

Penerapan Metode Decision Tree dalam Klasifikasi Tingkat Kemiskinan di Pulau Jawa

Muhammad Hisham¹, Nazla Virza Rahman², Rangga Widyo Putra³,
Veti Apriana⁴, Hanafi Eko Darono⁵

^{1,2,3,5} Sistem Informasi, Universitas Bina Sarana Informatika

⁴ Teknik Industri, Universitas Bina Sarana Informatika

¹ samsc405@gmail.com, ² virzarahman@gmail.com, ³ classrangga0306@gmail.com, ⁴ veti.vta@bsi.ac.id,

⁵ hanafi.haf@bsi.ac.id

INFORMASI ARTIKEL

Riwayat Artikel:

Diterima 09/02/2026

Direvisi 15/05/2026

Diterbitkan 30/06/2026

Kata Kunci :

Kemiskinan Pulau Jawa,
Provinsi, Decision Tree,
Orange, Klasifikasi

*Keyword : Poverty in Java
Island, Province, Decision
Tree, Orange,
Classification*

ABSTRAK

Indonesia merupakan salah satu negara yang memiliki jumlah penduduk miskin terbanyak, berada dalam daftar 100 negara terendah yang dianggap negara miskin di dunia. Badan Pusat Statistik mengatakan bahwa sebagian besar orang yang miskin di Indonesia tinggal di Pulau Jawa dengan Buku Statistik: Penduduk dan Kemiskinan Sektor Pertanian Tahun 2023, tentang Jumlah Penduduk Miskin di Pulau Jawa dari Tahun 2018-2022. Data yang digunakan dalam studi ini berasal dari sumber data sekunder yang telah terverifikasi kredibilitasnya, dengan menggunakan pendekatan kuantitatif untuk mengkaji masalah kemiskinan di Provinsi atau Kabupaten/Kota yang ada di Pulau Jawa. Metode yang digunakan adalah *Decision Tree*, sebagai representasi pemodelan dari permasalahan yang menunjukkan keputusan dan solusi. Hasilnya menunjukkan bahwa ada beberapa variabel yang paling berpengaruh pada kemiskinan, seperti persentase penduduk miskin dengan persentase 81.2%, dan hasil dari model *confusion matrix* yang digunakan menunjukkan kinerja yang baik dalam hal akurasi dan lainnya dengan persentase 0.891%. Metode *Decision Tree* ini sangat efektif untuk menemukan pola kemiskinan yang kompleks dengan mempertimbangkan faktor-faktor seperti ekonomi, pendidikan, kesehatan, dan pengangguran. Dengan cara ini, metode ini dapat membantu pemerintah membuat pilihan yang lebih baik dalam mengatasi isu kemiskinan.

ABSTRACT

Indonesia is one of the countries with the largest number of poor people, being on the list of the 100 lowest countries considered poor countries in the world. The Central Statistics Agency (BPS) stated that most of the poor people in Indonesia live on the island of Java, according to the Statistics Book: Population and Poverty in the Agricultural Sector in 2023, regarding the Number of Poor People in Java from 2018-2022. The data used in this study comes from secondary data sources whose credibility has been verified, using a quantitative approach to examine poverty issues in Provinces or Regencies/Cities on the island of Java. The method used is the Decision Tree, as a modeling representation of problems that show decisions and solutions. The results show that there are several variables that most influence poverty, such as the percentage of the poor population with a percentage of 81.2%, and the results of the confusion matrix model used show good performance in terms of accuracy and others with a percentage of 0.891%. This Decision Tree method is very effective in finding complex poverty patterns by considering factors such as the economy, education, health, and unemployment. In this way, this method can help the government make better choices in addressing poverty issues.

Penulis Koresponden:

Veti Apriana

Email: veti.vta@bsi.ac.id

Pendahuluan

Kemiskinan merupakan situasi dimana seseorang, keluarga, atau kelompok masyarakat tidak memiliki akses yang cukup terhadap sumber daya dan ekonomi untuk memenuhi kebutuhan dasar mereka (Purwanti, 2024). Bank Dunia mendefinisikan kemiskinan sebagai ketidakmampuan untuk mendapatkan makanan, tempat tinggal, obat-obatan, tidak dapat bersekolah, tidak bisa membaca, tidak memiliki pekerjaan, kondisi sanitasi yang buruk, kekurangan kebebasan dalam berpolitik, serta kekhawatiran tentang kehidupan masa depan (Ariefudin et al., 2023). Sudah banyak penelitian yang membahas tentang kemiskinan, namun studi tentang kemiskinan dengan sudut pandang yang lebih luas masih jarang dibahas oleh beberapa pihak. Penelitian oleh Jacobus dan rekan-rekannya pada tahun 2019, menunjukkan bagaimana karakteristik sosial ekonomi sebuah rumah tangga bisa mempengaruhi tingkat kemiskinan. Faktor faktor sosial ekonomi pada rumah tangga yang mempengaruhi kemiskinan sebenarnya sama saja, tidak tergantung pada konsep kemiskinan yang digunakan, baik itu kemiskinan uang maupun kemiskinan yang dilihat dari berbagai sisi lainnya (Salam et al., 2022). Selain itu, kenaikan harga yang tinggi, khususnya pada harga makanan, telah menyebabkan daya beli masyarakat menurun, serta menambah kesulitan bagi masyarakat miskin yang pengeluarannya sebagian besar dihabiskan untuk konsumsi bahan pangan (Tanjung et al., 2025). Masalah kemiskinan menjadi salah satu isu utama yang sangat diperhatikan oleh pemerintah di negara manapun. Salah satu aspek penting untuk mendukung upaya penanggulangan kemiskinan adalah dengan tersedianya data kemiskinan yang akurat. Data kemiskinan yang baik dapat digunakan untuk menilai kebijakan pemerintah terkait kemiskinan, membandingkan tingkat kemiskinan antar waktu dan daerah, serta menentukan target penduduk miskin dengan tujuan untuk memperbaiki keadaan mereka (Direktorat Statistik Ketahanan Sosial, 2024).

Indonesia merupakan salah satu negara dengan tingkat kemiskinan yang cukup tinggi. Oleh karena itu, pemerintah harus dengan cepat mengambil tindakan untuk mengatasi faktor-faktor yang menyebabkan munculnya kemiskinan (Hendayanti & Nurhidayati, 2021). Saat ini, Indonesia termasuk ke dalam peringkat 100 besar, negara paling miskin di dunia. Menurut Data dari *World Population Review (WPR)*, Indonesia berada di peringkat 73 sebagai negara termiskin di dunia. Kemiskinan ini menyebabkan kualitas sumber daya manusia menurun dikarenakan orang-orang tidak mampu mengakses atau mendapatkan pendidikan, kesehatan, dan gizi yang baik yang berimbas pada melemahnya produktivitas pembangunan manusia (Sangaji & Triono, 2023). Indeks Pembangunan Manusia (IPM) merupakan ukuran yang meliputi tiga aspek utama, yaitu Umur Harapan Hidup (Kesehatan), Rata-Rata dan Harapan Lama Sekolah (Pendidikan), dan Pengeluaran per Kapita (Ekonomi). Pembangunan manusia memiliki tujuan untuk memastikan setiap orang bisa mengakses dan memenuhi kebutuhan ini untuk mencapai kehidupan yang baik. Selain itu, pembangunan manusia juga bertujuan untuk menciptakan kesetaraan dalam mengakses sumber daya (Statistik, 2025). Sumber daya ini tidak hanya terkait pada aspek keuangan, tetapi juga meliputi berbagai jenis kekayaan yang dapat membantu meningkatkan kesejahteraan hidup masyarakat secara keseluruhan. Indonesia sendiri merupakan negara berkembang dengan kekayaan sumber daya alam, tetapi juga memiliki jumlah penduduk miskin terbanyak. Oleh karena itu, kemiskinan harus menjadi pusat perhatian pemerintah Indonesia karena menjadi salah satu hambatan untuk bisa maju dan berkembang (Triwulandari et al., 2023).

Pulau Jawa adalah pulau yang paling banyak dituju oleh orang-orang untuk berpindah tempat tinggal. Meskipun luas Pulau Jawa hanya 6,8% dari seluruh wilayah Indonesia, pulau ini ternyata dihuni oleh 57,5% penduduk Indonesia. Jika dibandingkan dengan pulau lainnya seperti Pulau Sumatera yang memiliki luas 25,2%, tetapi pulau ini hanya dihuni oleh 21,3% penduduk Indonesia. Ini menunjukkan bahwa ketidakmerataan penduduk di Pulau Jawa yang menjadi salah satu alasan terjadinya kemiskinan. Penyebab mengapa penduduk Indonesia memilih untuk pindah ke Pulau Jawa adalah karena infrastrukturnya. Infrastruktur tersebut sangat penting untuk membantu pertumbuhan ekonomi di suatu daerah (Mansur, 2024). Badan Pusat Statistik (BPS) menjelaskan bahwa mayoritas populasi masyarakat yang miskin di Indonesia berada di Pulau Jawa. Hal ini terjadi karena keterbatasan lapangan pekerjaan yang tersedia, sehingga menyebabkan banyak masyarakat tidak dapat memenuhi kebutuhan hidup mereka (Triwulandari et al., 2023).

Pada penelitian ini, penulis memfokuskan secara spesifik untuk mengklasifikasi tingkat kemiskinan di 6 provinsi yang berada di Pulau Jawa, yaitu DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, dan Banten dengan menggunakan metode *Decision Tree*. Dengan mengklasifikasikan tingkat kemiskinan di provinsi tersebut, penelitian ini mampu memetakan tingkat kemiskinan secara lebih terperinci sehingga dapat menjadi dasar rekomendasi kebijakan yang lebih tepat sasaran bagi pemerintah daerah maupun pemerintah pusat. Pada Buku Statistik: Penduduk dan Kemiskinan Sektor Pertanian Tahun 2023 (Wiratno, 2023), Jumlah Penduduk Miskin (Total Jiwa) di Pulau Jawa dari Tahun 2018-2022, dapat dilihat pada tabel dibawah ini:

Tabel 1. Data Statistik Penduduk Miskin di Pulau Jawa Tahun 2018-2022

No	Provinsi	Tahun				
		2018	2019	2020	2021	2022
1.	DKI Jakarta	372.260	326.300	496.840	498.290	494.930
2.	Jawa Barat	3.539.400	3.375.890	4.188.520	4.004.860	4.053.420
3.	Jawa Tengah	3.867.420	3.679.400	4.119.930	3.934.010	3.858.230
4.	DI Yogyakarta	450.250	440.890	503.140	474.490	463.630
5.	Jawa Timur	4.292.150	4.056.000	4.585.970	4.259.60	4.236.510
6.	Banten	668.740	641.420	857.640	852.280	829.660

Sumber: Badan Pusat Statistik (2023)

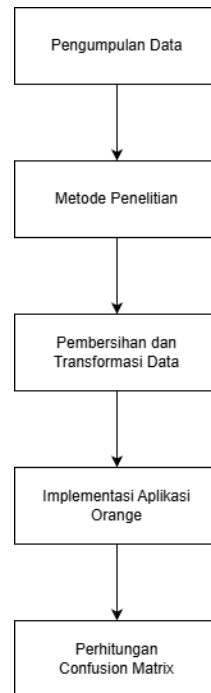
Kemiskinan tetap menjadi tantangan sosial ekonomi yang kompleks di Indonesia, di mana konsentrasi penduduk miskin secara absolut masih didominasi oleh wilayah Pulau Jawa. Meskipun memiliki pertumbuhan ekonomi yang relatif tinggi, ketidakesetaraan antarwilayah di dalam Pulau Jawa sendiri masih sangat signifikan. Oleh karena itu, diperlukan sebuah pendekatan analitis yang tidak hanya mampu menggambarkan peta kemiskinan, tetapi juga dapat mengeksplorasi pola tersembunyi dari faktor-faktor yang memengaruhinya.

Pada penelitian yang dilakukan oleh (Hendayanti & Nurhidayati, 2021), tentang “Klasifikasi Tingkat Keparahan Kemiskinan Provinsi di Indonesia dengan Analisis Diskriminan”, telah mendapatkan kesimpulan bahwa dari 21 Provinsi yang ada di Indonesia, terdapat hanya 12 Provinsi yang memiliki tingkat kemiskinan rendah. Sedangkan, 9 Provinsi lainnya diprediksi berada dalam kategori kemiskinan tinggi. Pada penelitian lainnya, oleh (Tanjung et al., 2025), menunjukkan bahwa Laju Pertumbuhan Penduduk berdampak negatif dan signifikan terhadap Pendapatan per Kapita, dan berdampak positif terhadap peningkatan Persentase Penduduk Miskin. Di sisi lain, Pendapatan per Kapita juga berpengaruh negatif dan signifikan terhadap kemiskinan, yang menunjukkan bahwa jika peningkatan Pendapatan per Kapita naik, maka dapat menurunkan tingkat kemiskinan. Penelitian tersebut didukung oleh penelitian yang dilakukan oleh (Triwulandari et al., 2023), tentang “Analisis Pengaruh Pendapatan Asli Daerah, Pertumbuhan Ekonomi dan Pengangguran terhadap Kemiskinan di Pulau Jawa Tahun 2013-2022”, menjelaskan bahwa pengangguran yang memiliki pengaruh besar terhadap tingkat kemiskinan. Oleh karena itu, penting bagi pemerintah untuk mengetahui tingkat kemiskinan di Indonesia, terutama pemerintah daerah dan provinsi. Dikarenakan hal ini, diperlukan suatu metode untuk mengklasifikasi tingkat kemiskinan di berbagai provinsi di Pulau Jawa dengan mengaitkan beberapa faktor-faktor yang dapat mempengaruhi kemiskinan.

Penelitian terdahulu banyak menggunakan metode statistik konvensional atau pengelompokan (*K-Means Clustering*) yang sering kali hanya berhenti pada tahap pemetaan tanpa memberikan aturan keputusan yang intuitif untuk kebijakan strategis. Metode *Decision Tree* dipilih karena memiliki keunggulan komparatif dalam mentransformasikan data multidimensi yang rumit menjadi struktur hierarki keputusan yang mudah diinterpretasikan oleh pemangku kebijakan. Berbeda dengan metode statistik konvensional, *Decision Tree* berperan dalam mengidentifikasi variabel-variabel paling dominan yang menentukan status kemiskinan suatu wilayah melalui serangkaian aturan keputusan yang terstruktur. Melalui pembentukan pohon keputusan, algoritma ini mampu mengidentifikasi variabel yang paling kontributif hingga variabel pendukung secara objektif berdasarkan nilai *Information Gain* atau *Gini Index*, sehingga menghasilkan model klasifikasi dengan akurasi tinggi dan aturan yang logis. Peranan *Decision Tree* dalam penelitian ini bukan sekadar sebagai alat prediksi, melainkan juga sebagai sarana eksplorasi faktor-faktor kemiskinan yang saling berkaitan, mulai dari dimensi ekonomi, kesehatan, pendidikan, hingga ketenagakerjaan, sehingga menghasilkan pemahaman yang lebih mendalam dan berbasis data tentang pola kemiskinan di tingkat provinsi pada kabupaten/kota.

Tujuan utama dari penelitian ini adalah melakukan klasifikasi tingkat kemiskinan di 6 provinsi yang berada di Pulau Jawa, yaitu DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, dan Banten menggunakan metode *Decision Tree*, dengan memberikan penekanan utama pada variabel Persentase Penduduk Miskin sebagai penentu utama status kemiskinan suatu provinsi di kabupaten/kota. Untuk menghasilkan model klasifikasi yang lebih komprehensif dan akurat, penelitian ini juga mengintegrasikan beberapa variabel pendukung klasifikasi lainnya, yaitu Indeks Pembangunan Manusia (IPM), Umur Harapan Hidup (UHH), Rata-Rata Lama Sekolah (RLS), Pengeluaran Per Kapita, dan Tingkat Pengangguran Terbuka (TPT). Melalui kombinasi variabel utama dan pendukung ini, *Decision Tree* diharapkan mampu menghasilkan pohon keputusan yang presisi dan dapat menghasilkan model klasifikasi yang akurat dan interpretatif untuk mendukung pengambilan keputusan dalam upaya penanggulangan kemiskinan di Pulau Jawa.

Metode Penelitian



Sumber: Penelitian (2026)

Gambar 1. Langkah-Langkah Penelitian

- 1) **Pengumpulan Data:** Setelah memilih tema dan topik permasalahan yang ingin diteliti, langkah berikutnya adalah mengumpulkan sumber data untuk dianalisis.
- 2) **Metode Penelitian:** Setelah mengumpulkan data, langkah berikutnya adalah memilih metode penelitian yang tepat untuk menentukan tujuan dari data yang telah diperoleh.
- 3) **Pembersihan dan Transformasi Data:** Di bagian ini, akan dilakukan pembersihan data untuk menemukan data yang kosong atau tidak valid, serta mengubah data tersebut jika diperlukan agar sesuai dengan kebutuhan metode penelitian yang akan digunakan.
- 4) **Implementasi Aplikasi Orange:** Setelah data dibersihkan, akan dilanjutkan dengan menerapkan metode penelitian yang digunakan dengan menggunakan aplikasi *Data Mining* yang bernama *Orange*.
- 5) **Perhitungan Confusion Matrix:** Tahap ini adalah melakukan perhitungan akurasi dan prediksi dengan menggunakan tabel yang bernama *Confusion Matrix*.

1. Sumber dan Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini didapatkan dari sumber data sekunder yang telah terverifikasi kredibilitasnya. Dalam penelitian ini, penulis menggunakan data yang diperoleh dari *platform* yang bernama *Kaggle*, sebuah situs resmi yang menyediakan dataset terbuka. Proses pengumpulan data secara *online* dari sumber yang sudah ada, seperti *dataset* di *Kaggle* ini, dapat disebut sebagai penelitian sekunder atau *Data Mining*. Hal ini dikarenakan data atau informasi yang diperoleh merupakan data besar (*Big Data*). Penulis menggunakan analisis dari *dataset* yang sangat besar dan kompleks ini untuk menemukan pola, tren, dan keputusan yang tersembunyi dengan mengubahnya dalam bentuk *dataset* kecil.

Populasi dalam *dataset* ini adalah klasifikasi kemiskinan di seluruh Provinsi dan Kabupaten/Kota di Indonesia. Namun, sampel yang digunakan dalam penelitian ini adalah klasifikasi kemiskinan di Provinsi atau Kabupaten/Kota yang berada di Pulau Jawa. Dalam proses mengumpulkan data ini, penulis menggunakan metode dokumentasi, yang mencakup pemilihan variabel, menetapkan batasan atau *threshold* untuk setiap indikator, hingga pembersihan dan transformasi data. Tahapan ini memastikan agar setiap data yang digunakan dapat membantu menentukan target tingkat kemiskinan yang relevan dengan situasi saat ini sehingga selaras dengan kebutuhan penelitian.

2. Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif untuk mengkaji isu permasalahan tentang kemiskinan di Provinsi atau Kabupaten/Kota yang ada di Pulau Jawa melalui analisis data secara sistematis. Salah satu metode modern untuk mengukur tingkat kemiskinan adalah dengan memanfaatkan *Data Mining*. *Data Mining* adalah proses analisis data dalam jumlah besar untuk menemukan pola-pola atau informasi yang tersembunyi (Setyawan et al., 2025). *Data Mining* dapat menggunakan metode seperti regresi, klasifikasi, pengelompokan, dan asosiasi untuk membantu dalam pengambilan keputusan yang lebih baik (Tawakal, Iqbal, Effendi, M.Makmun, Maulana Majid, 2025).

Klasifikasi merupakan salah satu metode yang sering digunakan dalam *Data Mining*. Salah satu teknik yang biasa digunakan untuk metode klasifikasi adalah *Decision Tree*. *Decision Tree*, yang juga dikenal sebagai Pohon Keputusan, adalah representasi pemodelan dari permasalahan yang menunjukkan keputusan dan solusi, serta berbagai keputusan yang membawa pada solusi. Konsep dasar *Decision Tree* adalah dengan mengubah sekumpulan data menjadi model pohon yang menggambarkan aturan-aturan keputusan. Pohon ini juga menunjukkan faktor-faktor kemungkinan atau probabilitas yang akan mempengaruhi pilihan-pilihan keputusan tersebut, dengan perkiraan hasil akhir yang mungkin akan kita dapatkan jika kita memilih keputusan tersebut (Aluf & Fatah, 2025). Metode ini populer digunakan karena hasilnya dapat divisualisasikan dalam bentuk pohon keputusan yang mudah dipahami, bahkan oleh orang-orang yang tidak memiliki pengetahuan teknis. *Decision Tree* bekerja dengan memisahkan data berdasarkan atribut yang memberikan informasi paling signifikan. Dalam penelitian ini, klasifikasi digunakan untuk menentukan tingkat kemiskinan di suatu Provinsi atau Kabupaten/Kota yang ada di Pulau Jawa berdasarkan variabel ekonomi, pendidikan, kesehatan, dan pengangguran. *Decision Tree* dipilih karena dapat mengidentifikasi faktor-faktor tersebut secara bertahap yang paling berpengaruh terhadap tingkat kemiskinan di suatu wilayah.

3. Variabel Penelitian

Variabel yang ada dalam penelitian ini dibagi menjadi dua kelompok utama, yaitu variabel dependen dan variabel independen. Variabel dependen berwujud status tingkat kemiskinan yang akan dibagi ke dalam kategori dan kelas tertentu sesuai dengan data yang tersedia, misalnya miskin atau tidak miskin. Di sisi lain, variabel independen mencakup berbagai indikator yang digunakan dalam kajian tentang topik kemiskinan, misalnya Persentase Penduduk Miskin atau Pengeluaran per Kapita. Setiap indikator tersebut diukur mengikuti standar umum *threshold* yang ditentukan untuk mengukur variabel tingkat kemiskinan seperti yang ada pada tabel dibawah ini:

Tabel 2. *Threshold* Indikator Kemiskinan

Indikator	Threshold
Persentase Penduduk Miskin	Jika lebih dari atau sama dengan 9%, maka dikategorikan miskin.
Indeks Pembangunan Manusia	Jika kurang dari 74%, maka dikategorikan miskin.
Umur Harapan Hidup	Jika kurang dari 74 tahun, maka dikategorikan miskin.
Rata-Rata Lama Sekolah	Jika kurang dari 9 tahun, maka dikategorikan miskin.
Pengeluaran per Kapita	Jika kurang dari Rp 9.600.000 per tahun, maka dikategorikan miskin.
Tingkat Pengangguran Terbuka	Jika lebih dari atau sama dengan 7%, maka dikategorikan miskin.

Sumber: Penelitian (2026)

Pada indikator Persentase Penduduk Miskin, merupakan indikator yang bisa menunjukkan suatu daerah itu miskin atau tidak secara spesifik. Namun, dalam penelitian ini, penulis tidak hanya akan menggunakan indikator Persentase Penduduk Miskin sebagai target utama untuk menentukan tingkat kemiskinan, tetapi juga akan mempertimbangkan indikator pendukung lainnya seperti Umur Harapan Hidup atau Pengeluaran per Kapita. Jika Persentase Penduduk Miskin dan terdapat tiga indikator lainnya

yang dinyatakan sebagai miskin, maka variabel Klasifikasi Kemiskinan akan dikategorikan sebagai miskin.

4. Confusion Matrix

Confusion Matrix adalah tabel yang digunakan dalam *Machine Learning* untuk menilai seberapa baik model klasifikasi bekerja. Tabel ini menunjukkan visualisasi *detail* tentang seberapa efektif model dalam memprediksi kategori yang berbeda dalam dataset. Dengan *Confusion Matrix*, hasilnya dapat dilihat dengan:

- a. *True Positive (TP)*: Jumlah kasus miskin dan diprediksi benar oleh model sebagai kategori miskin.
- b. *True Negative (TN)*: Jumlah kasus tidak miskin dan diprediksi benar oleh model sebagai kategori tidak miskin.
- c. *False Positive (FP)*: Jumlah kasus tidak miskin tetapi diprediksi salah oleh model sebagai kategori miskin.
- d. *False Negative (FN)*: Jumlah kasus miskin tetapi diprediksi salah oleh model sebagai kategori tidak miskin.

Dalam penelitian tentang tingkat kemiskinan, *Confusion Matrix* membantu kita untuk melihat dan memahami seberapa efektif model *Machine Learning* dalam mengidentifikasi daerah yang miskin dan tidak miskin. Ini sangat penting untuk menentukan dengan lebih akurat dan efisien dalam upaya penanggulangan kemiskinan (Danil et al., 2025). Hasil dari model klasifikasi dievaluasi dengan menggunakan beberapa matriks utama yang dihasilkan dari tabel *Confusion Matrix*, yaitu:

a. Accuracy

Accuracy adalah matriks yang menunjukkan seberapa banyak prediksi yang benar dari keseluruhan jumlah data yang telah diprediksi.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Keterangan:

- TP = *True Positive*
- TN = *True Negative*
- FP = *False Positive*
- FN = *False Negative*

b. Precision

Precision adalah perbandingan antara jumlah prediksi yang benar (*True Positive*) dengan keseluruhan jumlah data dari kategori positif (*True Positive* dan *False Positive*).

$$Precision = \frac{TP}{TP+FP}$$

Keterangan:

- TP = *True Positive*
- FP = *False Positive*

c. Recall

Recall adalah perbandingan antara jumlah prediksi yang benar (*True Positive*) dengan keseluruhan jumlah data yang seharusnya positif. (*True Positive* dan *False Negative*).

$$Recall = \frac{TP}{TP+FN}$$

Keterangan:

- TP = *True Positive*

FN = False Negative

d. *F1-Score*

F1-Score adalah matriks yang menggabungkan *Precision* dan *Recall* menjadi satu nilai yang menyeimbangkan kedua hal tersebut, digunakan untuk mengevaluasi dan menilai seberapa baik performa model klasifikasi bekerja pada *dataset* yang tidak seimbang.

$$F1 - Score = \frac{(2 TP)}{(2 TP) + FP + FN}$$

Keterangan:

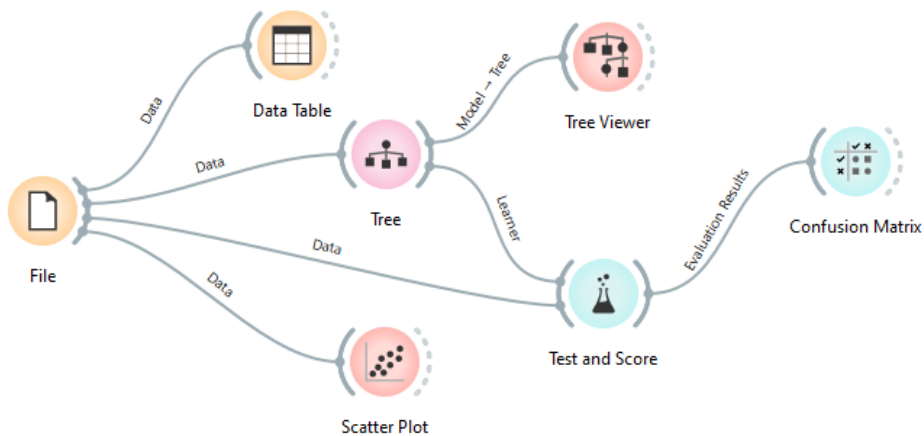
TP = True Positive

FP = False Positive

FN = False Negative

Hasil dan Pembahasan

Pada bagian ini, akan dijelaskan hasil dan pembahasan mengenai implementasi aplikasi *Orange* dengan menerapkan metode *Decision Tree* pada *dataset* Klasifikasi Tingkat Kemiskinan di Pulau Jawa. Proses pemodelan alur kerja beserta penjelasan tentang *widget* yang digunakan dapat dilihat pada Gambar 2:



Sumber: Penelitian (2026)

Gambar 2. Alur Kerja *Orange*

Widget File

Tabel 3. File *Dataset* Klasifikasi Tingkat Kemiskinan di Pulau Jawa

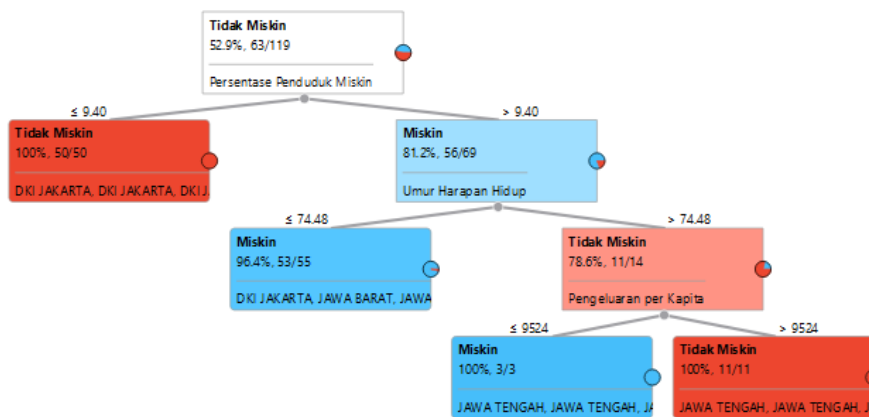
No	Name	Type	Role	Values
1.	Provinsi	(C) categorical	feature	Banten, DI Yogyakarta, DKI Jakarta, Jawa Barat, Jawa Tengah, Jawa Timur
2.	Persentase Penduduk Miskin	(N) numeric	feature	
3.	Indeks Pembangunan Manusia	(N) numeric	feature	
4.	Umur Harapan Hidup	(N) numeric	feature	
5.	Rata-Rata Lama Sekolah	(N) numeric	feature	

6.	Pengeluaran per Kapita	(N) numeric	feature	
7.	Tingkat Pengangguran Terbuka	(N) numeric	feature	
8.	Klasifikasi Kemiskinan	(C) categorical	target	Miskin, Tidak Miskin
9.	Kabupaten/Kota	(S) text	meta	

Sumber: Penelitian (2026)

Pada tabel diatas menunjukkan tampilan *widget file* aplikasi *Orange* menggunakan *dataset* Klasifikasi Tingkat Kemiskinan di Pulau Jawa. dengan kolom “*Name*” yang merupakan nama variabel yang ada di dalam dataset tersebut, kolom “*Type*” yang merupakan tipe data setiap variabelnya, kolom “*Role*” yang merupakan peran apa yang ingin kita berikan pada setiap variabelnya, dan kolom “*Values*” yang merupakan isi data dari setiap variabel. Dalam kasus ini, penulis ingin melihat variabel “Klasifikasi Kemiskinan” sebagai target untuk hasilnya.

Widget Tree Viewer



Sumber: Penelitian (2026)

Gambar 3. Visualisasi *Tree Viewer*

Pada Gambar 3, merupakan hasil visualisasi dari metode *Decision Tree* menggunakan *Orange*. Jika kursor diarahkan ke kolom paling atas, hasilnya akan menunjukkan variabel Klasifikasi Kemiskinan tentang miskin dan tidak miskin, dengan persentase masing-masing dari total 119 data. Kolom selanjutnya akan dibagi dengan variabel *Persentase Penduduk Miskin*, yang jika kurang dari atau sama dengan 9.40%, maka akan dikategorikan “Tidak Miskin” dengan persentase 100% dari total 50 data, dengan dibawahnya adalah nama Provinsinya. Sebaliknya, jika lebih dari 9.40%, maka akan dikategorikan “Miskin” dengan persentase 81.2% dengan 56 dari total 69 data. Begitu juga selanjutnya, untuk variabel “Umur Harapan Hidup”, dan Pengeluaran per Kapita.

Widget Scatter Plot

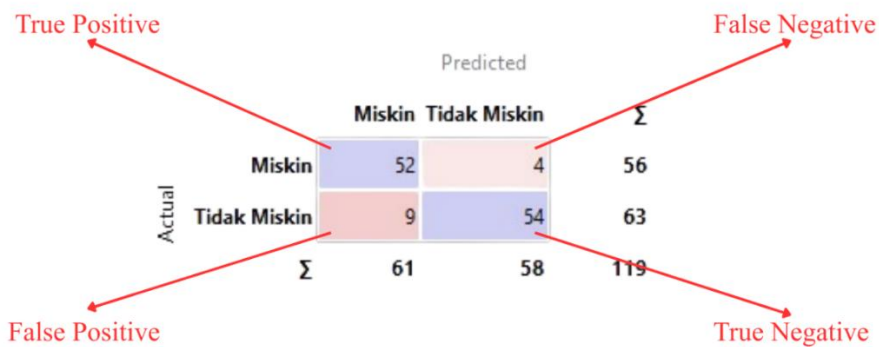
Pada Gambar 4, merupakan *Scatter Plot* yang menampilkan hubungan antara dua variabel dengan menggunakan titik-titik data di atas dua sumbu. Grafik ini digunakan untuk memvisualisasikan pola, tren, korelasi antar variabel, dan data yang tidak biasa dalam sekumpulan data. Berikut adalah hasil visualisasi *Scatter Plot* dari variabel Provinsi dan Klasifikasi Kemiskinan menggunakan *Orange*:



Sumber: Penelitian (2026)

Gambar 4. Scatter Plot

Widget Confusion Matrix



Sumber: Penelitian (2026)

Gambar 5. Confusion Matrix

Pada widget Test & Score, tabel Accuracy, Precision, Recall dan F1-Score dapat dilihat dan menghasilkan perhitungan berikut:

Tabel 4. Test & Score Aplikasi Orange

Model	Accuracy	Precision	Recall	F1-Score
Tree	0.891	0.894	0.891	0.891

Sumber: Penelitian (2026)

- 1) Accuracy
$$Accuracy = \frac{52+54}{52+54+9+4} = 0.890$$
- 2) Precision
$$Precision = \frac{52}{52+9} = 0.852$$
- 3) Recall
$$Recall = \frac{52}{52+4} = 0.928$$
- 4) F1-Score
$$F1 - Score = \frac{(2 \times 52)}{(2 \times 52) + 9 + 4} = 0.888$$

Dari hasil perhitungan diatas, dapat dilihat bahwa hasil perhitungan dengan tabel *Test & Score* aplikasi *Orange* memiliki hasil yang sedikit berbeda jika dilakukan dengan perhitungan secara manual. Namun, hasilnya tetap dianggap mendekati karena memiliki nilai yang tidak terlalu jauh berbeda.

Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, penelitian ini berhasil mengimplementasikan metode *Decision Tree* untuk mengklasifikasikan tingkat kemiskinan di 6 provinsi yang ada di Pulau Jawa, yaitu DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, dan Banten. Model *Decision Tree* yang terbentuk mampu mengekstrak pola keputusan yang jelas mengenai faktor-faktor yang paling memengaruhi kemiskinan di wilayah-wilayah tersebut. Hasil klasifikasi menunjukkan bahwa penekanan pada variabel Persentase Penduduk Miskin, Umur Harapan Hidup, dan Pengeluaran Per Kapita memegang peranan krusial sebagai indikator utama pembeda tingkat kemiskinan antar Kabupaten/Kota di suatu provinsi. Pohon keputusan yang dihasilkan memberikan kontribusi teoritis dan praktis berupa aturan klasifikasi yang dapat digunakan oleh pemerintah daerah di enam provinsi yang ada di Pulau Jawa untuk merancang strategi intervensi penurunan kemiskinan yang lebih tepat sasaran, efektif, dan berbasis data konkrit. Selain itu, integrasi variabel pendukung seperti Indeks Pembangunan Manusia (IPM), Rata-Rata Lama Sekolah, dan Tingkat Pengangguran Terbuka terbukti memperkuat akurasi model dalam memetakan karakteristik kemiskinan secara multidimensional. Pengujian performa model yang dievaluasi menggunakan *confusion matrix* menunjukkan tingkat akurasi sebesar 0.891, presisi sebesar 0.894, recall sebesar 0.891, dan Skor F-1 sebesar 0.891 dalam mengklasifikasikan tingkat kemiskinan.

Saran untuk penelitian selanjutnya dapat menambah variabel-variabel lain seperti Pertumbuhan Ekonomi Daerah, Persentase Rumah Tangga yang Memiliki Akses Terhadap Sanitasi atau Air Minum yang layak, atau Dampak Pandemi untuk meningkatkan generalisasi model di Pulau Jawa. Penelitian ini juga dapat mengintegrasikan *Orange* dengan alat visualisasi lain seperti *Power BI* untuk analisis yang lebih interaktif dan mendalam, sesuai dengan minat pengguna pada visualisasi data. Pemerintah dianjurkan untuk menggunakan model ini agar bisa mengutamakan intervensi di daerah yang memiliki label "Miskin" tinggi.

Referensi

- Aluf, M. W., & Fatah, Z. (2025). Klasifikasi Algoritma Decision Tree Untuk Tingkat Kemiskinan Di Indonesia. *Journal Of Computer Science And Technology (JOCSTEC)*, 3(1), 55–62. <https://doi.org/10.59435/jocstec.v3i1.440>
- Ariefudin, R., Alfarizzi, M., Machmudah, N., Aminah, E., Febiyanti, Y., Program Studi Matematika, M., Matematika, J., & Matematika dan Ilmu Pengetahuan, F. (2023). Analisis Diskriminan untuk Klasifikasi Tingkat Kemiskinan di Perkotaan Menurut Provinsi Berdasarkan Bagian Wilayah di Indonesia Tahun 2022 Discriminant Analysis for Classification of Poverty Levels in Urban Areas by Provinces Based on Regions in Indonesia. *Prosiding Seminar Nasional Matematika, Statistika, Dan Aplikasinya 2023*, 214.
- Danil, S., Rahaningsih, N., Dana, R. D., & Cirebon, K. (2025). *Model Pada Kab / Kota Di Indonesia Menggunakan Metode*. 13(2), 829–835.
- Direktorat Statistik Ketahanan Sosial. (2024). *Data Dan Informasi Kemiskinan Kabupaten/Kota Di Indonesia Tahun 2024*. Badan Pusat Statistik. <https://www.bps.go.id/publication/2024/11/29/d2848c3990f081182125a416/data-dan-informasi-kemiskinan-kabupaten-kota-tahun-2024.html>

- Hendayanti, N. P. N., & Nurhidayati, M. (2021). Klasifikasi Tingkat Keparahan Kemiskinan Provinsi Di Indonesia Dengan Analisis Diskriminan. *Math Educa Journal*, 5(1), 14–21. <https://doi.org/10.15548/mej.v5i1.2510>
- Mansur, Y. (2024). Analisis Perkembangan Penduduk Miskin, Karakteristik Kemiskinan dan Kedalaman Kemiskinan di Indonesia. *Jurnal EMT KITA*, 8(1), 18–31. <https://doi.org/10.35870/emt.v8i1.1930>
- Purwanti, E. (2024). Analisis Deskriptif Profil Kemiskinan Indonesia Berdasarkan Data BPS Tahun 2023. *AKADEMIK: Jurnal Mahasiswa Humanis*, 4(1), 1–10. <https://doi.org/10.37481/jmh.v4i1.653>
- Salam, A., Pratomo, D. S., & Saputra, P. M. A. (2022). Analisis kemiskinan pada rumah tangga di Jawa Timur melalui pendekatan multidimensi dan moneter. *Jurnal Kependudukan Indonesia*, 16(2), 127. <https://doi.org/10.14203/jki.v16i2.480>
- Sangaji, R. C., & Triono, T. A. (2023). Faktor Mempengaruhi Tingkat Kemiskinan di Indonesia. *Jurnal of Society Bridge*, 1, 59–67. <https://www.bk3s.org/ojs/index.php/jsb>
- Setyawan, A., Fitriani, A., Rilvani, E., Bangsa, U. P., & Bekasi, K. (2025). *Klasifikasi Kemiskinan Di Indonesia Dengan Decision Tree Menggunakan Rapidminer*. 3(7).
- Statistik, D. A. dan P. (2025). *Indeks Pembangunan Manusia 2024*. Badan Pusat Statistik. <https://www.bps.go.id/publication/2025/05/15/dd9a3ce7dfae1c733e46338f/indeks-pembangunan-manusia-2024.html>
- Tanjung, A. A., Ariza, D., Nababan, F., Siboro, R. P., Saputra, Z. E., Nasution, A. R., & Hidayat, N. (2025). Analisis Pengaruh Jumlah Penduduk Terhadap Kemiskinan Melalui Pendapatan Per Kapita Sebagai Variabel Mediasi. *Jurnal Publikasi Ekonomi Dan Akuntansi*, 5(2), 1–15. <https://doi.org/10.51903/jupea.v5i2.3828>
- Tawakal, Iqbal, Effendi, M.Makmun, Maulana Majid, A. (2025). *ANALISIS TINGKAT KEMISKINAN DENGAN ALGORITMA K-MEANS Abstraksi Keywords : Pendahuluan Tinjauan Pustaka*. 7(1), 112–119.
- Triwulandari, B. B., Boedirochminarni, A., & Firmansyah, M. (2023). Analisis Pengaruh Pendapatan Asli Daerah, Pertumbuhan Ekonomi dan Pengangguran terhadap Kemiskinan di Pulau Jawa Tahun 2013-2022. *JSHP : Jurnal Sosial Humaniora Dan Pendidikan*, 7(2), 175–187. <https://doi.org/10.32487/jshp.v7i2.1747>
- Wiratno, O. (2023). *Statistik: Penduduk dan Kemiskinan Sektor Pertanian Tahun 2023*. Pusat Data dan Sistem Informasi Pertanian Sekretariat Jenderal. https://satudata.pertanian.go.id/assets/docs/publikasi/Buku_Statistik_Penduduk_dan_Kemiskinan_Sektor_Pertanian_Tahun_2023_comp.pdf