

## Identifikasi Faktor Dominan yang Mempengaruhi Indeks Pembangunan Manusia Sumatera Selatan Menggunakan Algoritma C4.5

### *Identifying the Dominant Factors Influencing the Human Development Index of South Sumatra Using the C4.5 Algorithm*

**Farhana Puti Andayu\*, Fenny Purwani, Catur Eri Gunawan**

Universitas Islam Negeri Raden Fatah Palembang; Jl. Pangeran Ratu (Jakabaring), Kelurahan 5 Ulu, Kecamatan Seberang Ulu I, Kota Palembang, Sumatera Selatan, 30267  
Program Studi Sistem Informasi, Fakultas Saintek UIN Raden Fatah, Palembang  
e-mail: [farhanaputiandayu@radenfatah.ac.id](mailto:farhanaputiandayu@radenfatah.ac.id)\*, [fennypurwani\\_uin@radenfatah.ac.id](mailto:fennypurwani_uin@radenfatah.ac.id),  
[caturerig@radenfatah.ac.id](mailto:caturerig@radenfatah.ac.id)

#### **Abstrak**

*Indeks Pembangunan Manusia (IPM) merupakan indikator strategis untuk menilai kualitas pembangunan manusia, sementara Sumatera Selatan masih menghadapi ketimpangan antardaerah yang memerlukan analisis faktor penyebab yang lebih terstruktur. Penelitian ini menerapkan algoritma C4.5 untuk mengidentifikasi faktor dominan yang memengaruhi IPM menggunakan data 17 kabupaten/kota periode 2010–2024 (260 sampel). Proses KDD meliputi selection, preprocessing terhadap 10 missing values, transformation dengan mengonversi IPM numerik menjadi kategori, serta implementasi C4.5 dengan kedalaman maksimal tiga level. Hasil menunjukkan bahwa dimensi ekonomi menjadi determinan utama melalui pengeluaran riil per kapita (60,09%), diikuti harapan lama sekolah (16,83%), umur harapan hidup (11,98%), dan rata-rata lama sekolah (11,10%). Pohon keputusan menghasilkan delapan aturan IF–THEN, dan sistem analitik yang dikembangkan menyediakan visualisasi untuk pemanfaatan dataset pada penelitian berikutnya.*

**Kata kunci**— Algoritma C4.5, Indeks Pembangunan Manusia, Pohon Keputusan, Sumatera Selatan, Feature Importance

#### **Abstract**

*The Human Development Index (HDI) is a strategic indicator for assessing the quality of human development, while South Sumatra continues to experience inter-regional disparities that require a more structured analysis of their underlying factors. This study applies the C4.5 algorithm to identify the dominant determinants of HDI using data from 17 districts/cities for the 2010–2024 period (260 samples). The KDD process comprises selection, preprocessing of 10 missing values, transformation by converting numerical HDI values into categories, and the implementation of C4.5 with a maximum depth of three levels. The findings indicate that economic factors are the primary determinants, with real expenditure per capita contributing 60.09%, followed by expected years of schooling (16.83%), life expectancy (11.98%), and mean years of*

*schooling (11.10%). The decision tree generated eight IF–THEN rules, and the analytical system developed provides visualization features for future dataset utilization.*

**Keywords**— *C4.5 Algorithm, Human Development Index, Decision Tree, South Sumatra, Feature Importance*

## 1. PENDAHULUAN

Pembangunan manusia merupakan aspek utama dalam mewujudkan kesejahteraan sosial dan kemajuan ekonomi suatu daerah. Indeks Pembangunan Manusia (IPM) digunakan sebagai indikator komprehensif untuk mengukur keberhasilan pembangunan, meliputi dimensi kesehatan, pendidikan, dan ekonomi [1]. IPM tidak hanya menilai capaian material, tetapi juga memperhatikan kualitas hidup dan pemerataan antarwilayah [2].

Analisis IPM di Sumatera Selatan menjadi semakin penting mengingat provinsi ini merupakan salah satu wilayah strategis di Indonesia dengan potensi ekonomi yang besar dari sektor pertanian, pertambangan, dan industri. Namun, data BPS menunjukkan bahwa peningkatan IPM yang terjadi belum merata di seluruh kabupaten/kota. Wilayah perkotaan yang memiliki akses lebih baik terhadap layanan pendidikan dan kesehatan cenderung memiliki IPM lebih tinggi, sementara daerah perifer yang bergantung pada sektor pertanian masih menghadapi tantangan dalam peningkatan kualitas sumber daya manusia. Kesenjangan ini memerlukan pendekatan analisis yang mampu mengidentifikasi faktor-faktor spesifik yang paling berpengaruh, sehingga kebijakan pembangunan dapat diarahkan secara lebih tepat sasaran berdasarkan karakteristik masing-masing wilayah.

Meskipun IPM Indonesia menunjukkan peningkatan dari tahun ke tahun [3], ketimpangan antarwilayah masih menjadi tantangan utama, termasuk di Provinsi Sumatera Selatan. Berdasarkan data BPS [4], IPM Sumatera Selatan pada tahun 2024 mencapai 73,84, meningkat 0,66 poin dibanding tahun sebelumnya. Namun, terdapat disparitas antara kota dan kabupaten, wilayah perkotaan seperti Palembang dan Prabumulih menunjukkan nilai IPM tinggi, sementara kabupaten dengan ketergantungan pada sektor pertanian masih tergolong rendah. Ketimpangan ini mengindikasikan perlunya analisis faktor dominan yang berkontribusi terhadap variasi IPM [5].

Pendekatan statistik konvensional umumnya hanya memberikan gambaran deskriptif tanpa menunjukkan hubungan kausalitas antaratribut. Untuk itu, data mining hadir sebagai pendekatan komputasional yang mampu menemukan pola tersembunyi dalam data [6]. Misalnya, umur harapan hidup mencerminkan tingkat kesehatan secara umum, namun tidak secara langsung menunjukkan seberapa besar kontribusinya dibandingkan faktor lain terhadap pencapaian IPM. Begitu juga rata-rata lama sekolah dan harapan lama sekolah yang menunjukkan tingkat pendidikan, tanpa mengeksplorasi sejauh mana dominasi faktor ini terhadap keseluruhan capaian IPM di daerah tertentu [7]. Salah satu algoritma klasifikasi yang banyak digunakan adalah C4.5, yang dikembangkan oleh sebagai penyempurnaan dari ID3 [8]. C4.5 mampu menghasilkan pohon keputusan yang mudah dipahami dan diubah menjadi aturan *IF-THEN*[9].

Metode data mining dengan algoritma C4.5 dipilih dalam penelitian ini karena kemampuannya dalam menghasilkan model yang tidak hanya akurat tetapi juga mudah diinterpretasikan oleh *stakeholder* non-teknis. Berbeda dengan metode *machine learning* lain seperti *Random Forest* yang menghasilkan model terlalu kompleks, C4.5 menghasilkan *decision tree* yang struktur keputusannya dapat ditelusuri dan dipahami secara logis. Hal ini sangat penting dalam konteks kebijakan publik, di mana pembuat keputusan perlu memahami alasan di balik rekomendasi yang diberikan oleh sistem. Selain itu, C4.5 mampu menangani data numerik dan kategorikal secara bersamaan, serta dapat memberikan informasi tentang tingkat kepentingan relatif dari setiap faktor melalui mekanisme *feature importance*, yang menjadi kunci dalam penelitian ini untuk mengidentifikasi faktor dominan yang memengaruhi IPM.

# Identifikasi Faktor Dominan yang Mempengaruhi Indeks Pembangunan Manusia Sumatera Selatan Menggunakan Algoritma C4.5

---

Penelitian terdahulu telah menunjukkan efektivitas C4.5 dalam bidang sosial-ekonomi. Sirait & Azizah (2023) berhasil mengklasifikasikan IPM di Sumatera Barat dengan tingkat akurasi tinggi [10], sedangkan Srirahayu & Pribadie (2023) menemukan bahwa C4.5 lebih unggul dibanding *Naïve Bayes* dan KNN dalam menilai indikator Pembangunan [11]. Pratiwi dkk (2022) juga menegaskan keunggulan C4.5 dalam menghasilkan aturan klasifikasi dengan akurasi 97,83% [12].

Selain itu, faktor pendidikan dan ekonomi sering kali menjadi dimensi paling berpengaruh terhadap IPM [13] [13]. Penelitian Mahendra dkk (2025) menegaskan bahwa umur harapan hidup memiliki korelasi positif terhadap produktivitas ekonomi Masyarakat [14], sedangkan peningkatan rata-rata lama sekolah berdampak pada pengeluaran per kapita [15]. Hal ini menunjukkan bahwa pembangunan manusia tidak hanya bergantung pada ekonomi, tetapi juga kualitas pendidikan dan kesehatan masyarakat.

Berbeda dari studi terdahulu yang menerapkan algoritma C4.5 terutama untuk prediksi, penelitian ini diarahkan untuk membangun model yang mengungkap faktor-faktor utama yang memengaruhi IPM. Dengan demikian, diperoleh pemahaman yang lebih mendalam mengenai kondisi lokal melalui struktur aturan yang terbentuk dari data nyata.

Berdasarkan kesenjangan tersebut, penelitian ini berfokus untuk: (1) mengidentifikasi faktor dominan yang memengaruhi IPM di Sumatera Selatan menggunakan algoritma C4.5; (2) menghasilkan *decision rules* yang dapat menjadi dasar kebijakan pembangunan; dan (3) mengimplementasikan sistem berbasis *web* yang menampilkan hasil analisis dan visualisasi faktor dominan IPM secara otomatis.

## 2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan deskriptif kuantitatif dengan metode data mining berbasis algoritma C4.5. Data sekunder diperoleh dari Badan Pusat Statistik (BPS) Sumatera Selatan periode 2010-2024, mencakup 17 kabupaten/kota dengan total 270 observasi.

### 2.1 Proses Knowledge Discovery in Database (KDD)

Tahapan penelitian mengikuti pada proses KDD yang terdiri dari proses analisis dimulai dengan *selection*, yaitu pemilihan atribut berdasarkan indikator resmi IPM yang mencakup pengeluaran riil per kapita, harapan lama sekolah, rata-rata lama sekolah, dan umur harapan hidup. Tahap *preprocessing* dilakukan dengan menangani 10 *missing values* menggunakan metode *deletion* sehingga diperoleh 260 data bersih. Pada tahap *transformation*, nilai IPM numerik dikonversi ke dalam empat kategori: sangat tinggi ( $\geq 80$ ), tinggi (70–79.99), sedang (60–69.99), dan rendah ( $< 60$ ). Selanjutnya, tahap data mining menerapkan algoritma C4.5 untuk membentuk *decision tree* berdasarkan perhitungan *entropy* dan *gain ratio* dengan parameter *criterion = entropy* dan *max\_depth = 3*. Tahap terakhir, *interpretation*, mengubah struktur pohon keputusan menjadi aturan *IF-THEN* serta menghitung *feature importance* untuk menentukan atribut yang paling berpengaruh.

Pemilihan parameter *max\_depth = 3* dalam penelitian ini didasarkan pada pertimbangan keseimbangan antara kompleksitas model dan interpretabilitas hasil. Penelitian menunjukkan bahwa *decision tree* tanpa pembatasan kedalaman menghasilkan struktur yang terlalu kompleks dengan banyak cabang dan sulit dipahami oleh *stakeholder* non-teknis. Dengan pembatasan kedalaman maksimal tiga level, model yang dihasilkan tetap mampu menangkap pola penting dalam data sambil mempertahankan kesederhanaan struktur yang memudahkan ekstraksi *rules* dan pemahaman hierarki faktor-faktor pembentuk IPM. Validasi terhadap berbagai nilai *max\_depth* (2, 3, 4, dan 5) menunjukkan bahwa *max\_depth = 3* menghasilkan keseimbangan optimal antara jumlah *rules* yang dihasilkan dengan kejelasan interpretasi, menghasilkan 8 *rules* yang dapat dengan mudah diterjemahkan menjadi rekomendasi kebijakan yang lebih tepat sasaran.

## 2. 2 Implementasi Sistem

Sistem berbasis *web* dikembangkan menggunakan Python dan *framework* Flask, dengan antarmuka visual berbasis HTML. Fitur utama sistem meliputi: input data IPM tahunan oleh pengguna (BPS) dan *dashboard* hasil analisis yang menampilkan visualisasi *decision tree* dan grafik *feature importance*, ekstraksi otomatis *rules IF-THEN* (BPS dan Bappeda).

Arsitektur sistem dirancang dengan mempertimbangkan aspek keamanan, skalabilitas, dan kemudahan penggunaan. Sistem mengimplementasikan *role-based access control (RBAC)* yang membedakan hak akses antara Admin BPS, Kabid BPS, dan Bappeda untuk memastikan integritas data dan keamanan informasi. Proses upload data dilengkapi dengan validasi format *file* (CSV/XLSX), validasi struktur data, dan pengecekan ukuran *file* maksimal 16MB untuk mencegah *overload* sistem. Semua proses analisis dilakukan secara otomatis di *server-side* menggunakan *library scikit-learn* untuk algoritma C4.5, *pandas* untuk manipulasi data, dan *matplotlib* untuk visualisasi, sehingga pengguna tidak memerlukan instalasi *software* khusus dan dapat mengakses sistem melalui *web browser* standar. Hasil analisis disimpan dalam format yang terstruktur dan dapat diakses kembali oleh pengguna yang berwenang, mendukung transparansi dalam proses pengambilan keputusan pembangunan daerah.

## 3. HASIL DAN PEMBAHASAN

Hasil penelitian ini diperoleh melalui penerapan algoritma C4.5 terhadap data IPM kabupaten/kota di Provinsi Sumatera Selatan yang telah melalui seluruh tahapan *KDD*. Analisis dilakukan untuk mengidentifikasi atribut yang paling berpengaruh terhadap variasi IPM serta menghasilkan aturan klasifikasi (*rules*) yang dapat digunakan sebagai dasar interpretasi dan pengambilan keputusan.

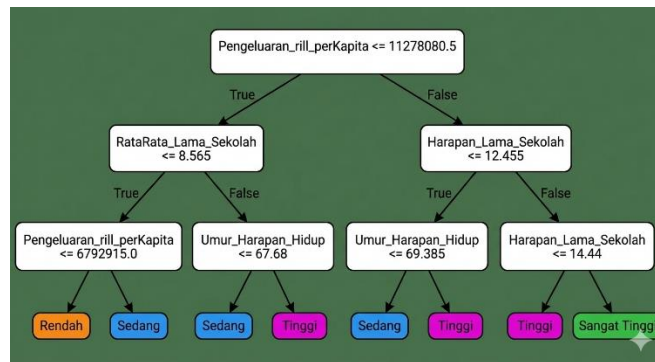
Bagian ini menyajikan hasil pengolahan data, visualisasi model pohon keputusan, serta pembahasan terhadap pola dan hubungan antaratribut yang terbentuk, baik dari sisi teoretis maupun praktis berdasarkan konteks pembangunan daerah. Dataset yang digunakan mencakup periode 15 tahun (2010-2024) dari 17 kabupaten/kota di Sumatera Selatan, memberikan representasi yang komprehensif terhadap dinamika pembangunan manusia di wilayah ini. Setelah proses *cleaning*, dari 270 observasi awal diperoleh 260 sampel yang valid dengan distribusi kategori IPM yang mencerminkan kondisi aktual: 196 sampel kategori "Sedang" (75,4%), 59 sampel kategori "Tinggi" (22,7%), 3 sampel kategori "Rendah" (1,2%), dan 2 sampel kategori "Sangat Tinggi" (0,8%). Distribusi ini menunjukkan bahwa mayoritas wilayah di Sumatera Selatan berada pada tahap pembangunan menengah, dengan hanya sebagian kecil yang telah mencapai kategori sangat tinggi atau masih tertinggal pada kategori rendah, yang konsisten dengan kondisi pembangunan daerah di Indonesia pada umumnya. Berikut sampel data yang digunakan.

	Umur_Harapan_Hidup	Harapan_Lama_Sekolah	RataRata_Lama_Sekolah	Pengeluaran_riil_perKapita	Indeks_Pembangunan_Manusia
0	68,34	11,03	7,82	8536200.0	64,44
1	67,54	11,66	8,38	8029220.0	64,13
2	67,94	9,79	6,74	8693740.0	61,04
3	67,67	10,59	7,49	7881250.0	62,12
4	63,92	11,71	8,28	7963400.0	62,31
5	66,79	9,93	7,06	7887950.0	59,69
6	67,97	10,22	7,51	7584010.0	61,79
7	68,17	10,21	7,02	7293300.0	60,31
8	65,68	10,76	7,45	6214010.0	58,88
9	67,65	10,78	6,91	9557600.0	63,36

Gambar 1. Sampel Data

Model pohon keputusan dengan kedalaman maksimal 3 level terbentuk pengeluaran riil per kapita sebagai atribut akar (*root node*).

## Identifikasi Faktor Dominan yang Mempengaruhi Indeks Pembangunan Manusia Sumatera Selatan Menggunakan Algoritma C4.5



**Gambar 2.** Pohon Keputusan

Adapun delapan aturan (*rules*) *IF-THEN* yang terbentuk dari pohon keputusan pada gambar 2, antara lain:

1. IF Pengeluaran\_riil\_perKapita  $\leq$  11278085 AND RataRata\_Lama\_Sekolah  $\leq$  8.57 AND Pengeluaran\_riil\_perKapita  $\leq$  6792915 THEN Rendah
2. IF Pengeluaran\_riil\_perKapita  $\leq$  11278085 AND RataRata\_Lama\_Sekolah  $\leq$  8.57 AND Pengeluaran\_riil\_perKapita  $>$  6792915 THEN Sedang
3. IF Pengeluaran\_riil\_perKapita  $\leq$  11278085 AND RataRata\_Lama\_Sekolah  $>$  8.57 AND Umur\_Harapan\_Hidup  $\leq$  67.68 THEN Sedang
4. IF Pengeluaran\_riil\_perKapita  $\leq$  11278085 AND RataRata\_Lama\_Sekolah  $>$  8.57 AND Umur\_Harapan\_Hidup  $>$  67.68 THEN Tinggi
5. IF Pengeluaran\_riil\_perKapita  $>$  11278085 AND Harapan\_Lama\_Sekolah  $\leq$  12.45 AND Umur\_Harapan\_Hidup  $\leq$  69.38 THEN Sedang
6. IF Pengeluaran\_riil\_perKapita  $>$  11278085 AND Harapan\_Lama\_Sekolah  $\leq$  12.45 AND Umur\_Harapan\_Hidup  $>$  69.38 THEN Tinggi
7. IF Pengeluaran\_riil\_perKapita  $>$  11278085 AND Harapan\_Lama\_Sekolah  $>$  12.45 AND Harapan\_Lama\_Sekolah  $\leq$  14.44 THEN Tinggi
8. IF Pengeluaran\_riil\_perKapita  $>$  11278085 AND Harapan\_Lama\_Sekolah  $>$  12.45 AND Harapan\_Lama\_Sekolah  $>$  14.44 THEN Sangat Tinggi

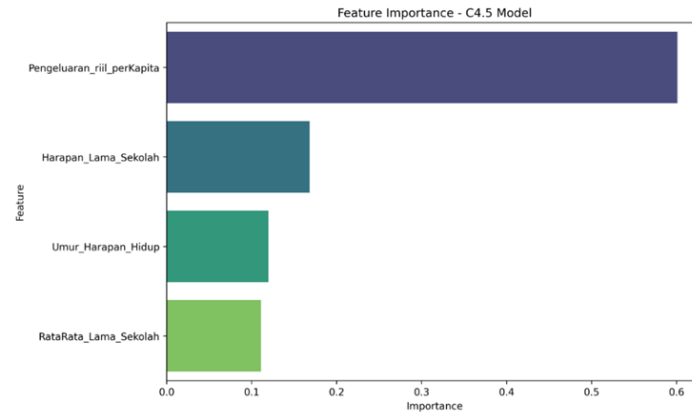
Adapun delapan aturan (*rules*) *IF-THEN* yang terbentuk dari pohon keputusan menunjukkan hubungan antara indikator sosial ekonomi dengan kategori IPM dan mengungkap bahwa peningkatan daya beli masyarakat menjadi kunci peningkatan pembangunan manusia. Struktur *decision tree* yang terbentuk menunjukkan hierarki yang jelas dalam pengaruh faktor-faktor pembentuk IPM. Pada level pertama (*root node*), pengeluaran riil per kapita dengan *threshold* Rp11.278.085 menjadi pembeda utama yang memisahkan daerah dengan daya beli rendah-sedang dari daerah dengan daya beli tinggi.

Pada level kedua, untuk daerah dengan pengeluaran rendah, rata-rata lama sekolah menjadi faktor pembeda dengan *threshold* 8.57 tahun, sementara untuk daerah dengan pengeluaran tinggi, harapan lama sekolah dengan *threshold* 12.45 tahun menjadi faktor kunci.

Pada level ketiga, umur harapan hidup berperan sebagai faktor penentu final dengan berbagai *threshold* (67.68 dan 69.38 tahun) tergantung pada jalur yang dilalui. Pola hierarki ini mengkonfirmasi bahwa meskipun ekonomi menjadi fondasi utama, kombinasi dengan faktor pendidikan dan kesehatan diperlukan untuk mencapai kategori IPM yang lebih tinggi, memberikan panduan strategis bagi pemerintah daerah dalam memprioritaskan alokasi anggaran pembangunan

Aturan-aturan ini merepresentasikan hubungan antara indikator sosial ekonomi dengan kategori IPM dan menunjukkan bahwa peningkatan daya beli masyarakat menjadi kunci peningkatan pembangunan manusia. Dari hasil perhitungan, diperoleh nilai kontribusi (*feature*

*importance*): pengeluaran riil per kapita (60,09%), harapan lama sekolah (16,83%), umur harapan hidup (11,98%), dan rata-rata lama sekolah (11.10%).



**Gambar 3.** *Feature Importance*

Hasil *feature importance* ini memberikan wawasan penting tentang prioritas pembangunan di Sumatera Selatan. Dominasi pengeluaran riil per kapita (60,09%) mengindikasikan bahwa peningkatan kesejahteraan ekonomi masyarakat harus menjadi fokus utama, yang dapat dicapai melalui program-program seperti pengembangan UMKM, peningkatan produktivitas sektor pertanian, dan pembukaan lapangan kerja baru. Kontribusi kumulatif dari kedua indikator pendidikan mencapai 27,93%, menunjukkan pentingnya investasi pada sektor pendidikan, dengan penekanan lebih besar pada peningkatan akses pendidikan (harapan lama sekolah) dibandingkan peningkatan pencapaian historis. Kontribusi umur harapan hidup sebesar 11,98% menunjukkan bahwa meskipun kesehatan penting, faktor ini lebih berperan sebagai *enabler* yang efeknya optimal ketika kondisi ekonomi dan pendidikan sudah memadai. Temuan ini sejalan dengan teori pembangunan manusia yang menekankan bahwa peningkatan kapabilitas ekonomi menjadi prasyarat untuk mengakses layanan kesehatan dan pendidikan berkualitas.

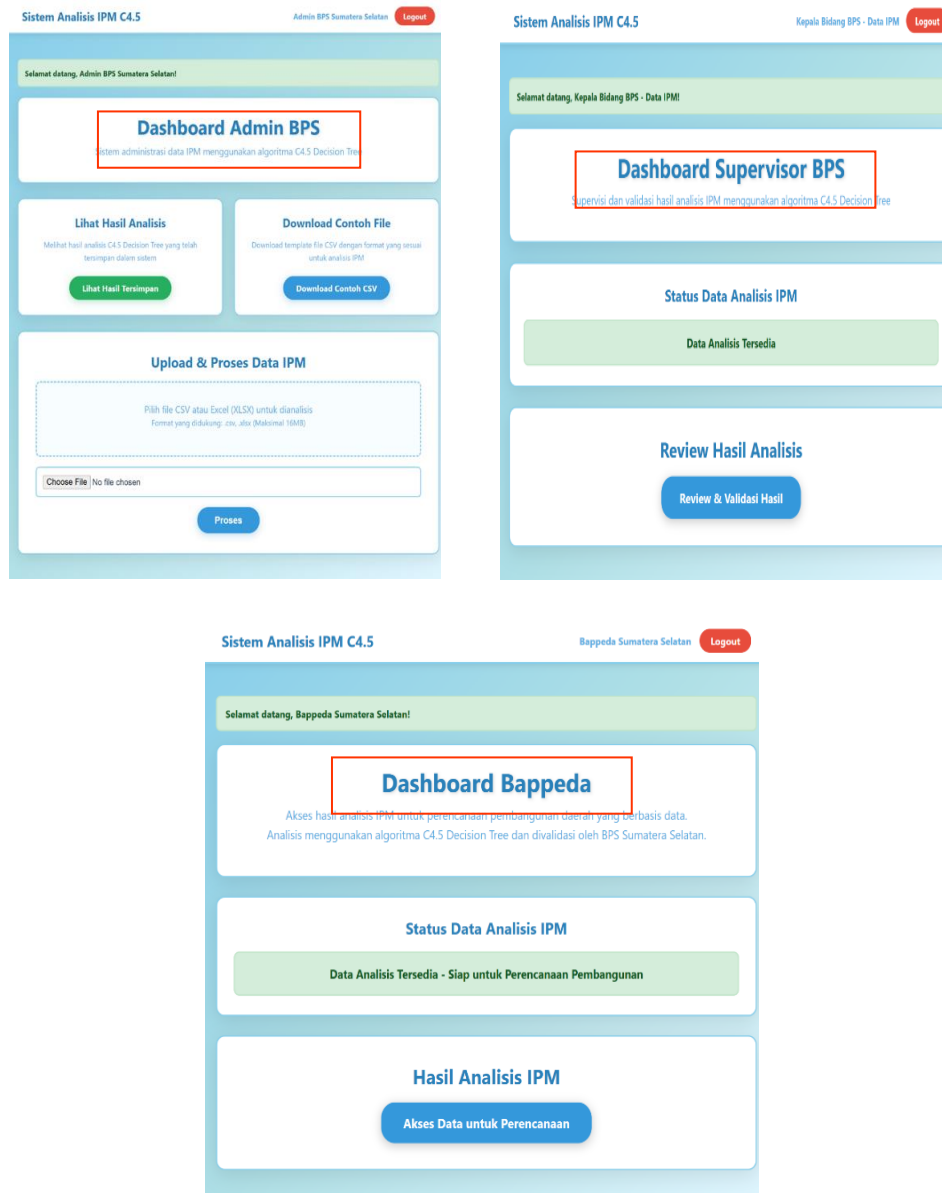
Sistem yang dikembangkan berhasil menampilkan pohon keputusan (*decision tree*) untuk memperlihatkan cabang keputusan dari atribut utama hingga hasil klasifikasi, aturan-aturan yang terbentuk untuk merepresentasikan pohon keputusan dalam bentuk yang lebih mudah dipahami, dan *feature importance* (Gambar 3) untuk menampilkan bobot kontribusi setiap atribut terhadap IPM dalam bentuk grafik batang dan tabel.

Antarmuka sistem terdiri dari tiga peran pengguna, yaitu admin BPS, kepala bidang BPS, dan Bappeda. Masing-masing memiliki hak akses terhadap sistem aplikasi tersebut. Sistem yang dibangun memiliki tiga aktor utama, yaitu Bappeda, Admin BPS, dan Kabid BPS. Masing-masing aktor memiliki hak akses berbeda sesuai perannya dalam proses analisis IPM. Sistem analisis C4.5 dirancang untuk memberikan kemudahan bagi pengguna dalam mengunggah data IPM, melakukan proses analisis, dan melihat hasil analisis. Admin BPS bertugas mengunggah *file* data IPM ke sistem, Bappeda berperan dalam mengakses hasil untuk perencanaan kebijakan, sedangkan Kabid BPS berperan sebagai pihak yang melakukan peninjauan hasil analisis. Setiap aktor memiliki hak akses berbeda sesuai dengan tanggung jawabnya dalam proses pengelolaan data IPM.

Implementasi sistem dilakukan menggunakan antarmuka berbasis *web* agar dapat diakses oleh pihak BPS dan Bappeda dengan mudah. Setiap pengguna memiliki tampilan *dashboard* yang berbeda sesuai dengan perannya. Implementasi sistem dengan *role-based access* ini dirancang untuk mendukung *workflow* kolaboratif dalam proses analisis dan pemanfaatan data IPM. Admin BPS memiliki akses penuh untuk menginput data, memproses analisis, dan mendownload *template*, memastikan bahwa data yang dianalisis selalu *up-to-date* dan akurat. Kabid BPS dapat melakukan *review* dan validasi hasil analisis sebelum data digunakan untuk keperluan strategis. Bappeda sebagai *end-user* dapat mengakses hasil analisis yang telah divalidasi untuk keperluan

## Identifikasi Faktor Dominan yang Mempengaruhi Indeks Pembangunan Manusia Sumatera Selatan Menggunakan Algoritma C4.5

perencanaan pembangunan, dengan visualisasi yang intuitif memudahkan pemahaman terhadap faktor-faktor kunci yang perlu diprioritaskan. Sistem juga menyimpan *history* analisis sehingga memungkinkan perbandingan tren antar periode, mendukung evaluasi efektivitas kebijakan pembangunan yang telah diimplementasikan. Dengan pendekatan ini, sistem tidak hanya berfungsi sebagai tool analisis teknis, tetapi juga sebagai platform kolaborasi antar-institusi dalam ekosistem perencanaan pembangunan daerah. Beberapa halaman pengguna disediakan sesuai peran masing-masing pengguna (Gambar 4).



**Gambar 4.** Tampilan *Dashboard* Admin BPS, Kabid BPS, dan Bappeda

Gambar 4 menampilkan implementasi antarmuka sistem analisis IPM berbasis algoritma C4.5 yang dirancang dengan pembagian peran pengguna, yaitu Admin BPS, *Supervisor* BPS, dan Bappeda. *Dashboard* Admin BPS berfungsi sebagai modul pengelolaan dan pemrosesan data IPM, termasuk pengunggahan dataset dan pelaksanaan analisis secara otomatis menggunakan

algoritma C4.5 Decision Tree. Dashboard Supervisor BPS digunakan untuk melakukan pengawasan serta validasi hasil analisis guna menjamin akurasi dan konsistensi data. Sementara itu, dashboard Bappeda menyediakan akses terhadap hasil analisis IPM yang telah divalidasi sebagai dasar perencanaan, evaluasi, dan pengambilan keputusan pembangunan daerah berbasis data. Rancangan sistem ini mendukung proses analisis yang terstruktur dan efisien dalam evaluasi pembangunan manusia.

### Hasil Analisis C4.5 Decision Tree

[Kembali ke Halaman Utama](#)

Analisis berhasil diselesaikan dengan algoritma C4.5 Decision Tree

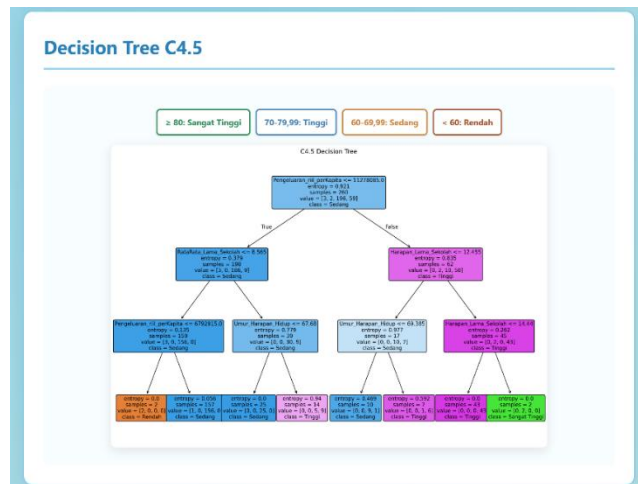
#### Data Yang Diproses

Data berikut digunakan untuk training model C4.5 Decision Tree

Umur Harapan Hidup	Harapan Lama Sekolah	Rata-Rata Lama Sekolah	Pengeluaran riil perKapita	Indeks Pembangunan Manusia	IPM Kategori
58.34	11.03	7.82	6536200.0	54.44	Sedang
57.54	11.66	8.36	6030200.0	64.13	Sedang
57.94	9.79	6.74	6691240.0	61.04	Sedang
57.67	10.59	7.49	7881250.0	62.12	Sedang
61.92	11.71	8.28	7963400.0	62.87	Sedang
66.79	9.93	7.06	7887950.0	59.69	Rendah
57.97	10.22	7.51	7584010.0	61.79	Sedang
66.17	10.21	7.82	7293800.0	60.31	Sedang
55.68	10.76	7.45	6214010.0	58.88	Rendah

total: 260 data siap dianalisis

\* Data di atas akan menghasilkan decision tree, rules, dan feature importance di bawah ini



### Rule IF-THEN Decision Tree

≥ 80: Sangat Tinggi
70-79,99: Tinggi
60-69,99: Sedang
< 60: Rendah

Total 8 Rule berhasil diekstrak dari pohon keputusan

- Rule 1:**  
IF Pengeluaran\_ril\_perKapita <= 11278085 AND RataRata\_Lama\_Sekolah <= 8.57 AND Pengeluaran\_ril\_perKapita <= 6792915 THEN Rendah
- Rule 2:**  
IF Pengeluaran\_ril\_perKapita <= 11278085 AND RataRata\_Lama\_Sekolah <= 8.57 AND Pengeluaran\_ril\_perKapita > 6792915 THEN Sedang
- Rule 3:**  
IF Pengeluaran\_ril\_perKapita <= 11278085 AND RataRata\_Lama\_Sekolah > 8.57 AND Umur\_Harapan\_Hidup <= 67.68 THEN Sedang
- Rule 4:**  
IF Pengeluaran\_ril\_perKapita <= 11278085 AND RataRata\_Lama\_Sekolah > 8.57 AND Umur\_Harapan\_Hidup > 67.68 THEN Tinggi
- Rule 5:**  
IF Pengeluaran\_ril\_perKapita > 11278085 AND Harapan\_Lama\_Sekolah <= 12.45 AND Umur\_Harapan\_Hidup <= 69.38 THEN Sedang
- Rule 6:**  
IF Pengeluaran\_ril\_perKapita > 11278085 AND Harapan\_Lama\_Sekolah <= 12.45 AND Umur\_Harapan\_Hidup > 69.38 THEN Tinggi
- Rule 7:**  
IF Pengeluaran\_ril\_perKapita > 11278085 AND Harapan\_Lama\_Sekolah > 12.45 AND Harapan\_Lama\_Sekolah <= 14.44 THEN Tinggi
- Rule 8:**  
IF Pengeluaran\_ril\_perKapita > 11278085 AND Harapan\_Lama\_Sekolah > 12.45 AND Harapan\_Lama\_Sekolah > 14.44 THEN Sangat\_Tinggi

# Identifikasi Faktor Dominan yang Mempengaruhi Indeks Pembangunan Manusia Sumatera Selatan Menggunakan Algoritma C4.5



Ranking	Fitur	Importance Score	Persentase
1	Pengeluaran_riil_perKapita	0.6009	60.1%
2	Harapan_Lama_Sekolah	0.1683	16.8%
3	Umur_Harapan_Hidup	0.1198	12.0%
4	RataRata_Lama_Sekolah	0.1110	11.1%

Catatan: Semakin tinggi nilai importance, semakin besar pengaruh faktor tersebut dalam menentukan kategori IPM

**Gambar 5.** Halaman Hasil Analisis

Tampilan-tampilan pada Gambar 5 menampilkan hasil analisis berupa data yang diproses, pohon keputusan, *rule IF-THEN*, dan grafik *feature importance*. Berdasarkan hasil implementasi antarmuka, sistem ini mampu memfasilitasi proses analisis IPM dengan baik. Seluruh pengguna dapat menjalankan fungsi masing-masing secara terintegrasi, mulai dari pengunggahan data, pemrosesan menggunakan algoritma C4.5, hingga peninjauan hasil analisis. Tampilan antarmuka yang sederhana dan responsif juga memudahkan pengguna non-teknis seperti staf Bappeda untuk memahami hasil klasifikasi dengan cepat

Pada bagian awal, sistem menampilkan dataset sebanyak 260 sampel yang telah melalui tahap *preprocessing*, memberikan transparansi terhadap basis data yang digunakan dalam model. Implementasi algoritma C4.5 menghasilkan struktur pohon keputusan yang intuitif, di mana atribut Pengeluaran Riil per Kapita muncul sebagai *root node* atau pemecah utama, mengindikasikan signifikansi variabel ekonomi dalam klasifikasi IPM.

Transformasi model dari bentuk grafis ke dalam delapan aturan logika *IF-THEN* bertujuan untuk meningkatkan *interpretability* sistem bagi pengguna manajerial. Melalui representasi ini, pengguna dapat mengidentifikasi alur klasifikasi kategori IPM, mulai dari 'Rendah' hingga 'Sangat Tinggi' tanpa harus memiliki keahlian teknis dalam algoritma pembelajaran mesin. Validasi terhadap pengaruh variabel dipertegas melalui tabel *feature importance*, yang menunjukkan bahwa Pengeluaran Riil per Kapita memiliki kontribusi dominan sebesar 60,1%, disusul oleh variabel pendidikan dan kesehatan. Secara keseluruhan, integrasi antarmuka yang responsif ini tidak hanya berfungsi sebagai alat pengolah data, tetapi juga sebagai instrumen pendukung keputusan (*Decision Support System*) yang efektif bagi staf Bappeda dalam melakukan pemetaan dan intervensi strategis terhadap capaian pembangunan manusia di daerah.

## 4. KESIMPULAN

Penelitian ini berhasil mengidentifikasi faktor dominan yang memengaruhi IPM di Sumatera Selatan menggunakan algoritma C4.5. Hasil utama yang diperoleh yaitu:

1. Faktor ekonomi (pengeluaran riil per kapita) merupakan atribut paling dominan (60,09%), diikuti dimensi pendidikan dan kesehatan.
2. Model C4.5 menghasilkan delapan aturan *IF-THEN* yang menunjukkan hubungan langsung antara atribut ekonomi dan kategori IPM.
3. Sistem berbasis *web* yang dikembangkan mampu menampilkan hasil analisis dan visualisasi *feature importance* secara otomatis.

Temuan ini memberikan kontribusi praktis bagi pemerintah daerah dalam merancang kebijakan berbasis bukti (*evidence-based policy*) untuk peningkatan kualitas pembangunan manusia di Sumatera Selatan.

Implikasi praktis dari penelitian ini adalah tersedianya panduan prioritas pembangunan yang terukur dan berbasis data untuk Sumatera Selatan. *Decision rules* yang dihasilkan dapat digunakan untuk mengidentifikasi posisi masing-masing kabupaten/kota dalam spektrum pembangunan manusia dan menentukan intervensi yang paling tepat. Sebagai contoh, daerah dengan IPM kategori "Sedang" yang memiliki pengeluaran per kapita rendah namun tingkat

pendidikan cukup baik (sesuai *rule* 3 dan 4) dapat diprioritaskan untuk program peningkatan ekonomi masyarakat seperti pengembangan industri kreatif atau pariwisata berbasis komunitas. Sementara itu, daerah dengan ekonomi baik namun akses pendidikan terbatas (*rule* 5 dan 6) perlu diprioritaskan untuk pembangunan infrastruktur pendidikan dan program beasiswa. Sistem yang dikembangkan juga dapat diadopsi oleh provinsi lain di Indonesia dengan melakukan pada data input, menjadikannya sebagai model untuk analisis IPM berbasis data mining di tingkat daerah, mendukung implementasi pembangunan berbasis data di era digitalisasi pemerintahan.

## 5. SARAN

Sistem analisis IPM berbasis algoritma C4.5 yang dikembangkan dalam penelitian ini berpotensi diimplementasikan oleh BPS sebagai alat bantu analisis dan evaluasi pembangunan manusia. Melalui kemampuan otomatisasi, sistem dapat digunakan untuk analisis rutin terhadap data IPM yang dipublikasikan secara berkala, sehingga memudahkan identifikasi tren, pola, serta faktor dominan pembangunan manusia secara lebih efisien dan objektif. Hasil analisis tersebut juga dapat mendukung BPS dalam merumuskan rekomendasi kebijakan yang lebih tepat sasaran.

Untuk pengembangan selanjutnya, sistem ini dapat ditingkatkan melalui perluasan dataset dengan penambahan variabel pendukung seperti tingkat pengangguran, kemiskinan, dan PDRB per sektor, serta penerapan algoritma *ensemble* seperti *Random Forest* atau *Gradient Boosting* guna membandingkan kinerja dan meningkatkan *robustness* analisis. Selain itu, pengembangan modul prediktif untuk memproyeksikan IPM berdasarkan skenario kebijakan, integrasi dengan sistem informasi geografis (GIS) untuk visualisasi spasial, serta penambahan fitur *early warning system* untuk deteksi dini penurunan indikator IPM juga menjadi arah pengembangan yang relevan. Kolaborasi berkelanjutan antara akademisi, BPS, dan Bappeda perlu diperkuat agar sistem tetap adaptif terhadap kebutuhan perencanaan pembangunan daerah.

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Badan Pusat Statistik Provinsi Sumatera Selatan atas penyediaan data, serta kepada Universitas Islam Negeri Raden Fatah Palembang atas dukungan dalam pelaksanaan penelitian ini. Terima kasih juga kepada dosen pembimbing atas arahan selama proses penyusunan karya ilmiah ini.

## DAFTAR PUSTAKA

- [1] UNDP, *Human Development Report*. New York, 2024.
- [2] Kementerian PPN/ Bappenas, “Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020-2024,” 2019.
- [3] BPS, “Indeks Pembangunan Manusia (IPM) Provinsi Sumatera Selatan tahun 2024 mencapai 73,84 meningkat 0,66 poin atau 0,90 persen dibandingkan tahun sebelumnya yang sebesar 73,18,” Badan Pusat Statistik Provinsi Sumatera. Accessed: Jul. 02, 2025. [Online]. Available: <https://sumsel.bps.go.id/id/pressrelease/2024/12/02/814/indeks-pembangunan-manusia--ipm--provinsi-sumatera-selatan-tahun-2024-mencapai-73-84-meningat-0-66-poin-atau-0-90-persen-dibandingkan-tahun-sebelumnya-yang-sebesar-73-18.html>
- [4] BPS, *Indeks Pembangunan Manusia 2023*, vol. 18. Badan Pusat Statistik, 2024.
- [5] Y. W. Simarmata and D. D. Iskandar, “Pengaruh Pengeluaran Pemerintah, Investasi, Jumlah Penduduk, Kemiskinan Terhadap Pertumbuhan Ekonomi dan IPM: Analisa Two Stage Least Square untuk Kasus Indonesia,” *J. Din. Ekon. Pembang.*, vol. 5, no. 1, pp. 78–94, 2022, doi: 10.14710/jdep.5.1.78-94.
- [6] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques*, vol. 9, no. 27 Special Issue. 2012.

## Identifikasi Faktor Dominan yang Mempengaruhi Indeks Pembangunan Manusia Sumatera Selatan Menggunakan Algoritma C4.5

- 
- [7] M. Sari and P. Purnadi, "Pemodelan Indeks Pembangunan Manusia Provinsi Jawa Barat, Jawa Timur Dan Jawa Tengah Tahun 2019 Dengan Menggunakan Metode Regresi Logistik Ordinal," *J. Gaussian*, vol. 10, no. 1, pp. 149–158, 2021, doi: 10.14710/j.gauss.v10i1.30022.
- [8] J. R. Quinlan, *C4.5: Programs for Machine Learning*. California: Morgan Kaufmann, 1988. [Online]. Available: <https://www.semanticscholar.org/paper/Programs-for-Machine-Learning-Publishers/dec46fb79cde2823c13dea9a6b604e1dfcb435cd>
- [9] D. Jollyta, Prihandoko, A. Hajjah, E. Haerani, and M. Siddik, *Algoritma Klasifikasi untuk Pemula*, 1st ed. Yogyakarta: Deepublish, 2023.
- [10] W. Sirait and N. Azizah, "Prediction of the Human Development Index for Equitable Development in West Sumatra Province Using the C4.5 Algorithm," *Sink. J. dan Penelit. Tek. Inform.*, vol. 8, no. 4, pp. 2179–2189, 2023, doi: 10.33395/sinkron.v8i4.12968.
- [11] A. Srirahayu and L. S. Pribadie, "Review Paper Data Mining Klasifikasi Data Mining," *J. Ilm. Inform. Glob.*, vol. 14, no. 1, pp. 7–12, 2023, doi: 10.36982/jiig.v14i1.2981.
- [12] N. W. O. Pratiwi, N. W. Utami, and I. G. J. E. Putra, "Klasifikasi Penentuan Penerima Bantuan Sosial Tunai (BST) Menggunakan Algoritma C4.5 Di Desa Keramas, Gianyar, Bali," *JINTEKS J. Inform. Teknol. dan Sains*, vol. 4, no. 3, pp. 101–107, 2022, doi: 10.51401/jinteks.v4i3.1667.
- [13] R. V. Siregar, P. K. D. Lubis, F. Azkiah, and A. Putri, "Peran Penting Pendidikan dalam Pembentukan Sumber Daya Manusia Cerdas di Era Digitalisasi Menuju Smart Society 5.0," *IJEDR Indones. J. Educ. Dev. Res.*, vol. 2, no. 2, pp. 1408–1418, 2024, doi: 10.57235/ijedr.v2i2.2621.
- [14] J. Mahendra, Alimin, and H. Pasrizal, "Pengaruh Tingkat Kesehaan , Pendidikan, dan Pendapatan Terhadap Produktivitas Tenaga Kerja di Kota Padang Panjang," *J. Ekon. Syariah Pelita Bangsa*, vol. 10, no. 01, pp. 122–133, 2025, doi: 10.37366/jespb.v10i01.2367.
- [15] I. Arofah and S. Rohimah, "Analisis Jalur Untuk Pengaruh Angka Harapan Hidup, Harapan Lama Sekolah, Rata-Rata Lama Sekolah Terhadap Indeks Pembangunan Manusia Melalui Pengeluaran Riil Per Kapita di Provinsi Nusa Tenggara Timur," *J. Sainika Unpam*, vol. 2, no. 1, pp. 76–87, 2019, doi: 10.32493/jsmu.v2i1.2920.
-