

BAB I

PENDAHULUAN

1.1 Latar Belakang

Keberhasilan sebelum tahun 1990 lebih sering diukur menggunakan indikator ekonomi seperti pertumbuhan ekonomi dan pendapatan per kapita, sehingga dimensi distribusi kesejahteraan dan kualitas hidup masyarakat kurang mendapat perhatian (Todaro dan Smith, 2015). Negara-negara berkembang kemudian mulai menyusun kebijakan pembangunan yang tidak hanya berorientasi pada pertumbuhan ekonomi, tetapi juga pada pengurangan kemiskinan serta peningkatan kesejahteraan sosial sebagai upaya memperbaiki kualitas hidup penduduk. Kebutuhan untuk menilai keberhasilan pembangunan secara lebih komprehensif mendorong munculnya konsep pembangunan manusia, yang tidak hanya berfokus pada aspek ekonomi dan pembangunan infrastruktur, tetapi juga pada peningkatan kualitas hidup individu dan masyarakat secara menyeluruh.

Amartya Sen, ekonom asal India dan peraih Nobel Ekonomi tahun 1998, merupakan salah satu tokoh penting yang mempopulerkan konsep pembangunan manusia melalui karyanya yang berjudul *Development as Freedom*. Buku tersebut menegaskan bahwa pembangunan seharusnya dipahami sebagai proses memperluas kebebasan substantif individu, bukan sekadar peningkatan pendapatan atau produk agregat. Pembangunan manusia dalam perspektif ini tidak hanya berkaitan dengan pertumbuhan ekonomi, tetapi juga dengan peningkatan kemampuan dan pilihan hidup seseorang, termasuk akses terhadap pendidikan, kesehatan, serta partisipasi sosial dan politik (Sen, 1999).

United Nations Development Programme (UNDP) menerbitkan *Human Development Report* pertama pada tahun 1990 yang menjadi tonggak penting dalam pengembangan konsep pembangunan manusia (UNDP, 1990). Laporan tersebut memperkenalkan Indeks Pembangunan Manusia (IPM) atau Human Development Index (HDI) sebagai ukuran kemajuan suatu negara berdasarkan dimensi harapan hidup, pendidikan, dan standar hidup layak, bukan semata-mata produk domestik bruto per kapita. Perkembangan pembangunan manusia selanjutnya menempatkan dalam kerangka pembangunan berkelanjutan yang menekankan keseimbangan antara pertumbuhan ekonomi, keadilan sosial, dan kelestarian lingkungan, sehingga mampu meningkatkan kualitas hidup secara menyeluruh dan berkelanjutan.

Konsep kemampuan atau kapabilitas merujuk pada kebebasan individu untuk memilih berbagai bentuk kehidupan yang dianggap bernilai, yang mencakup akses terhadap pendidikan, kesehatan, kesempatan kerja, serta partisipasi dalam kehidupan sosial dan politik (Sen, 1999). Pembangunan ideal berorientasi pada peningkatan kemampuan individu dalam mencapai tujuan hidup yang mereka anggap penting. Indikator keberhasilan pembangunan tidak hanya terbatas pada peningkatan pendapatan per kapita, tetapi juga mencakup aspek kesehatan, pendidikan, kebebasan, dan keadilan sosial, serta menjaga kelestarian lingkungan agar kesejahteraan dapat dipertahankan lintas generasi.

Pembangunan manusia memiliki tujuan yang lebih luas dibandingkan sekadar pertumbuhan ekonomi. Upaya tersebut mencakup peningkatan harapan hidup melalui perluasan akses layanan kesehatan, perbaikan sanitasi, serta pemenuhan gizi yang memadai. Dimensi pendidikan menekankan pentingnya akses yang adil dan merata terhadap pendidikan berkualitas agar masyarakat memperoleh

pengetahuan dan keterampilan yang dibutuhkan. Dimensi kesejahteraan ekonomi berfokus pada peningkatan pendapatan dan pengurangan kemiskinan sehingga kebutuhan dasar dapat terpenuhi. Dimensi kebebasan menegaskan hak individu untuk mengambil keputusan penting dalam kehidupan serta berpartisipasi dalam proses sosial, politik, dan ekonomi, termasuk mendorong kesetaraan gender dan pemberdayaan kelompok secara menyeluruh.

Perkembangan teknologi serta ketersediaan data pada era digital mendorong pemanfaatan *machine learning* dalam pengolahan dan analisis data pembangunan manusia, termasuk data IPM di Indonesia (Pamungkas dan Widiyanto, 2023). Algoritma *machine learning* dapat menghasilkan kinerja yang berbeda meskipun menggunakan data yang sama, sehingga diperlukan pendekatan yang mampu menggabungkan keunggulan masing-masing algoritma agar hasil klasifikasi lebih akurat dan stabil (Mahmuda, 2024). *Ensemble learning* menawarkan solusi melalui penggabungan beberapa model dalam satu kerangka kerja untuk memperoleh kinerja yang lebih baik dibandingkan model tunggal (Zhou, 2012).

Ensemble learning mengombinasikan keluaran beberapa model melalui berbagai pendekatan, seperti *bagging*, *voting*, *AdaBoost*, dan *stacking*, dengan tujuan meningkatkan akurasi serta kemampuan generalisasi model. Penelitian terbaru menunjukkan bahwa algoritma *Support Vector Machine* (SVM), *Random Forest*, dan *AdaBoost* dalam kerangka *ensemble learning* mampu memberikan kinerja klasifikasi yang baik pada berbagai kasus, termasuk data sosial dan pembangunan (Arumnisaa *et al.*, 2023). Hasil penelitian tersebut menunjukkan bahwa *Random Forest* memiliki akurasi tertinggi dalam klasifikasi IPM di

Indonesia, sehingga berpotensi dikembangkan lebih lanjut sebagai bagian dari *ensemble classifier*.

Data IPM sering menghadapi masalah ketidakseimbangan kelas, misalnya jumlah wilayah dengan kategori IPM sangat tinggi lebih sedikit dibandingkan kategori lainnya, sehingga model cenderung bias terhadap kelas mayoritas (He dan Garcia, 2009). Teknik *Synthetic Minority Over-sampling Technique* (SMOTE) digunakan untuk mengatasi permasalahan tersebut dengan membangkitkan sampel sintesis pada kelas minoritas sehingga distribusi data menjadi lebih seimbang (Chawla *et al.*, 2002). Penerapan SMOTE terbukti mampu meningkatkan kinerja *Random Forest* pada data tidak seimbang, dengan peningkatan *accuracy* sebesar 3,45%, *precision* 4,8%, *sensitivity* 7,1%, dan *F1-score* 4,8% (Erlin, 2022).

Evaluasi performa serta pemilihan *hyperparameter* model klasifikasi memerlukan prosedur *cross validation* yang mampu memberikan estimasi kinerja yang stabil serta mengurangi risiko *overfitting*. *GridSearchCV* menerapkan *cross validation* pada berbagai kombinasi *hyperparameter* sehingga diperoleh konfigurasi model terbaik pada data latih (Pedregosa *et al.*, 2011). Penggunaan *GridSearchCV* berbasis *cross validation* pada klasifikasi IPM dengan *Support Vector Machine* (SVM), *Random Forest*, dan *AdaBoost* terbukti memberikan estimasi akurasi dan stabilitas model yang baik (Arumnisaa *et al.*, 2023).

Khofifah (2021) menunjukkan bahwa variabel ekonomi, pendidikan, dan kesehatan berpengaruh signifikan terhadap IPM kabupaten/kota di Indonesia. Temuan tersebut menegaskan pentingnya pengembangan metode analisis dan klasifikasi yang lebih komprehensif untuk memetakan variasi pembangunan manusia sebagai dasar perencanaan pembangunan. Penelitian lain menunjukkan

bahwa penggunaan *machine learning* dan *ensemble classifier* mampu meningkatkan akurasi klasifikasi IPM, dengan nilai *accuracy* sebesar 85,23%, *specificity* 71,63%, *sensitivity* 95,05% (Arumnisaa *et al.*, 2023).

Arta (2023) menunjukkan bahwa penerapan SMOTE pada klasifikasi IPM dapat meningkatkan akurasi dan sensitivitas model *Random Forest* dan SVM pada data tidak seimbang. Integrasi SMOTE dengan *Random Forest* dan algoritma lain menjadi penting ketika data memiliki ketimpangan kelas yang signifikan, karena mampu meningkatkan kemampuan model dalam mengenali kelas minoritas (Erlin, 2022).

Penelitian ini bertujuan untuk menerapkan *ensemble classifier* dengan mengombinasikan algoritma *Support Vector Machine (SVM)*, *Random Forest*, dan *AdaBoost* dalam klasifikasi data IPM di Indonesia tahun 2024. Pendekatan yang digunakan mencakup penerapan SMOTE untuk mengatasi ketidakseimbangan kelas serta penggunaan *GridSearchCV* berbasis *cross validation* dalam pemilihan *hyperparameter* dan evaluasi awal performa model. Model yang dihasilkan diharapkan mampu memberikan klasifikasi IPM yang akurat dan stabil, serta menjadi dasar analitis dalam perumusan kebijakan pembangunan manusia yang lebih tepat sasaran dan efektif.

1.2 Rumusan Masalah

Penelitian ini berangkat dari kebutuhan untuk mengidentifikasi penerapan pendekatan *ensemble classifier* yang mengombinasikan algoritma *Support Vector Machine (SVM)*, *Random Forest*, dan *AdaBoost* dalam melakukan klasifikasi di Indonesia berdasarkan Indeks Pembangunan Manusia (IPM). Penelitian ini juga mempersoalkan efektivitas teknik *Synthetic Minority Over-sampling Technique*

(SMOTE) dalam menangani ketidakseimbangan kelas pada data IPM sehingga model klasifikasi mampu mempelajari setiap kategori secara lebih seimbang. Penelitian ini selain itu mempertanyakan metode evaluasi dan pemilihan *hyperparameter* model menggunakan *GridSearchCV* berbasis *cross validation* dapat memastikan keakuratan dan kestabilan hasil klasifikasi IPM.

1.3 Batasan Masalah

Batasan masalah terhadap data dan metode yang digunakan dalam penelitian ini, yang dibuat penulis berdasarkan permasalahan yang telah dirumuskan, yaitu sebagai berikut:

1. Data penelitian menggunakan data sekunder Indeks Pembangunan Manusia (IPM) kabupaten/kota di Indonesia tahun 2024 yang diperoleh dari Badan Pusat Statistik (BPS).
2. Variabel independen (bebas) yang digunakan meliputi Tingkat Partisipasi Angkatan Kerja, Upah Minimum Kabupaten/Kota, Rata-rata Pengeluaran per Kapita, Jumlah Sekolah SMA dan SMK, Jumlah Murid SMA dan SMK, Indeks Pembangunan Literasi, Jumlah Puskesmas, serta Persentase Rumah Tangga yang Memiliki Akses Sanitasi Layak. IPM digunakan sebagai variabel target yang dikonversi ke dalam kategori kelas.
3. Penanganan ketidakseimbangan kelas pada data IPM dilakukan menggunakan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) yang diterapkan pada data latih.
4. Model klasifikasi dibangun menggunakan tiga algoritma dasar, yaitu *Support Vector Machine* (SVM), *Random Forest*, dan *AdaBoost*, yang dikombinasikan

dalam kerangka *Stacking Ensemble Classifier* dengan *Logistic Regression* sebagai *final estimator*.

5. Proses pelatihan model menggunakan pembagian data menjadi data latih dan data uji. Evaluasi performa setiap algoritma serta model *stacking* dilakukan menggunakan matriks akurasi, presisi, *recall*, dan *F1-score*.
6. Pencarian *hyperparameter* dilakukan menggunakan metode *GridSearchCV* untuk memperoleh kombinasi parameter terbaik pada model.

1.4 Tujuan Penelitian

Tujuan penelitian berdasarkan rumusan masalah yang ditentukan yaitu:

1. Menganalisis penerapan pendekatan *Stacking Ensemble Classifier* yang menggabungkan algoritma *Support Vector Machine*, *Random Forest*, dan *AdaBoost* dalam klasifikasi Indeks Pembangunan Manusia (IPM) kabupaten/kota di wilayah Pulau Jawa tahun 2024..
2. Mengkaji efektivitas teknik *Synthetic Minority Over-sampling Technique* (SMOTE) dalam menangani ketidakseimbangan kelas pada data IPM kabupaten/kota di wilayah Pulau Jawa.
3. Mengevaluasi performa model dasar (*Support Vector Machine*, *Random Forest*, dan *AdaBoost*) serta model *Stacking Ensemble* yang dihasilkan dengan menggunakan *GridSearchCV* berbasis *cross validation* dan matriks akurasi, presisi, *recall*, serta *F1-score* dalam klasifikasi IPM.