

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kabupaten Subang merupakan salah satu wilayah di Provinsi Jawa Barat yang menunjukkan kecenderungan peningkatan jumlah penduduk dari tahun ke tahun. Berdasarkan data dari Badan Pusat Statistik (BPS) Kabupaten Subang, jumlah penduduk pada tahun 2023 mencapai sekitar 1,64 juta jiwa dengan rata-rata laju pertumbuhan penduduk sebesar 0,67% per tahun selama lima tahun terakhir. Kondisi tersebut menunjukkan bahwa dinamika kependudukan di Kabupaten Subang terus mengalami perubahan sehingga memerlukan perhatian dalam proses perencanaan pembangunan daerah. Pemerintah daerah membutuhkan informasi mengenai jumlah penduduk pada periode mendatang sebagai dasar dalam menyusun perencanaan pembangunan yang lebih efektif, efisien dan terarah[1].

Selama ini, prediksi pertumbuhan penduduk dilakukan menggunakan pendekatan demografi maupun metode statistik konvensional yang memanfaatkan asumsi pertumbuhan tertentu berdasarkan data historis. Pendekatan tersebut telah banyak digunakan dalam penyusunan prediksi penduduk, namun memiliki keterbatasan ketika dihadapkan pada hubungan yang semakin kompleks antar variabel kependudukan. Di sisi lain, perkembangan sistem administrasi kependudukan menghasilkan berbagai indikator yang berpotensi memberikan informasi tambahan mengenai kondisi demografi masyarakat[2].

Seiring dengan berkembangnya teknologi pengolahan data, *data mining* menjadi salah satu pendekatan yang banyak dimanfaatkan untuk membangun model prediksi. *Data mining* tidak hanya mampu menemukan pola dan hubungan antar variabel, tetapi juga membangun model yang dapat digunakan untuk memperkirakan nilai pada periode berikutnya[3]. Dalam konteks penelitian, pendekatan data mining dimanfaatkan untuk membangun model prediksi jumlah penduduk[4], sehingga hasil prediksi tersebut dapat digunakan sebagai dasar dalam menganalisis pertumbuhan penduduk Kabupaten Subang pada periode mendatang dengan memanfaatkan data yang diperoleh dari Dinas Kependudukan dan Pencatatan Sipil Kabupaten subang dari periode 2023 sampai dengan 2025.

Salah satu algoritma machine learning yang memiliki kemampuan baik dalam membangun model prediksi pada data tabular adalah *Extreme Gradient Boosting* (XGBoost). XGBoost dipilih karena mampu memodelkan hubungan nonlinier, menangkap interaksi antarvariabel, bekerja secara efisien pada data tabular, serta menyediakan analisis *feature importance* sehingga sesuai dengan karakteristik data administrasi kependudukan yang digunakan dalam penelitian [5]. Berbagai penelitian terdahulu menunjukkan bahwa XGBoost memiliki performa yang baik pada berbagai kasus prediksi. Namun demikian, penelitian-penelitian tersebut umumnya berfokus pada pengujian performa algoritma pada domain yang berbeda atau menggunakan variabel prediktor yang relatif terbatas. Pemanfaatan indikator administrasi kependudukan sebagai variabel prediktor untuk membangun model prediksi jumlah penduduk pada tingkat pemerintah daerah relatif jarang dilakukan. Maka terdapat peluang dalam melakukan untuk mengeksplorasi potensi data administrasi kependudukan sebagai sumber informasi prediktif[6].

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan pembangunan model *data mining* menggunakan algoritma XGBoost untuk memprediksi jumlah penduduk Kabupaten Subang pada periode berikutnya dengan memanfaatkan berbagai indikator administrasi kependudukan sebagai variabel prediktor. Hasil prediksi jumlah penduduk tersebut selanjutnya digunakan sebagai dasar dalam menganalisis pertumbuhan penduduk selama satu tahun mendatang sehingga mampu memberikan informasi yang lebih objektif sebagai pendukung proses perencanaan pembangunan daerah[7].

Penelitian ini diharapkan memberikan kontribusi dalam tiga aspek. Secara praktis, penelitian ini menyediakan model prediksi jumlah penduduk yang dapat dimanfaatkan sebagai pendukung perencanaan pembangunan daerah. Secara metodologis, penelitian ini menunjukkan pemanfaatan indikator administrasi kependudukan sebagai variabel prediktor dalam pembangunan model prediksi berbasis algoritma XGBoost. Secara akademik, penelitian ini memberikan bukti empiris mengenai performa algoritma XGBoost pada data kependudukan daerah serta mengidentifikasi variabel-variabel yang paling berpengaruh terhadap hasil prediksi melalui analisis *feature importance*[8].

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan diatas, dapat dirumuskan permasalahan pada penelitian sebagai berikut:

1. Bagaimana mengimplementasikan algoritma XGBoost pada model prediksi yang dibangun terhadap data historis pertumbuhan penduduk?
2. Bagaimana kinerja algoritma XGBoost dalam memprediksi pertumbuhan penduduk berdasarkan metrik evaluasi regresi berupa *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan koefisien determinasi (R^2 Score)?
3. Bagaimana melakukan analisis *feature importance* pada model XGBoost untuk mengetahui variabel administrasi kependudukan yang paling berpengaruh dalam prediksi pertumbuhan penduduk?

1.3 Tujuan Penelitian

Tujuan yang hendak dicapai dan telah dirumuskan berdasarkan rumusan masalah di atas adalah sebagai berikut:

1. Mengimplementasikan algoritma XGBoost terhadap model prediksi pertumbuhan penduduk.
2. Menganalisis kinerja algoritma XGBoost dalam memprediksi pertumbuhan penduduk berdasarkan metrik evaluasi regresi berupa *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan koefisien determinasi (R^2 Score).
3. Mengidentifikasi variabel administrasi kependudukan yang memiliki kontribusi paling berpengaruh dalam prediksi jumlah penduduk berdasarkan analisis *feature importance* pada algoritma XGBoost.

1.4 Manfaat Penelitian

Adapun manfaat penelitian yang didapatkan dari penelitian yang dilakukan sebagai berikut:

1. Bagi Praktisi

Memberikan informasi berupa hasil prediksi pertumbuhan penduduk di Kabupaten Subang. Informasi tersebut dapat digunakan sebagai bahan pertimbangan dalam perencanaan pembangunan, penyusunan kebijakan,

serta penyediaan sarana dan prasarana yang berkaitan dengan kebutuhan masyarakat di masa mendatang.

2. Bagi Peneliti

Menambah pengetahuan dan pemahaman mengenai penerapan metode *data mining*, khususnya dalam penggunaan algoritma XGBoost untuk memprediksi pertumbuhan penduduk. Selain itu, penelitian ini juga memberikan pengalaman dalam proses pengolahan data, pembuatan model dan evaluasi hasil prediksi.

3. Bagi Akademik

Menjadi referensi bagi penelitian selanjutnya yang berkaitan dengan prediksi menggunakan metode *data mining*, serta memberikan kontribusi dalam pengembangan ilmu pengetahuan di bidang *data mining* dan analisis prediksi.

1.5 Batasan Masalah

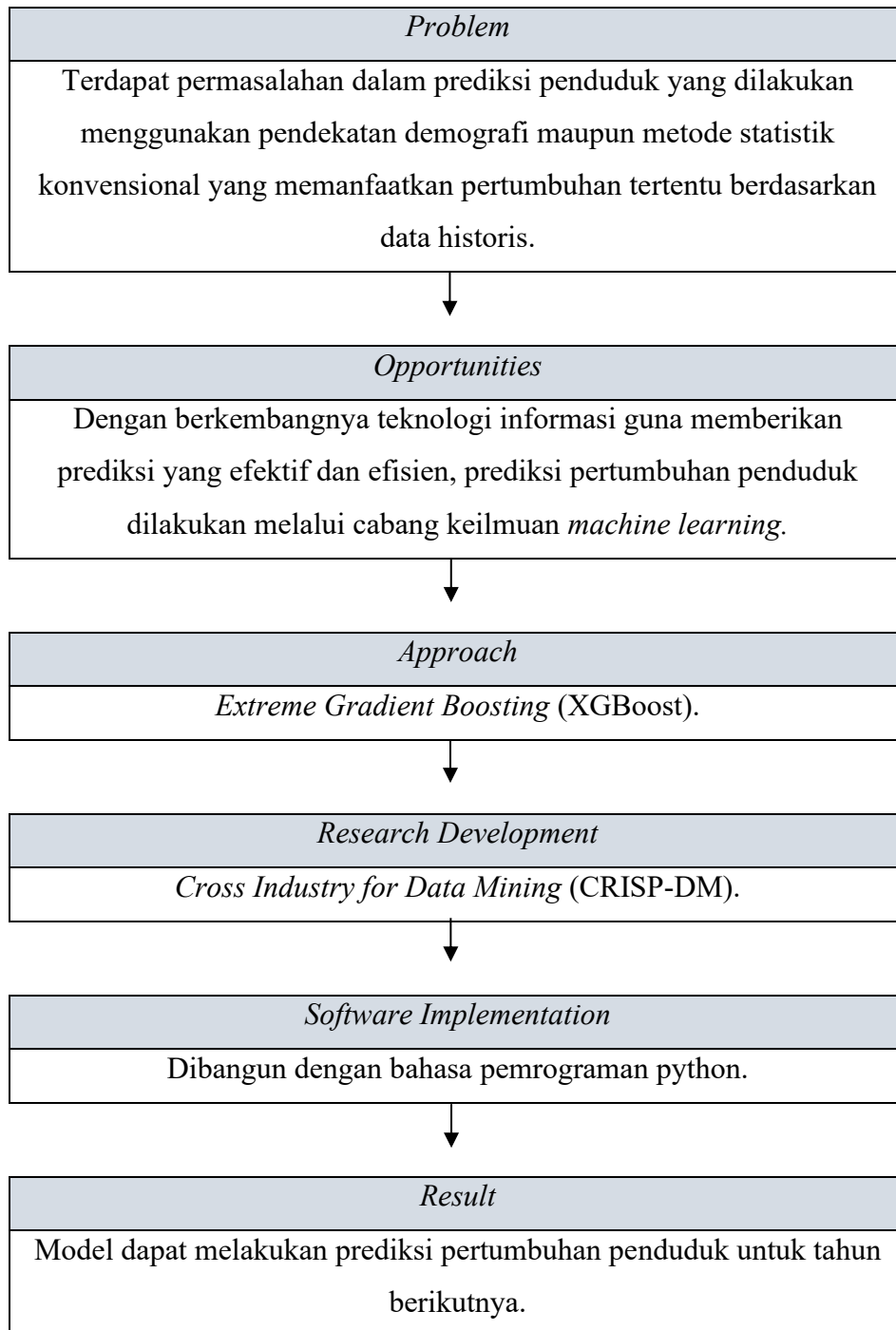
Batasan-batasan masalah yang ditetapkan pada penelitian ini sebagai berikut:

1. Data yang digunakan dalam penelitian adalah data Kependudukan Kabupaten Subang berdasarkan periode waktu 2023 sampai dengan 2025 yang diperoleh dari Dinas Kependudukan dan Pencatatan Sipil Kabupaten Subang.
2. Penelitian berfokus pada pembangunan model prediksi pertumbuhan penduduk berdasarkan data historis yang disediakan oleh instansi.
3. Hasil prediksi dilakukan untuk tahun selanjutnya berdasarkan ketersediaan data historis kependudukan.
4. Penelitian berfokus terhadap pengumpulan data, pemrosesan data dan pelatihan model sehingga menghasilkan model prediksi tanpa memerlukan *interface*.

1.6 Kerangka Pemikiran

Kerangka pemikiran dari tugas akhir ini dapat diuraikan sebagai tabel berikut:

Tabel 1.1 Kerangka pemikiran sistem



1.7 Sistematika Penulisan

Sistematika penulisan dalam penelitian tugas akhir untuk membuat aplikasi atau perangkat lunak dibagi terhadap 5 bab yang menggambarkan dari keseluruhan isi penelitian yang dilakukan. Berikut adalah sistematika penulisannya:

BAB I PENDAHULUAN

Bab ini membahas latar belakang penelitian yang menjelaskan alasan dilakukannya penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, kerangka pemikiran dan sistematika penulisan laporan.

BAB II KAJIAN LITERATUR

Melalui *State of The Art* yang menguraikan penelitian-penelitian terdahulu yang relevan, yang berfungsi sebagai pendukung sekaligus bahan perbandingan untuk menunjukkan pembaharuan penelitian yang dilakukan. Dalam bab kedua ini memuat landasan teori yang mencakup konsep, teori dan dasar keilmuan yang diperlukan sebagai acuan dalam mendukung proses perancangan dan pelaksanaan penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan tahapan penelitian yang dilakukan menggunakan metode CRISP-DM, mulai dari *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, hingga *deployment*. Selain itu, dijelaskan juga sumber dan teknik pengumpulan data, proses praproses data, penerapan algoritma XGBoost, metode evaluasi model, serta alat dan bahan yang digunakan dalam penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dan pembahasan dari implementasi model prediksi pertumbuhan penduduk menggunakan algoritma XGBoost berdasarkan jumlah data penduduk, tingkat kelahiran, kematian dan migrasi. Hasil prediksi dianalisis menggunakan metrik evaluasi untuk mengetahui tingkat akurasi model.

BAB V PENUTUP

Bab ini berisi kesimpulan yang diperoleh dari hasil penelitian sesuai dengan tujuan yang telah dirumuskan. Selain itu, disampaikan juga saran dan rekomendasi yang dapat digunakan untuk pengembangan penelitian selanjutnya.