

BAB I PENDAHULUAN

Bab I berisi tentang latar belakang, rumusan masalah, manfaat dan tujuan, batasan masalah dan sistematika penulisan.

1.1 Latar Belakang

Kemiskinan merupakan kondisi seseorang yang tidak mampu memenuhi kebutuhan pokoknya serta kondisi kebijakan yang menyebabkan ia lebih miskin dibandingkan sekitarnya (BAPPENAS, 2018)[1]. Kemiskinan merupakan masalah kompleks yang mempengaruhi kualitas hidup masyarakat. Menurut data Badan Pusat Statistik (BPS), jumlah penduduk miskin di Indonesia pada tahun 2022 mencapai 27,54 juta jiwa atau sekitar 10,02% dari total penduduk. Angka ini menunjukkan bahwa kemiskinan masih menjadi perhatian serius di Indonesia. Kemiskinan memiliki dampak yang luas terhadap berbagai aspek kehidupan masyarakat, seperti kesehatan, pendidikan, dan kesempatan kerja.

Pemerintah Indonesia telah melakukan berbagai upaya untuk menanggulangi kemiskinan, namun penurunan tingkat kemiskinan secara signifikan masih menjadi tantangan. Hal ini disebabkan oleh banyaknya faktor yang mempengaruhi kemiskinan, termasuk ketimpangan distribusi pendapatan, akses terbatas ke pendidikan dan layanan kesehatan, serta kondisi infrastruktur yang belum merata di berbagai wilayah[2]. Pada dasarnya, terdapat berbagai aspek dalam mengartikannya, umumnya dalam standar kemiskinan yang sering digunakan adalah aspek ekonomi. Hal ini menunjukkan perlunya pendekatan yang lebih sistematis dan berbasis data untuk memahami dan menanggulangi masalah kemiskinan.

Di era yang modern sekarang, pastinya teknologi informasi dan ketersediaan data sudah berkembang sangat pesat, metode berbasis data seperti pembelajaran mesin (*Machine Learning*) mulai banyak digunakan di semua kalangan untuk analisis dan pengambilan Keputusan. Sebagai metode dalam pembelajaran mesin (*Machine Learning*) seperti *Regresi Logistik Biner*, analisis diskriminan, serta *Decision Tree*. [2] Pada penelitian menggunakan metode *Regresi Logistik Biner* dalam Penentuan Ketepatan Klasifikasi Tingkat

Kedalaman Kemiskinan Provinsi-Provinsi di Indonesia mempunyai keakuratan prediksi secara menyeluruh sebesar 85.3%[3]. Pada penelitian lain Klasifikasi tingkat keparahan kemiskinan provinsi di Indonesia Dengan *analisis diskriminan* memiliki akurasi 67,7%[4]. Penelitian lain Klasifikasi Tingkat Kemiskinan di Indonesia Menggunakan Algoritma *Decision Tree* memiliki nilai akurasi 93%[5]. Berdasarkan beberapa penelitian terdahulu, yang menggunakan metode *Decision Tree* memiliki nilai akurasi paling tinggi, namun *decision tree* masih dapat dikembangkan dengan menggunakan metode *Random Forest*, karena setiap *decision tree* dilatih menggunakan subset data latih, sehingga model *Random Forest* mampu menghasilkan hasil yang lebih stabil dan akurat. Dalam klasifikasi, prediksi akhir ditentukan berdasarkan mayoritas voting dari hasil klasifikasi setiap *decision tree*. [6] Algoritma *Random Forest* sangat sesuai untuk klasifikasi pada dataset besar yang memiliki banyak fitur sehingga dapat menghasilkan nilai akurasi yang lebih baik.

Penelitian ini akan mengimplementasikan metode *Random Forest* untuk mengklasifikasikan tingkat kemiskinan di Indonesia[7]. Dengan memanfaatkan dataset yang tersedia di Kaggle, penelitian ini akan menganalisis berbagai indikator kemiskinan di berbagai wilayah di Indonesia. Tujuan utama dari penelitian ini adalah untuk mengidentifikasi faktor-faktor yang paling berpengaruh terhadap kemiskinan dan memberikan rekomendasi yang dapat digunakan untuk mengembangkan strategi penanggulangan kemiskinan yang lebih efektif dan efisien[5].

Sehingga hasil yang diperoleh diharapkan dapat berguna dan membantu pemerintah untuk menentukan sasaran serta kebijakan penanggulangan kemiskinan yang efektif dan terprogram. Tujuan lain penelitian ini untuk mengetahui hasil dan tingkat akurasi metode *random forest* dalam melakukan klasifikasi apakah memiliki nilai akurasi yang lebih tinggi atau malah lebih rendah dibandingkan dengan metode klasifikasi yang lain.[1] Dari uraian latar belakang di atas, maka penulis mengambil judul Klasifikasi Tingkat Kemiskinan di Indonesia menggunakan Metode *Random Forest*.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Berdasarkan latar belakang yang telah dijelaskan di atas permasalahan yang ingin diselesaikan dalam penelitian ini yaitu hasil dari mengklasifikasikan data tingkat kemiskinan yang ada di Indonesia dengan metode *random forest*.

1.2.2 Solusi Permasalahan

Metode yang akan digunakan dalam penelitian ini yaitu *Random Forest*.

1.2.3 Pertanyaan Penelitian

Berdasarkan masalah dan metode yang disajikan, pertanyaan penelitian sebagai berikut:

Berapa hasil akurasi pada klasifikasi data tingkat kemiskinan di Indonesia menggunakan metode *Random Forest*?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan Penelitian

Penelitian ini memiliki tujuan untuk memperoleh hasil akurasi pada klasifikasi tingkat kemiskinan di Indonesia.

1.3.2 Manfaat Penelitian

Manfaat penelitian ini yaitu untuk mengetahui hasil kinerja metode *random forest* dan mengetahui perbandingan metode split data 90:10%, 80:20%, 70:30%, 60:40% dan 50:50% pada klasifikasi menggunakan metode *random forest*.

1.4 Batasan-batasan

Berdasarkan pada perumusan masalah yang sudah disebutkan di atas, maka batasan masalah penelitian ini adalah sebagai berikut:

1. Metode yang digunakan pada pengklasifikasian ini adalah Metode *Random Forest*,
2. Dataset yang digunakan dalam penelitian ini diambil dari dataset yang ada di website Kaggle. Dataset tersebut mempunyai data sebanyak 514 data.

3. Dataset yang diklasifikasi yaitu dataset yang mencakup sebanyak 10 atribut dan 514 data.

1.5 Sistematika Penulisan

Sistematika penulisan skripsi ini mencakup beberapa bab, yang disusun sebagai berikut:

- **BAB 1 PENDAHULUAN**

Bab ini berisi latar belakang, perumusan masalah yang meliputi permasalahan, metode usulan, dan pertanyaan penelitian, kemudian selanjutnya terdapat tujuan dan manfaat dari penelitian, batasan-batasan penelitian dan sistematika penulisan.

- **BAB 2 KAJIAN PUSTAKA**

Bab ini berisi mengenai landasan teori atau referensi mengenai materi metode yang digunakan pada penelitian ini, diantaranya meliputi rujukan paper yang digunakan terkait klasifikasi tingkat kemiskinan di Indonesia.

- **BAB 3 METODE PENELITIAN**

Bab ini berisi penjelasan tentang arsitektur penelitian yang berisi tentang langkah-langkah atau proses dalam memecahkan masalah yang sedang diteliti menggunakan metode yang telah ditetapkan.

- **BAB 4 HASIL DAN PEMBAHASAN**

Bab ini berisi hasil penelitian, meliputi tahapan perancangan sistem disertai code, hasil running, kemudian terdapat tampilan yang dihasilkan, hasil pengujian sistem.

- **BAB 5 PENUTUP**

Bab ini berisi kesimpulan dari penelitian yang telah dilakukan dan saran untuk penelitian selanjutnya.