

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemiskinan adalah masalah sosial dan ekonomi menjadi perhatian utama seluruh negara di dunia, termasuk di Indonesia. Dampak akibat dari kemiskinan yang semakin luas akan mempengaruhi kualitas hidup sehingga mempengaruhi stabilitas sosial dan ekonomi suatu negara. Tingkat kemiskinan di Indonesia menjadi perhatian serius meskipun telah mengalami penurunan setiap tahunnya. Penduduk yang hidup dalam kemiskinan di Indonesia berjumlah 25,90 juta orang dengan proporsi penduduk miskin sebanyak 14,90 persen di tahun 2023. “Besaran penduduk miskin tersebut mendekati batasan kemiskinan kronis merupakan suatu tantangan tersendiri bagi stakeholder di Indonesia untuk bisa menurunkan lebih kecil lagi. Usaha untuk menurunkan tingkat kemiskinan ke tingkat yang lebih baik perlu dukungan kebijakan yang tepat” (Badan Pusat Statistik, 2023).

Penyebab tingginya angka kemiskinan di Indonesia dipengaruhi oleh beberapa faktor, seperti besarnya jumlah penduduk miskin, rendahnya tingkat pengeluaran perkapita, tingkat pendidikan, tingkat pengangguran terbuka, laju pertumbuhan penduduk dan indeks pembangunan manusia dapat mempengaruhi kesejahteraan masyarakat. Dampak tidak hanya dirasakan oleh individu, tetapi juga memicu ketidakstabilan sosial dan politik setiap kabupaten dan kota seluruh Indonesia.

Metode yang digunakan dalam klasifikasi diantaranya *Support Vector Machine* dan Regresi Logistik. “Metode klasifikasi terdiri dari tiga metode yaitu metode klasifikasi linier (*Linier Classification Models*), nonlinier (*Nonlinier Classification Models*), dan metode klasifikasi berbasis pohon dan aturan (*Trees/Rule-Based Models*)” (Woro Maharsi & Hadi, 2022).

Analisis data menjadi alat yang efektif untuk memahami pola kemiskinan. Metode prediksi berbasis *Machine Learning* seperti *Support Vector Machine* dan Regresi Logistik sangat relevan untuk penelitian ini. Kedua pendekatan ini memiliki kelebihan tersendiri dalam menangani masalah klasifikasi. Dengan menggunakan variabel jumlah penduduk miskin, persentase penduduk miskin (%), pengeluaran per kapita disesuaikan (ribu rupiah/Orang/tahun), angka harapan hidup

(tahun), tingkat pengangguran terbuka (%), rata-rata lama sekolah (tahun), laju pertumbuhan penduduk (tahun), indeks pembangunan manusia (tahun), Pendapatan Dalam Rasio Bruto (PDRB) per kapita atas dasar harga berlaku menurut kabupaten/kota (ribu rupiah), angka melek huruf penduduk usia 15 tahun keatas (%) sebagai variabel *Input* (X) dan kolom target untuk variabel *Output* (Y) adalah status kemiskinan yang diklasifikasikan berdasarkan nilai garis kemiskinan nasional yang ditetapkan oleh BPS, yaitu Rp569.299 perkapita perbulan. Kemudian mengukur kinerja kedua metode berdasarkan matrik akurasi, *precision*, *recall*, *F1-score*, *ROC Curve* dan *AUC*.

Penelitian sebelumnya mengenai prediksi kemiskinan yang menggunakan *Support Vector Machine* dan *Regresi Logistik* diantaranya “Metode (SVM) pada Klasifikasi dan Karakterisasi tingkat kedalaman kemiskinan di Provinsi Jawa Timur” (MELIANA NIKE RAKHMASARI, 2022). “Penerapan *Support Vector Machine* untuk klasifikasi kabupaten tertinggal di Provinsi Maluku” (Latuconsina et al., 2024). “Analisis *Regresi Logistik Biner* terhadap data indeks kedalaman kemiskinan di Indonesia” (Fitri, 2022).

Penelitian mengenai pemilihan metode klasifikasi antara *SVM* dan *Regresi Logistik* untuk mengidentifikasi status kemiskinan rumah tangga di wilayah Provinsi Lampung” (Woro Maharsi & Hadi, 2022). Pada penelitian tersebut menggunakan Metode *SVM* dan *Regresi Logistik* secara terpisah dan tidak secara langsung pada konteks prediksi kemiskinan di Indonesia. Penggunaan variabel yang digunakan sedikit dan menggunakan dataset 38 Provinsi. Penelitian ini menggunakan 10 variabel, dataset 416 Kabupaten dan 98 Kota di Indonesia. Menurut beberapa penelitian sebelumnya, metode *SVM* memiliki tingkat akurasi prediksi yang cukup tinggi, di atas 80%, sedangkan *Regresi Logistik* memiliki tingkat akurasi di atas 70%.

Solusi yang diusulkan terletak pada dampaknya terhadap pengambilan keputusan berbasis data. Dengan memilih metode prediksi yang lebih akurat, efektivitas dan efisiensi dalam pengalokasian sumber daya program pengentasan kemiskinan dapat lebih ditingkatkan. Hasil penelitian ini juga berperan dalam memperkaya pengembangan ilmu pengetahuan, khususnya melalui penerapan machine learning dan statistik di sektor sosial ekonomi.

1.2 Perumusan Masalah

Penelitian ini merumuskan permasalahan sebagai berikut:

1. Bagaimana membangun model prediksi tingkat kemiskinan di Indonesia menggunakan metode *Support Vector Machine (SVM)* dan Regresi Logistik?
2. Bagaimana performa model yang dibangun menggunakan metode *Support Vector Machine (SVM)* dan Regresi Logistik dalam melakukan prediksi tingkat kemiskinan di Indonesia?
3. Bagaimana hasil perhitungan perbandingan tingkat kemiskinan menggunakan metode *Support Vector Machine (SVM)* dengan metode Regresi Logistik?

1.3 Batasan Masalah

Peneliti menetapkan beberapa batasan sebagai berikut:

1. Data yang digunakan sebagai variabel *Input X* yaitu jumlah penduduk miskin, Persentase Penduduk Miskin (%), Pengeluaran Per kapita Disesuaikan (Ribuan Rupiah/Orang/Tahun), Angka Harapan Hidup (Tahun), Tingkat Pengangguran Terbuka (%), Rata-rata Lama Sekolah (Tahun), Laju Pertumbuhan Penduduk (Tahun), Indeks Pembangunan Manusia (Tahun), PDRB Per kapita Atas Dasar Harga Berlaku menurut Kabupaten/Kota (ribu rupiah), Angka Melekat Huruf Penduduk Usia 15 Tahun Keatas (%).
2. Kolom target untuk variabel *Output (Y)* adalah status kemiskinan yang ditentukan berdasarkan nilai garis kemiskinan nasional yang ditetapkan oleh BPS, yaitu Rp.569.299 per kapita per bulan. Daerah dengan pengeluaran per kapita di bawah Rp569.299 dikategorikan sebagai “Miskin” (kelas 1), sedangkan daerah dengan pengeluaran per kapita sama dengan atau di atas Rp569.299 dikategorikan sebagai “Tidak Miskin” (kelas 0).
3. Algoritma dalam penelitian ini terbatas pada *SVM* dan Regresi Logistik.

1.4 Tujuan dan Manfaat

Tujuan

Dalam penelitian ini penulis bertujuan untuk:

1. Membangun model prediksi tingkat kemiskinan di Indonesia menggunakan metode *Support Vector Machine (SVM)* dan metode *Regresi Logistik*.

2. Menganalisis performa model yang dibangun menggunakan metode *Support Vector Machine (SVM)* dan metode *Regresi Logistik* dalam memprediksi tingkat kemiskinan di Indonesia.
3. Membandingkan efektivitas kedua algoritma tersebut berdasarkan matrik evaluasi seperti akurasi, presisi, recall, F1-score, Roc Curve dan Auc.

Manfaat

Dalam penelitian ini terdapat beberapa manfaat diantaranya:

1. Memberikan rekomendasi kepada pemerintah dalam membuat kebijakan berbasis data yang tepat sasaran untuk mengurangi tingkat kemiskinan.
2. Meningkatkan pemahaman tentang faktor-faktor yang mempengaruhi tingkat kemiskinan di Indonesia.
3. Memberikan kontribusi akademik berupa studi kasus penerapan algoritma pembelajaran mesin dalam permasalahan sosial ekonomi khususnya dibidang kemiskinan.
4. Menyediakan model prediksi yang dapat menjadi dasar pengembangan lebih lanjut untuk aplikasi praktis seperti dashboard pemantauan kemiskinan atau system pendukung keputusan.

1.5 Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan analisis perbandingan antar algoritma. Adapun tahapan penelitian yang dilakukan adalah sebagai berikut:

1. Pengumpulan Data

Penelitian ini menggunakan data yang diperoleh dari Badan Pusat Statistik Indonesia (BPS). Data ini mencakup variabel yang relevan untuk prediksi tingkat kemiskinan seperti Jumlah Penduduk Miskin, Persentase Penduduk Miskin (%), Pengeluaran Per kapita Disesuaikan (Ribuan Rupiah/Orang/Tahun), Angka Harapan Hidup (Tahun), Tingkat Pengangguran Terbuka (%), Rata-rata Lama Sekolah (Tahun), Laju Pertumbuhan Penduduk (Tahun), Indeks Pembangunan Manusia (Tahun), PDRB Per kapita Atas Dasar Harga Berlaku menurut Kabupaten/Kota (Ribuan Rupiah), Angka Melekat Huruf Penduduk Usia 15 Tahun Keatas (%).

2. Persiapan Data

- Melakukan preprocessing data, seperti penanganan nilai yang hilang (missing values), normalisasi data, pembagian dataset menjadi data training dan data testing.
- Menentukan kolom input (X) dan output (Y) sesuai dengan batasan masalah yang telah ditentukan.

3. Implementasi Metode Prediksi

Metode SVM dan metode Regresi Logistik akan melakukan pelatihan (training) dan pengujian (testing) model untuk memprediksi tingkat kemiskinan berdasarkan variabel yang sudah ditentukan.

4. Analisis dan Perbandingan

- Evaluasi model: menggunakan matrik evaluasi seperti akurasi, *precision*, *recall*, *F1-score*, ROC Curve dan AUC untuk mengukur performa kedua algoritma dalam memprediksi tingkat kemiskinan.
- Perbandingan Metode: menganalisis kedua algoritma berdasarkan hasil evaluasi untuk menentukan metode yang lebih efektif.