

PERBANDINGAN KRITERIA DECISION TREE PADA PENGETAHUAN MASYARAKAT PADA PEMILIHAN UMUM PRESIDEN INDONESIA

Siti Masripah¹, Lestari Yusuf^{2*}

Sistem Informasi Akuntansi Kampus Kota Bogor¹
Universitas Bina Sarana Informatika¹
www.bsi.ac.id¹
siti.stm@bsi.ac.id¹

Sistem Informasi²
Universitas Nusa Mandiri²
www.nusamandiri.ac.id²
lestari.lyf@nusamandiri.ac.id^{2*}



Abstract—In the Presidential election, gen z as a new voter, must know in advance who the presidential candidate will be in the 2024 election as well as the election process, because if voters do not know and do not understand, it will cause the wrong choice will result in their votes cannot be used even to abstain. What factors cause millennials and gen z generations to not know about elections can be determined using a decision tree. Therefore, in this study, a questionnaire was given to millennials and gen z generation to find out whether voters know the presidential candidate to be elected. The data from the questionnaire is processed to become training data and testing data with a ratio of 70:30. Then measure the accuracy level using the C4.5 algorithm with a comparison of splitting criteria, namely gain ratio, information gain and gini index. By knowing the right splitting criteria, the decision tree model can help overcome the problem of overfitting in the data. Overfitting occurs when the model is too complex in memorizing training data, thus failing to generalize well to read new data. The calculation results show the difference in accuracy values between Gain ratio, Information gain and Gini index, namely 81.67%, 83.33% and 83.33%. It can be concluded that for the use of Algorithm C4.5 splitting criteria Gain ratio and Gini index have the same accuracy value for accuracy measurement in this study.

Keywords: classification, decision tree, general election, president.

Abstrak—Pada pemilu Presiden, gen z sebagai pemilih baru, harus mengetahui terlebih dahulu siapa bakal calon presiden pada pemilu tahun 2024 begitupun dengan proses pemilunya, karena jika pemilih tidak tahu dan tidak paham maka akan menimbulkan salah pilih akan mengakibatkan suara mereka tidak bisa digunakan bahkan sampai golput. Faktor-faktor apa yang menyebabkan generasi milenial dan generasi gen z tidak mengetahui pemilu bisa ditentukan menggunakan decision tree. Oleh karena itu pada penelitian ini kuesioner diberikan kepada generasi millennial dan gen z untuk mengetahui apakah pemilih tahu akan calon presiden yang dipilih. Data hasil kuesioner tersebut diolah untuk menjadi data *training* dan data *testing* dengan perbandingan 70:30. Lalu ukur Tingkat akurasi menggunakan Algoritma C4.5 dengan perbandingan kriteria *splitting*, yaitu *gain ratio*, *information gain* dan *gini index*. Dengan mengetahui *splitting criteria* yang tepat, model *decision tree* dapat membantu mengatasi masalah *overfitting pada data*. *Overfitting* terjadi ketika model terlalu rumit dalam menghafal data pelatihan, sehingga gagal dalam menggeneralisasi dengan baik untuk membaca data baru. Hasil perhitungan menunjukkan perbedaan nilai akurasi antara *Gain ratio*, *Information gain* dan *Gini index* yaitu 81,67 %, 83,33% dan 83,33%. Dapat disimpulkan bahwa untuk penggunaan Algoritma C4.5 *splitting criteria Gain ratio* dan *Gini index* memiliki nilai akurasi yang sama untuk pengukuran akurasi pada penelitian ini.

Kata Kunci: klasifikasi, pohon keputusan, pemilihan umum, presiden.

PENDAHULUAN

Pemilihan presiden di Indonesia dilakukan setiap lima tahun sekali, seperti yang dinyatakan dalam UUD 1945 pasal 22E ayat 1-6. Regulasi tersebut menjadikan pilpres sebagai kegiatan yang signifikan untuk menentukan kepemimpinan lima tahun ke depan (Maharani & Santosa, 2022). Sebagai pemilih pemula, keberadaan Gen-Z juga sangat diperhitungkan dalam Pilpres 2024 (Febriana, 2023). Namun dalam prosesnya, Komisi Pemilihan Umum (KPU) menemukan bahwa pemilih belum memahami pemilihan presiden kali ini, seperti penjelasannya dalam (Fitratun Komariah, 2023).

Pengetahuan masyarakat pada pemilu Presiden salah satunya adalah mengetahui siapa bakal calon presiden pada pemilu tahun 2024. Ketidaktahuan tentang pemilihan presiden akan mengakibatkan pemilih menjadi golput atau memilih pasangan calon presiden secara asal-asalan. Dampak dari golput akan mengurangi legitimasi pemilu seperti yang dijelaskan dalam (Susthira, 2019). Minimnya pemahaman tentang pemilu membuat KPU harus lebih optimal dalam melakukan sosialisasi bakal calon presiden kepada pemilih. Berdasarkan permasalahan tersebut dilakukan klasifikasi data dengan melakukan perbandingan kriteria *splitting* pada Algoritma C4.5

Klasifikasi dapat menggunakan beberapa metode yang terdapat dalam supervised learning. Algoritma ini memerlukan masukan berupa contoh pelatihan, label pelatihan, dan atribut. C4.5, yang merupakan peningkatan dari ID3, memiliki kemampuan tambahan untuk menangani masalah seperti nilai yang hilang, data yang bersifat continue, dan proses pemangkasan (Putra, 2023). Algoritma C4.5 juga dianggap lebih baik dalam menentukan gizi balita dalam artikel (Prasetio, 2021).

Selain itu, algoritma C4.5 juga digunakan oleh (Istiawan & Khikmah, 2019) dalam penelitiannya saat melakukan prediksi lahan kritis pada kawasan budidaya pertanian di daerah aliran sungai yang memiliki nilai akurasi sebesar 92,47%. Selain itu, algoritma C4.5 juga dapat menyelesaikan permasalahan pengukuran kepuasan. Dalam penelitian yang ditulis oleh (Widiastuti et al., 2022) dijelaskan bahwa algoritma dari level kepuasan pelayanan akademik di Universitas Jenderal Achmad Yani sebesar 90%. Angka tersebut menandakan bahwa algoritma decision tree baik untuk digunakan untuk menyelesaikan masalah klasifikasi.

Pada penelitian ini, klasifikasi algoritma C4.5 pada pre-processingnya akan menggunakan *splitting* criteria atau seleksi fitur. Seleksi fitur yang digunakan pada penelitian ini adalah *Information gain*, *Gain ratio* dan *Gini index*. Menurut Tangirala

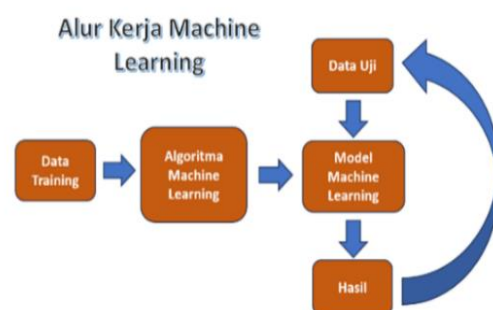
pada penelitiannya (Tangirala, 2020) pada saat membangun decision tree tujuan dari setiap simpul adalah untuk memilih *splitting* criteria yang optimal agar dapat membagi dataset menjadi subset yang seragam. Proses ini bertujuan untuk menghasilkan nilai bobot atau tingkat kepentingan untuk setiap atribut dalam dataset (M. Adib Al Karomi, Abdul Kharis, 2019). Begitupun pada artikel yang ditulis oleh Fajrian (Fajriyan et al., 2022) kriteria *gain ratio* mendapatkan hasil yang lebih baik dibandingkan *information gain* dengan tingkat akurasi perbandingan sebesar 70,06%.

Berdasarkan latar belakang tersebut, pada penelitian ini dilakukan penambahan kriteria *splitting* yang digunakan pada algoritma C4.5 dan untuk mengetahui nilai akurasi yang tertinggi pada pengetahuan Masyarakat akan pemilu presiden dengan mengetahui siapa bakal calon presiden dengan tiga kriteria *splitting* yaitu *Information gain*, *Gain ratio* dan *Gini index*. Penelitian mengenai perbandingan *splitting* criteria inipun bertujuan agar pembaca dapat mengetahui nilai akurasi jika menggunakan *splitting* criteria akan lebih baik atau tidak.

BAHAN DAN METODE

Penelitian ini menggunakan *Decision tree*, bagian dari Algoritma *Machine Learning*. Alur kerja *Machine Learning* yang pertama adalah memasukkan data *Training* ke dalam algoritma *machine learning* untuk dilatih. Hasil pengolahan data akan menghasilkan pengetahuan baru (Model *Machine Learning*).

Setelah mendapatkan hasil berupa pengetahuan, beberapa data lain disiapkan untuk pengujian. Tujuannya adalah untuk melihat apakah model *machine learning* sudah cukup atau harus diperbaiki. Proses tersebut ditampilkan pada Gambar 1.



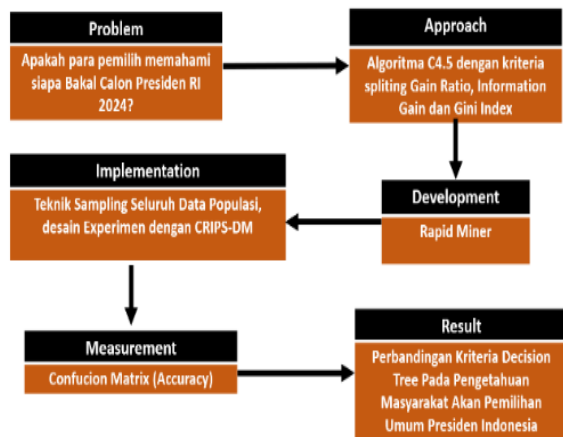
Sumber: (Masripah & Yusuf, 2024)

Gambar 1. Alur Kerja Machine Learning

Berdasarkan alur kerja *Machine Learning*, penelitian ini menggunakan metode teknik Data

Mining menggunakan algoritma C4.5 dengan beberapa kriteria pemisahan *Information gain*, *Gain ratio* dan *Gini index*. Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang dikembangkan oleh J. Ross Quinlan, salah satu algoritma induksi pohon keputusan (Nurwulandari et al., 2022). Data mining adalah sebuah proses, maka prosesnya pun harus sesuai dengan prosedur, yaitu proses Crisp-DM (*Cross-Standard Industry Process for Data Mining*) sebagai sebuah proses yang utuh (Masripah et al., 2022), data preprocessing, model building, model *evaluation*, dan yang terakhir adalah model deployment.

Penjelasan mengenai kerangka kerja penelitian dapat dilihat pada Gambar 2.



Sumber: (Masripah & Yusuf, 2024)

Gambar 2. Kerangka pemikiran Kerangka kerja penelitian

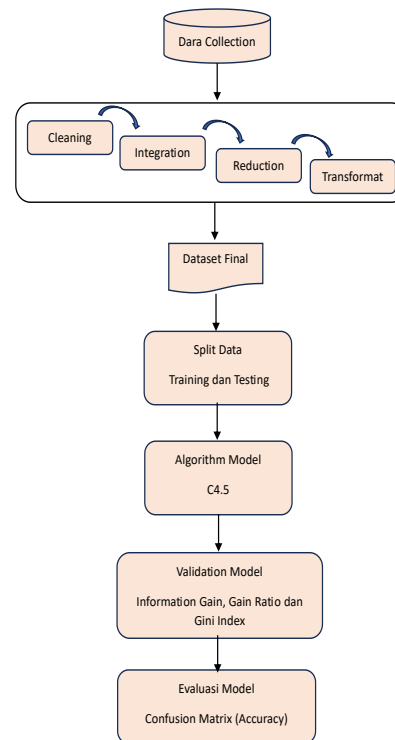
Pada Gambar 2 dijelaskan bahwa permasalahan dari penelitian ini adalah Apakah para pemilih memahami siapa Bakal Calon Presiden RI 2024?. Metode yang digunakan dalam penelitian ini adalah menggunakan Algoritma C4.5 dengan menggunakan *splitting* criteria. *Splitting* criteria digunakan untuk mengukur seberapa baik sebuah atribut. *Splitting* criteria yang akan digunakan adalah:

1. *Information gain*. Merupakan metode pemisahan kriteria paling sederhana. *Information gain* melakukan perangkingan dalam aplikasi yang berkategori teks (Kurniabudi et al., 2021).
2. *Gain ratio*, merupakan pengembangan dari *information gain*. *Gain ratio* baik digunakan untuk data yang memiliki dua class dengan data numerik. *Gain ratio* juga dapat mengurangi bias pada data. Perbaikan data juga merupakan fitur pada *gain ratio* dengan cara memperbaiki informasi intristik dari atribut yang ada (Fajriyan et al., 2022).

3. *Gini index*, merupakan *splitting* criteria yang fokus pada *biner split* pada setiap atribut data. Setelah mendapatkan *biner split Gini index* akan melakukan penjumlahan *impurity* setiap atributnya (Kurniawan, 2020).

Perhitungan penelitian ini menggunakan tools RapidMiner, dengan teknik CRISP-DM, untuk pengukuran akurasi dilakukan dengan Confusion Matrix (Akurasi) sehingga hasil penelitian yang didapatkan adalah Algoritma C4.5 dalam pengetahuan pemilih pada Pemilihan Calon Presiden RI 2024.

Adapun tahapan penelitian yang dilakukan ditampilkan pada gambar 3.



Sumber: (Masripah & Yusuf, 2024)

Gambar 3. Tahap Penelitian

Adapun penjelasan dari tahap penelitian pada gambar 3 sebagai berikut:

1. Tahap pencarian data dilakukan dengan menyebarkan kuesioner dari tanggal 11 september – 6 Oktober 2023 yang korespondennya adalah masyarakat yang memiliki hak pilih pada pemilu 2024.
2. Tahap preprocessing data dilakukan untuk menghindari data yang bermasalah, data yang tidak diperlukan, sehingga data dapat digunakan dengan baik atau menjadi data final.
3. Data yang ada dipisahkan untuk mendapatkan data training dan data testing dengan perbandingan data 70 banding 30.

4. Model yang digunakan adalah Algoritma C4.5.
5. Model validasi yang digunakan adalah *Information gain*, *Gain ratio* dan *Gini index*.
6. Model evaluasi yang digunakan adalah Confusion Matrix (Akurasi).

HASIL DAN PEMBAHASAN

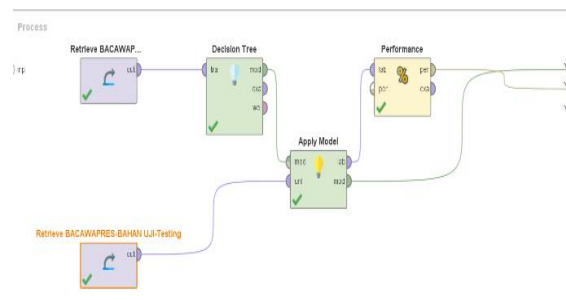
Hasil penelitian menggunakan Algoritma C4.5 dengan perbandingan kriteria spilling yaitu *Information gain*, *Gain ratio* dan *Gini index*. Atribut pada penelitian ini ditampilkan pada tabel 1.

Tabel 1. Atribut

Atribut	Keterangan
TL	Tahun Lahir
Usia	Usia
Gen	Generasi
JK	Jenis Kelamin
Status	Status
Pendidikan	Pendidikan
Lokasi	Lokasi Tinggal
Pekerjaan	Pekerjaan
Aktif M	Aktif di masyarakat
Aktif S	Aktif di medsos
Medsos	Media sosial yang dimiliki
P1	Apakah tahu pemilu diadakan tahun 2024
P2	Apakah tahu dan paham siapa bakal calon presiden

Sumber: (Masripah & Yusuf, 2024)

Dengan jumlah dataset sebanyak 196 data, dibagi menjadi data training dan data testing, untuk 136 data training dan untuk 60 data testing, dengan label P2. Pada gambar 4 terlihat model dari *RapidMiner* yaitu setelah data dibaca oleh *rapidminer*, lalu pilih metode *decision tree*, pada tahap inilah kriteria pemisah dipilih, perbandingan antara *Information gain*, *Gain ratio* dan *Gini index*. Kemudian dilakukan *testing* data dengan *apply model* dan Evaluasi *performance* untuk mendapatkan nilai akurasi.



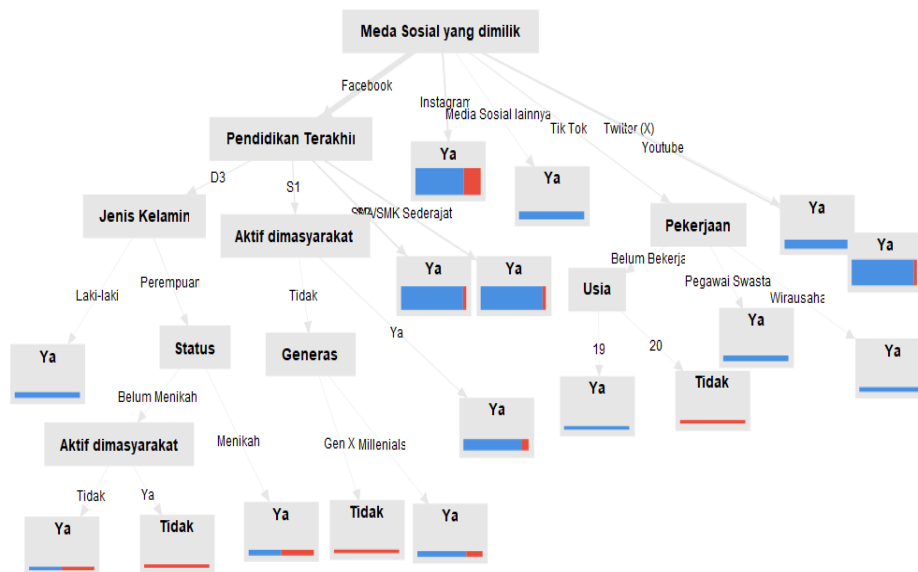
Sumber: (Masripah & Yusuf, 2024)

Gambar 4. Pemodelan dengan *Rapid Miner*

Berikut hasil pengolahan data Algoritma C4.5 dengan kriteria *splitting*.

1. *Information gain*

Pada kriteria *splitting information gain*, dapat terlihat pohon keputusan yang dihasilkan, Dimana yang menjadi akar utama adalah media sosial yang dimiliki, akar pohon dapat dilihat pada gambar 5.



Sumber: (Masripah & Yusuf, 2024)

Gambar 5. Akar Pohon *Information gain*

Berikut penjelasan akar pohon dengan kriteria *Information gain* :

IF Media Sosial yang dimiliki = Facebook AND Pendidikan Terakhir = D3 AND Jenis Kelamin = Perempuan AND status = belum menikah AND aktif di masyarakat = tidak THEN P2 : Ya

IF Media Sosial yang dimiliki = Facebook AND Pendidikan Terakhir = D3 AND Jenis Kelamin = Perempuan AND status = belum menikah AND aktif di masyarakat = Ya THEN P2 : Tidak

IF Media Sosial yang dimiliki = Facebook AND Pendidikan Terakhir = S1 AND aktif di masyarakat = tidak AND generasi = Gen X THEN P2 : Tidak

IF Media Sosial yang dimiliki = Facebook AND Pendidikan Terakhir = S1 AND aktif di masyarakat = tidak AND generasi = milenials THEN P2 : Ya

IF Media Sosial yang dimiliki = Facebook AND Pendidikan Terakhir = S1 AND aktif di masyarakat = Ya THEN P2 : Ya

IF Media Sosial yang dimiliki = Facebook AND Pendidikan Terakhir = S2 THEN P2 : Ya

IF Media Sosial yang dimiliki = Facebook AND Pendidikan Terakhir = SMA/SMK Sederajat THEN P2 : Ya

IF Media Sosial yang dimiliki = Tik tok AND Pekerjaan = belum bekerja AND Usia = 19 THEN P2 : Ya

IF Media Sosial yang dimiliki = Tik tok AND Pekerjaan = belum bekerja AND Usia = 20 THEN P2 : Tidak

IF Media Sosial yang dimiliki = Tik tok AND Pekerjaan = pegawai swasta THEN P2 : Ya

IF Media Sosial yang dimiliki = Tik tok AND Pekerjaan = wirausaha THEN P2 : Ya

IF Media Sosial yang dimiliki = Youtube THEN P2 : Ya

IF Media Sosial yang dimiliki = Twiter THEN P2 : Ya

Kriteria *splitting Information gain* menghasilkan nilai akurasi sebesar 81,67%, menghasilkan *Confusion matrix Information gain* seperti pada tabel 2.

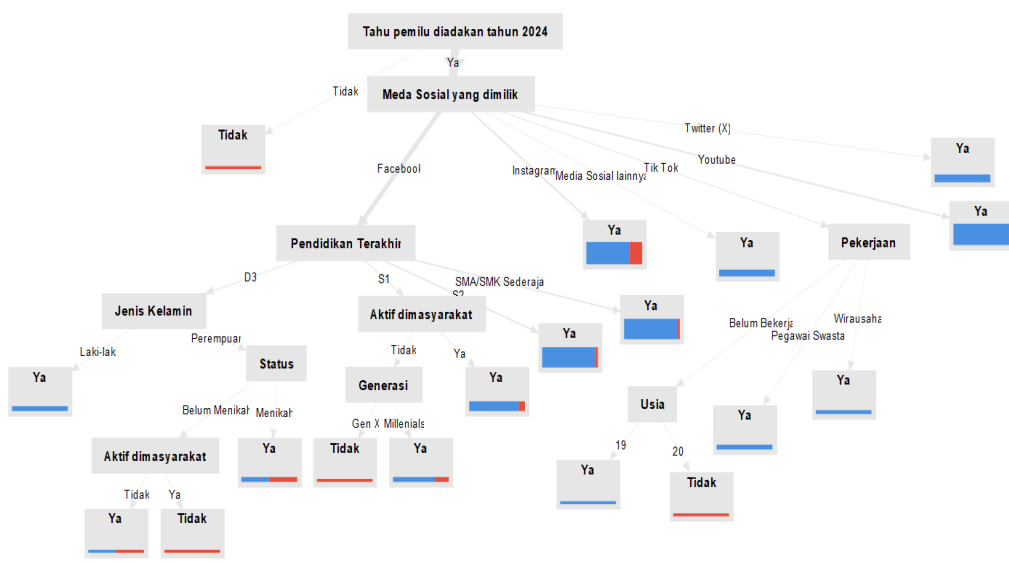
Tabel 2. Evaluasi *Confusion matrix Information gain*

	true Tidak	true Ya
pred. Tidak	0	2
pred. Ya	9	49
class recall	0.00%	96.08%

Sumber: (Masripah & Yusuf, 2024)

2. Gain ratio

Pada kriteria *Gain ratio*, akar yang terpilih paling atas pada akar pohon adalah Tahu pemilu diadakan tahun 2024 (P1) terlihat pada gambar 6.



Sumber: (Masripah & Yusuf, 2024)

Gambar 6. Pohon Keputusan dengan Kriteria *Information gain*

Berikut penjelasan akar pohon dengan kriteria *Gain ratio* :

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = D3 AND Jenis kelamin = Laki-laki THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = D3 AND Jenis kelamin = Perempuan AND status = Belum menikah AND Aktif di masyarakat = Tidak THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = D3 AND Jenis kelamin = Perempuan AND status = Belum menikah AND Aktif di masyarakat = Ya THEN P2 = Tidak

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = D3 AND Jenis kelamin = Perempuan AND status = menikah THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = S1 AND aktif di masyarakat = Tidak AND Generasi = Gen X THEN P2 = Tidak

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = S1 AND aktif di masyarakat = Tidak AND Generasi = Milenials THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = S1 AND aktif di masyarakat = Ya THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = S2 THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = SMA/SMK sederajat THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Instagram THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Media sosial lainnya THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Tik Tok AND Pekerjaan = belum bekerja AND usia = 18 THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Tik Tok AND Pekerjaan = belum bekerja AND usia = 20 THEN P2 = Tidak

IF P1 = Ya AND Media sosial yang dimiliki = Tik Tok AND Pekerjaan = pegawai swasta THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Tik Tok AND Pekerjaan = wirausaha THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Youtube THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Twitter THEN P2 = Ya

IF P1 = Tidak THEN P2 = Tidak

Kriteria *splitting Gain ratio* menghasilkan nilai akurasi sebesar 83.33%, menghasilkan Confusion matrix *gain ratio* terlihat pada tabel 3.

Tabel 3. Evaluasi *Confusion matrix Gain ratio*

	True Tidak	True Ya
Pred. Tidak	1	2
Pred. Ya	8	49
Class recall	11.11%	96.08%

Sumber: (Masripah & Yusuf, 2024)

3. Gini index

Pada kriteria *Gini index*, akar yang terpilih paling atas pada akar pohon adalah Tahu pemilu diadakan tahun 2024 (P1) terlihat pada gambar 7. Adapun penjelasan akar pohon dengan kriteria *Gini index* sebagai berikut.

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = D3 AND Jenis kelamin = Laki-laki THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = D3 AND Jenis kelamin = Perempuan AND status = Belum menikah AND Aktif di masyarakat = Tidak THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = D3 AND Jenis kelamin = Perempuan AND status = Belum menikah AND Aktif di masyarakat = Ya THEN P2 = Tidak

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = D3 AND Jenis kelamin = Perempuan AND status = menikah THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = S1 AND aktif di masyarakat = Tidak AND Generasi = Gen X THEN P2 = Tidak

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = S1 AND aktif di masyarakat = Tidak AND Generasi = Milenials THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = S1 AND aktif di masyarakat = Ya THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = S2 THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Facebook AND Pendidikan terakhir = SMA/SMK sederajat THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Instagram THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Media sosial lainnya THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Tik Tok AND Pekerjaan = belum bekerja AND usia = 18 THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Tik Tok AND Pekerjaan = belum bekerja AND usia = 20 THEN P2 = Tidak

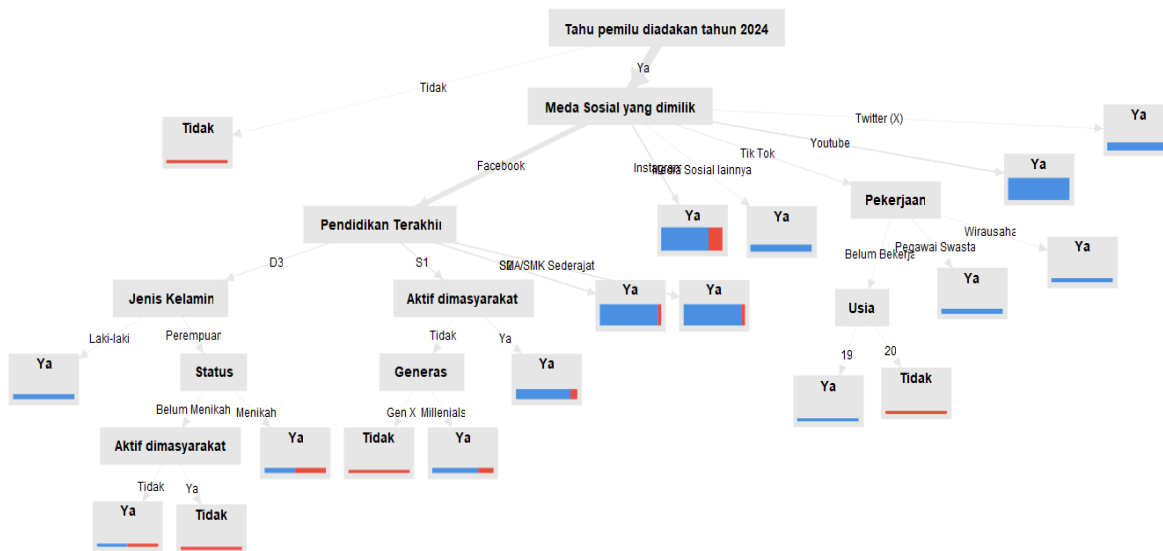
IF P1 = Ya AND Media sosial yang dimiliki = Tik Tok AND Pekerjaan = pegawai swasta THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Tik Tok AND Pekerjaan = wirausaha THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Youtube THEN P2 = Ya

IF P1 = Ya AND Media sosial yang dimiliki = Twitter THEN P2 = Ya

IF P1 = Tidak THEN P2 = Tidak



Sumber: (Masripah & Yusuf, 2024)

Gambar 7. Pohon Keputusan dengan Kriteria *Gini index*

Kriteria *splitting Gini index* menghasilkan nilai akurasi sebesar 83.33%, menghasilkan *Confusion matrix gini index* terlihat pada tabel 4.

Tabel 4. Evaluasi *Confusion matrix Gini index*

	True Tidak	True Ya
Pred. Tidak	1	2
Pred. Ya	8	49
Class Recall	11.11%	96.08%

Sumber: (Masripah & Yusuf, 2024)

4. Hasil Perbandingan

Hasil perbandingan kriteria *splitting* untuk nilai akurasi pada ketiga kriteria, dapat dilihat pada tabel 5.

Tabel 5. Perbandingan nilai akurasi kriteria *splitting*

	Information gain	Gain ratio	Gini index
kriteria	81,67%	83,33%	83,33%

Sumber: (Masripah & Yusuf, 2024)

Pada tabel 5, menunjukkan bahwa hasil nilai akurasi pada perbandingan tiga kriteria *splitting* dengan nilai tinggi adalah kriteria *splitting Gain ratio* dan *Gini index* dengan nilai 83,33%.

Nilai akurasi dari *gain ratio* dan *gini index* menunjukkan bahwa dua kriteria tersebut lebih bisa memperbaiki data-data ada sehingga menghasilkan tingkat akurasi yang lebih baik dalam penggunaannya. Maka penggunaan *decision tree* C4.5 akan lebih maksimal tingkat akurasinya jika menggunakan *splitting criteria gain ratio* dan *gini index*.

KESIMPULAN

Dari hasil penelitian didapat akurasi yang dihasilkan dari Algoritma C4.5 dengan perbandingan tiga kriteria *splitting* yaitu *information gain* sebesar 81,67% sedangkan untuk *Gain ratio* dan *Gini index* menghasilkan nilai yang sama yaitu 83,33%. Dengan mengetahui hasil perbandingan akurasi tiga kriteria tersebut, pengguna dapat mengetahui *splitting* kriteria mana yang lebih baik digunakan untuk meningkatkan nilai akurasi. Mengetahui kriteria yang tepat pun akan membuat model *decision tree* dapat membantu mengatasi masalah overfitting. Berdasarkan hasil tersebut dapat disimpulkan bahwa untuk penggunaan Algoritma C4.5 kriteria *splitting Gain ratio* dan *Gini index* memiliki nilai akurasi yang lebih baik.

Berdasarkan hasil pengujian dan kesimpulan yang didapat terdapat beberapa saran pada penelitian ini diperlukan untuk menerapkan teknik optimasi untuk kriteria *splitting* agar hasil akurasi lebih optimal. Penelitian berikutnya pun dapat melakukan perbandingan dengan beberapa Algoritma lainnya, agar nilai akurasi menjadi lebih baik.

REFERENSI

- Fajriyan, F. N., Ahsan, M., & Harianto, W. (2022). Komparasi Tingkat Akurasi *Information gain* Dan *Gain ratio* Pada Metode K-Nearest Neighbor. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(1), 386–391.
- Febriana, S. (2023, July 1). Indikator Politik: Suara Milenial dan Gen Z di Pemilu 2024 Mencapai 53%. *Metrotvnews.Com*. <https://www.metrotvnews.com/play/b3JCyDxg-indikator-politik-suara-milenial-dan-gen-z-di-pemilu-2024-mencapai-53>
- Fitratun Komariah. (2023). KPU Jakarta Temukan Banyak Pemilih Kurang Paham Pemilu. *Rri.Co.Id*. <https://www.rri.co.id/pusat-pemberitaan/pemilu/401599/kpu-jakarta-temukan-banyak-pemilih-kurang-paham-pemilu>
- Istiawan, D., & Khikmah, L. (2019). Implementation of C4.5 Algorithm for Critical Land Prediction in Agricultural Cultivation Areas in Pemali Jratun Watershed. *Indonesian Journal of Artificial Intelligence and Data Mining*, 2(2), 67. <https://doi.org/10.24014/ijaidm.v2i2.7569>
- Kurniabudi, K., Harris, A., & Mintaria, A. E. (2021). Komparasi *Information gain*, *Gain ratio*, CFS-Bestfirst dan CFS-PSO Search Terhadap Performa Deteksi Anomali. *Jurnal Media Informatika Budidarma*, 5(1), 332.
- Kurniawan, H. (2020). Deteksi Twitter Bot menggunakan Klasifikasi Decision Tree. *Jurnal Sustainable: Jurnal Hasil Penelitian Dan Industri Terapan*, 9(1), 31–37.
- M. Adib Al Karomi, Abdul Kharis, I. (2019). Optimasi Algoritma Naive Bayes Dengan *Information gain ratio* Untuk Menangani Dataset Berdimensi Tinggi. *Seminar Nasional Edusaintek*, 37–43.
- Maharani, T., & Santosa, B. (2022). Pakar Hukum: UUD Sudah Kunci Pemilu Dilaksanakan 5 Tahun Sekali, Tak Etis Ada Amendemen. *Kompas.Com*. <https://nasional.kompas.com/read/2022/03/09/12031701/pakar-hukum-uud-sudah-kunci-pemilu-dilaksanakan-5-tahun-sekali-tak-etis-ada>
- Masripah, S., Nurwulandari, D. A., & Saputra, R. A. (2022). Pencarian Criteria *Splitting* Terbaik Pada Algoritma C4. 5 Untuk Mengukur Pemilihan Pembelajaran Pada Era Pandemi Covid-19. *Jurnal Larik: Ladang Artikel Ilmu Komputer*, 2(1), 1-7.
- Nurwulandari, D. A., Masripah, S., & Saputra, R. A. (2022). Optimasi Algoritma C4.5 Untuk Mengukur Keputusan Pembelajaran Daring Berbasis Particle Swarm Optimization (PSO). *IJCIT (Indonesian Journal on Computer and Information Technology)*, 7(2), 103–110.
- Prasetyo, A. (2021). Simulasi Penerapan Metode Decision Tree (C4.5) Pada Penentuan Status Gizi Balita. *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, 4(3), 209–214.
- Putra, A. R. (2023). The Implementation of Data Mining Techniques for Predicting Student Study Period Using the C4. 5 Algorithm: Penerapan Teknik Data Mining Terhadap Prediksi *Indonesian Journal of Electrical Engineering and ...*, 3(December), 96–100.
- Susthira, K. M. (2019). *Golput Merugikan Sekaligus Mendeligitimasi Pemilu*. <https://mediaindonesia.com/politik-dan->

hukum/226065/golput-merugikan-
sekaligus-mendeligitimasi-pemilu

Tangirala, S. (2020). Evaluating the impact of *Gini index* and *information gain* on classification using decision tree classifier algorithm. *International Journal of Advanced Computer Science and Applications*, 11(2), 612–619.

Widiastuti, T., Karsa, K., & Juliane, C. (2022). Evaluasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Metode Klasifikasi Algoritma C4.5. *Technomedia Journal*, 7(3), 364–380.