



Analisis Dataset Pemilu KPU Menggunakan Algoritma Decision Tree Di Rapidminer

Maulana Fansyuri^{1*}, Murni¹², Teguh Arifin³, Ahmad Farhan⁴, Ramdan Tiplahi⁵, Aldi Muhammad Rifaldi⁶

^{1,2,3,4,5,6}Program Studi S1 Teknik Informatika ,Fakultas Ilmu Komputer ,Universitas Pamulang Kota Tangerang Selatan , Banten ,Indonesia

Email : ^{1*}dosen02359@unpam.ac.id, ²murniptr8@gmail.com

(* : coressponding author)

Abstrak - Pemilu merupakan pilar utama demokrasi yang membutuhkan pengelolaan data secara efisien dan akurat untuk menjamin transparansi dan meningkatkan kepercayaan publik. Penelitian ini bertujuan untuk menganalisis data hasil Pemilu yang disediakan oleh Komisi Pemilihan Umum (KPU) melalui algoritma decision tree menggunakan aplikasi RapidMiner. Dataset yang digunakan diperoleh dari Kaggle dan mencakup berbagai informasi terkait hasil Pemilu. Penelitian ini difokuskan pada pengidentifikasian pola data, analisis prediktif, serta evaluasi performa model dengan pendekatan berbasis data. Evaluasi dilakukan menggunakan metrik akurasi, precision, dan recall untuk mengukur sejauh mana algoritma decision tree mampu menghasilkan prediksi yang relevan dan akurat. Hasil penelitian menunjukkan bahwa algoritma decision tree dapat mengidentifikasi pola data penting yang berkontribusi pada analisis prediktif hasil Pemilu. Selain itu, metrik evaluasi memberikan gambaran yang jelas tentang keandalan model dalam memprediksi data dengan tingkat presisi dan akurasi yang signifikan. Penelitian ini diharapkan dapat memberikan wawasan baru bagi pengelolaan data Pemilu di masa depan serta mendukung proses pengambilan keputusan berbasis data dalam sistem demokrasi.

Kata Kunci: Data Mining, Pemilu KPU, Decision Tree, RapidMiner, Analisis Data.

Abstract - Elections are the cornerstone of democracy, requiring efficient and accurate data management to ensure transparency and enhance public trust. This research aims to analyze election data provided by the General Election Commission (KPU) using the decision tree algorithm with RapidMiner. The dataset used is obtained from Kaggle and includes various information related to election results. This research focuses on identifying data patterns, predictive analysis, and evaluating model performance using a data-driven approach. Evaluation is conducted using accuracy, precision, and recall metrics to measure how well the decision tree algorithm can produce relevant and accurate predictions. The results show that the decision tree algorithm can identify important data patterns that contribute to predictive analysis of election results. Additionally, the evaluation metrics provide a clear picture of the model's reliability in predicting data with significant precision and accuracy. This research is expected to provide new insights for future election data management and support data-driven decisionmaking processes in the democratic system.

Keywords: Data Mining, KPU Election, Decision Tree, RapidMiner, Data Analysis .

1. PENDAHULUAN

Pemilu adalah momen penting dalam demokrasi yang mencerminkan partisipasi rakyat dalam menentukan arah <https://www.kaggle.com/datasets/akuncanva/data-pemilukpu> pembangunan bangsa. Dalam era digital, data Pemilu yang besar dan kompleks memerlukan pengelolaan yang efisien untuk mendukung analisis yang akurat. Data mining menjadi alat yang efektif untuk mengungkap pola, tren, dan hubungan tersembunyi dalam data besar. Algoritma Decision Tree adalah salah satu metode data mining yang populer karena kemampuannya dalam menghasilkan aturan keputusan yang mudah dipahami.

Penelitian ini bertujuan untuk menganalisis dataset Pemilu KPU yang diunduh dari Kaggle menggunakan algoritma Decision Tree melalui aplikasi RapidMiner. Fokus penelitian meliputi identifikasi pola data, analisis prediktif, dan evaluasi performa model berdasarkan metrik akurasi, precision, dan recall.



2. METODOLOGI

2.1 Dataset

Data set yang diperoleh dari Kaggle , data hasil pemilu dari Komisi Pemilihan Umum (KPU). Data mencakup informasi jumlah suara, partisipasi pemilih, dan hasil suara untuk setiap kandidat berdasarkan wilayah. Berikut link dataset yang diambil dari Kaggle : <https://www.kaggle.com/datasets/akuncanva/data-pemilu-kpu> , atau jika terjadi kendala dalam link tersebut maka gunakan link geogle drive berikut :

https://drive.google.com/drive/u/2/folders/1KUPLW4Hov8PwgOAJh4A_DBbQEJA3IEjz

2.2 Pra-pemrosesan Data

Pra-pemrosesan dilakukan untuk memastikan kualitas data sebelum dianalisis:

- Mengatasi Missing Values:** Mengganti nilai kosong dengan rata-rata atau modus, sesuai dengan jenis data.
- Normalisasi Data:** Untuk memastikan skala data seragam.
- Pembagian Data:** Dataset dibagi menjadi data pelatihan (80%) dan data pengujian (20%).

2.3 Penerapan Algoritma Decision Tree

Algoritma Decision Tree diterapkan di RapidMiner menggunakan operator **Decision Tree**. Proses melibatkan:

- Memuat dataset ke RapidMiner.
- Menentukan atribut target (misalnya, partai pemenang).
- Melatih model pada data pelatihan.
- Menguji model pada data pengujian.

2.4 Evaluasi Performa Model

Evaluasi dilakukan menggunakan metrik berikut:

- Akurasi:** Proporsi prediksi yang benar dibandingkan total prediksi.
- Precision:** Tingkat ketepatan prediksi kelas positif.
- Recall:** Proporsi data kelas positif yang berhasil diprediksi dengan benar.

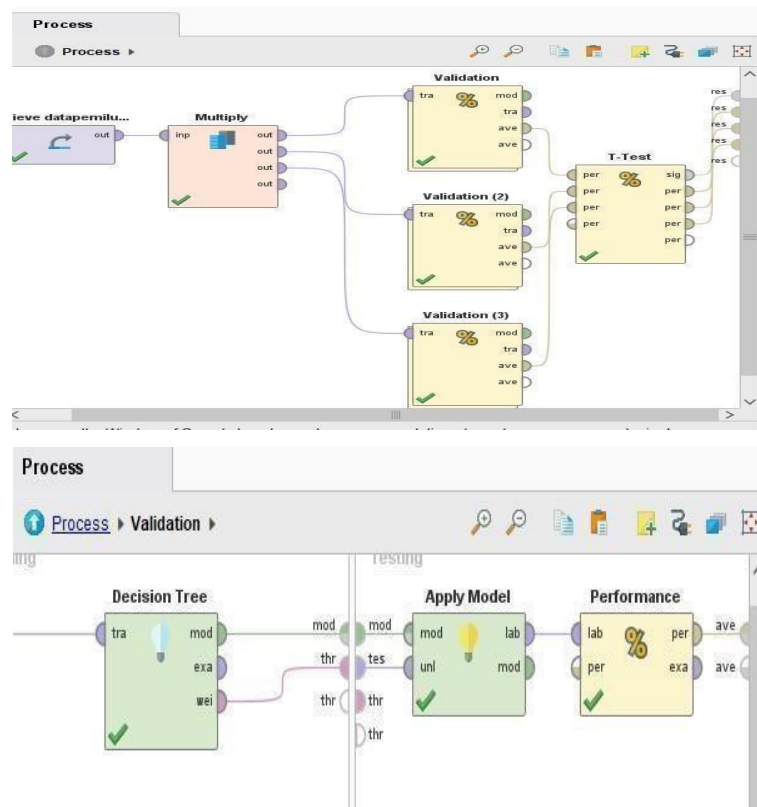
3. HASIL DAN PEMBAHASAN

3.1 Hasil Model Decision Tree

Hasil model menunjukkan struktur pohon keputusan dengan atribut seperti partisipasi pemilih dan distribusi suara sebagai faktor dominan. Decision Tree membantu memvisualisasikan aturan yang relevan dalam data Pemilu KPU.

3.2 Performa Algoritma

Hasil dari RapidMiner menunjukkan performa sebagai berikut:



- Akurasi

Table View Plot View

accuracy: 91.34%

	true TIDAK	true YA	class precision
pred. TIDAK	107	5	95.54%
pred. YA	6	9	60.00%
class recall	94.69%	64.29%	

- Precision

Table View Plot View

precision: 60.00% (positive class: YA)

	true TIDAK	true YA	class precision
pred. TIDAK	107	5	95.54%
pred. YA	6	9	60.00%
class recall	94.69%	64.29%	



JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 9 Februari Tahun 2025
ISSN 3025-0919 (media online)
Hal 1756-1760

- Recall

PerformanceVector (Performance (3))

Table View Plot View

recall: 64.29% (positive class: YA)

	true TIDAK	true YA	class precision
pred. TIDAK	107	5	95.54%
pred. YA	6	9	60.00%
class recall	94.69%	64.29%	

- PerformanceVektor

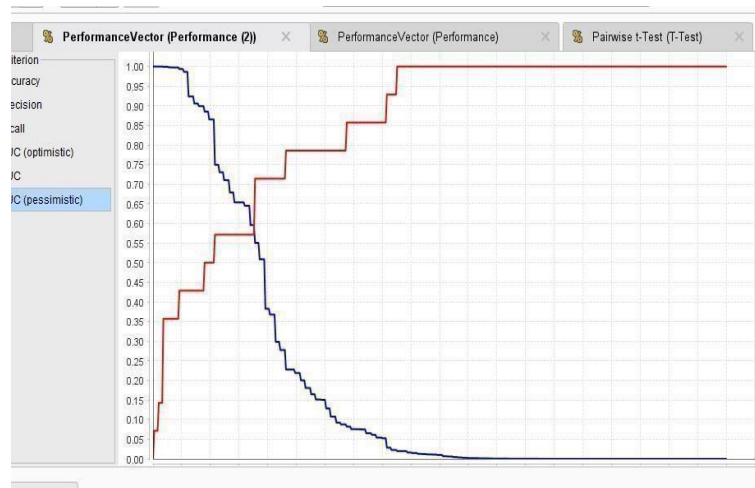
Result History

PerformanceVector

```

PerformanceVector:
accuracy: 91.34%
ConfusionMatrix:
True:  TIDAK  YA
TIDAK: 107    5
YA:    6      9
precision: 60.00% (positive class: YA)
ConfusionMatrix:
True:  TIDAK  YA
TIDAK: 107    5
YA:    6      9
recall: 64.29% (positive class: YA)
ConfusionMatrix:
True:  TIDAK  YA
TIDAK: 107    5
YA:    6      9
AUC (optimistic): 0.951 (positive class: YA)
AUC: 0.919 (positive class: YA)
AUC (pessimistic): 0.888 (positive class: YA)
  
```

- AUC (pessimistic)





JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 9 Februari Tahun 2025
ISSN 3025-0919 (media online)
Hal 1756-1760

3.3 Evaluasi Performa Model

Setelah diterapkan pada data pengujian, performa model dievaluasi dengan hasil berikut:

- **Akurasi:** 91,34 % • **Precision:** 60,00%
- **Recall:** 64,29 %

Hasil ini menunjukkan bahwa model memiliki kinerja yang baik dalam memprediksi partai pemenang berdasarkan data Pemilu KPU.

3.4 Diskusi

Decision Tree terbukti menjadi metode yang efektif untuk menganalisis data Pemilu KPU. Visualisasi pohon keputusan memberikan wawasan tentang atribut-atribut kunci yang memengaruhi hasil pemilu. Penggunaan RapidMiner mempermudah proses analisis, mulai dari pra-pemrosesan hingga evaluasi performa model.

4. KESIMPULAN

Penelitian ini berhasil mengelola dan menganalisis dataset Pemilu KPU menggunakan algoritma Decision Tree di RapidMiner. Hasil analisis menunjukkan performa model yang tinggi dalam mengidentifikasi pola data dan membuat prediksi. Penelitian ini dapat dijadikan referensi untuk pengambilan keputusan strategis terkait pemilu di masa depan.

REFERENCES

- (n.d.). (n.d.). *RapidMiner Documentation*. Retrieved from <https://docs.rapidminer.com>.
- Han, J. K. (2011). Concepts and Techniques. Morgan Kaufmann. *Data Mining*.
- KPU, D. P. (2024). *Kaggle*. (n.d.). Retrieved from <https://www.kaggle.com/datasets/akuncanva/data-pemilu-kpu>.