melakukan estimasi regresi polinomial lokal mencakup titik lokal, *bandwidth*, derajat polinomial lokal, dan fungsi kernel. Pada implementasinya, fungsi kernel Gaussian lebih umum digunakan karena mampu

menghasilkan estimasi yang lebih halus (*smooth*). Penentuan model terbaik dapat dilakukan dengan menggunakan sejumlah metode optimasi, seperti nilai *Mean Squared Error* (MSE), *Cross Validation* (CV), dan *Generalized Cross Validation* (GCV).

Penerapan regresi nonparametrik polinomial lokal telah dilakukan oleh sejumlah peneliti terdahulu. Rory dan Diana (2020) dengan regresi polinomial lokal pada data jumlah kasus baru harian Covid-19 di Jakarta menunjukkan bahwa model polinomial lokal derajat 1 dengan kernel Gaussian memberikan hasil terbaik dengan nilai R^2 sebesar 83,58%. Sementara itu, Oktaviani *et al.* (2020) membandingkan metode analisis regresi linier dengan metode regresi polinomial lokal dalam memodelkan hubungan antara berat badan balita dengan umur. Hasil penelitian tersebut menunjukkan bahwa model regresi polinomial lokal lebih baik dengan nilai R^2 sebesar 95,44%.

Penelitian-penelitian sebelumnya masih terbatas pada penggunaan satu variabel prediktor. Keterbatasan ini mendorong penulis untuk melakukan pengembangan dengan menggunakan dua variabel prediktor. Penelitian ini akan membahas pemodelan angka buta huruf di Indonesia dengan pendekatan regresi nonparametrik polinomial lokal multivariabel. Model polinomial lokal multivariabel digunakan karena terdapat lebih dari satu variabel prediktor, yaitu rata-rata lama sekolah dan persentase penduduk miskin. Derajat polinomial lokal yang digunakan merupakan derajat 0, 1, 2, dan 3. Rentang *bandwidth* yang digunakan pada derajat 0, 1, dan 2 adalah 1 hingga 4 dengan ketelitian 0,3, sedangkan pada derajat 3 berada pada rentang 1 hingga 5 dengan ketelitian 0,3. Pemilihan model terbaik dilakukan berdasarkan nilai MSE terkecil.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, permasalahan dalam penelitian ini dirumuskan sebagai berikut:

1. Bagaimana pemodelan regresi nonparametrik polinomial lokal multivariabel antara rata-rata lama sekolah dan persentase penduduk miskin terhadap angka buta huruf di Indonesia?
2. Bagaimana kebaikan model regresi nonparametrik polinomial lokal multivariabel dalam menjelaskan hubungan antara rata-rata lama sekolah dan persentase penduduk miskin terhadap angka buta huruf di Indonesia?
3. Bagaimana kinerja model regresi nonparametrik polinomial lokal multivariabel antara rata-rata lama sekolah dan persentase penduduk miskin terhadap angka buta huruf di Indonesia?

1.3 Batasan Masalah

Berdasarkan rumusan masalah yang telah disampaikan, batasan masalah dalam penelitian ini sebagai berikut:

1. Data yang digunakan merupakan data rata-rata lama sekolah, persentase penduduk miskin, dan angka buta huruf di Indonesia pada tahun 2024 yang diperoleh dari Badan Pusat Statistik (BPS).
2. Penentuan titik lokal, *bandwidth*, dan derajat polinomial optimal dilakukan dengan menggunakan metode *Mean Squared Error* (MSE).
3. Kinerja model regresi diuji dengan menggunakan *Symmetric Mean Absolute Percentage Error* (SMAPE).

4. Fungsi kernel yang digunakan adalah fungsi kernel Gaussian.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah disampaikan, maka tujuan penulis dalam penelitian ini sebagai berikut:

1. Menentukan pemodelan regresi nonparametrik polinomial lokal multivariabel antara rata-rata lama sekolah dan persentase penduduk miskin terhadap angka buta huruf di Indonesia.
2. Menentukan kebaikan model regresi nonparametrik polinomial lokal multivariabel dalam menjelaskan hubungan antara rata-rata lama sekolah dan persentase penduduk miskin terhadap angka buta huruf di Indonesia.
3. Menentukan kinerja model regresi nonparametrik polinomial lokal multivariabel antara rata-rata lama sekolah dan persentase penduduk miskin terhadap angka buta huruf di Indonesia.