

ANALISIS FAKTOR PENENTU DALAM ULASAN DAN REKOMENDASI TERHADAP KEPUTUSAN KONSUMEN DI MEDIA SOSIAL PADA MARKETPLACE SHOPEE MENGGUNAKAN ALGORITMA C4.5

Mutiara Novita Angie, Dita Azizah, Annisa Hertyanti, Mutya Endah Sukmawati

Sistem Informasi, Universitas Bina Sarana Informatika

Jalan Kramat Raya 98 Jakarta, Indonesia

19221141@bsi.ac.id

ABSTRAK

Sebelum melakukan pembelian, konsumen sering kali mencari ulasan dan rekomendasi dari pengguna lain untuk memastikan kualitas produk yang akan dibeli. Ulasan yang menggambarkan pengalaman pengguna dan rekomendasi dari sumber terpercaya berperan penting dalam mempengaruhi keputusan pembelian konsumen. Melalui platform media sosial, konsumen dapat dengan mudah mengakses berbagai macam ulasan, rekomendasi, dan informasi terkait produk yang mereka minati. Penelitian ini bertujuan untuk mencari tahu faktor yang paling unggul dari keputusan konsumen sebelum melakukan pembelian suatu produk. Faktor-faktor yang dianalisis meliputi pengalaman pribadi, kepercayaan terhadap sumber informasi, kualitas informasi yang diberikan, kuantitas ulasan, jenis platform yang digunakan, dan pengaruh sosial. Metode yang digunakan dalam penelitian ini yaitu data mining Algoritma C4.5 menggunakan tool *RapidMiner* dengan memberikan kuesioner kepada 100 responden sebagai perwakilan dari pengguna Media Sosial. Kriteria responden meliputi pengguna Media Sosial yang aktif melakukan pembelian online di *marketplace* Shopee serta pengguna yang pernah membaca ulasan atau mendapatkan rekomendasi sebelum melakukan pembelian. Hasil dari penelitian ini menunjukkan bahwa faktor paling unggul yang mempengaruhi keputusan pembelian konsumen adalah faktor Kuantitas Ulasan dengan Tingkat akurasi sempurna yaitu dengan nilai 100 %.

Kata kunci : Algoritma C4.5, Ulasan, Rekomendasi, Media Sosial, Keputusan Pembelian

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi yang pesat telah membawa perubahan dalam perilaku konsumen, terutama dalam konteks pembelian produk secara online. Konsumen tidak lagi sekadar mengandalkan informasi yang disediakan oleh penjual, tetapi juga aktif mencari informasi melalui ulasan dan rekomendasi dari pengguna lain di media sosial. Ulasan dan rekomendasi ini berfungsi sebagai sumber informasi yang sangat berharga, memberikan konsumen gambaran yang lebih jelas tentang produk sebelum mereka membuat keputusan pembelian.

Ulasan konsumen atau Online customer review adalah ulasan yang diberikan oleh konsumen terkait dengan informasi dari evaluasi suatu produk tentang berbagai macam aspek, dengan adanya informasi ini konsumen bisa mendapatkan kualitas dari produk yang dicari dari ulasan dan pengalaman yang ditulis oleh konsumen yang telah membeli produk dari penjual online [1].

Penelitian terdahulu yang dilakukan oleh Rahmawati dkk ditemukan bahwa ulasan online, baik dari segi kualitas maupun kuantitas, sangat berpengaruh terhadap kepercayaan konsumen dalam membeli produk secara online. Pengalaman pribadi yang dibagikan melalui ulasan menciptakan rasa kepercayaan yang lebih besar, sehingga konsumen merasa lebih yakin dalam membuat keputusan pembelian [2].

Selain itu, Arif dkk menekankan pentingnya pengaruh sosial dalam keputusan pembelian. Rekomendasi dari teman, keluarga, atau pengguna

yang dianggap memiliki kredibilitas tinggi dapat meningkatkan keyakinan konsumen terhadap suatu produk. Rekomendasi tersebut seringkali bersifat persuasif, terutama jika didukung oleh ulasan yang positif dan rinci. Selain itu, aspek teknis seperti kepercayaan terhadap platform marketplace, kualitas informasi dalam ulasan, dan jumlah rekomendasi juga mempengaruhi keputusan pembelian[3].

Tujuan dari penelitian ini adalah untuk menganalisis faktor yang paling unggul dari keputusan konsumen sebelum melakukan pembelian suatu produk. Faktor-faktor yang dianalisis meliputi pengalaman pribadi, kepercayaan terhadap sumber informasi, kualitas informasi yang diberikan, kuantitas ulasan, jenis platform yang digunakan, dan pengaruh sosial. Metode yang digunakan dalam penelitian ini meliputi analisis data mining menggunakan Algoritma C4.5 pada tool *RapidMiner*.

2. TINJAUAN PUSTAKA

Beberapa teori yang digunakan dalam penelitian ini demi mendukung baik perhitungan dan pembahasan terkait dengan Algoritma C4,5.

2.1. E - Commerce

E-commerce adalah suatu kegiatan membeli atau menjual secara elektronik yang dilakukan pada jaringan internet. E-commerce terlihat lebih nyata, dengan adanya kebutuhan penjual dan pembeli untuk melakukan transaksi [4].

2.2. Keputusan Pembelian

Keputusan pembelian adalah tahap dalam proses pengambilan keputusan pembelian dimana konsumen benar benar membeli. Pengambilan keputusan merupakan suatu kegiatan individu secara langsung terlibat dalam mendapatkan dan mempergunakan barang yang ditawarkan [5].

2.3. Data Mining

Data mining adalah suatu proses ekstraksi atau penggalian data dan informasi yang besar, yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting [6].

2.4. Macam-Macam Data Mining [7]

- a. Klasifikasi
Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, sedang dan rendah.
- b. Estimasi
Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.
- c. Prediksi
Prediksi hampir sama dengan klasifikasi dan estimasi. Hanya saja dalam prediksi, nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat digunakan untuk prediksi.
- d. Pengklasteran
Pengklasteran merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Pengklasteran berbeda dengan klasifikasi karena tidak adanya variabel target. Pengklasteran tidak mengklasifikasi, mengestimasi, atau memprediksi nilai dari variabel target namun pengklasteran melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan.
- e. Asosiasi
Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja. Atribut yang dimaksudkan dalam hal ini adalah faktor-faktor yang memengaruhi pengambilan keputusan.

2.5. Tahapan Data Mining [8]

Data Mining merupakan bagian yang terintegrasi dari Knowledge Discovery in Databases (KDD). Bertujuan untuk proses transformasi data mentah menjadi informasi berguna. Jika digambarkan secara detail tahapan KDD menjadi 5 tahap. Berikut Tahapannya:

- a. Seleksi
- b. Preprocessing
- c. Transformasi
- d. Data Mining
- e. Interpretasi & Evaluasi

2.6. Algoritma C4.5

Algoritma C4.5 merupakan metode yang dapat digunakan untuk melakukan pembentukan pohon keputusan. Decision tree tersebut mampu menghasilkan keputusan yang kompleks menjadi lebih sederhana, sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan[9] Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang di ciptakan oleh J. Rose Quinlan. Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

- a. Pilih atribut sebagai akar
- b. Buat cabang untuk tiap-tiap nilai
- c. Bagi kasus dalam cabang
- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama [10].

2.7. Decision Tree

Decision tree merupakan salah satu metode klasifikasi yang menggunakan representasi struktur pohon (tree) dimana setiap node merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut, dan daun merepresentasikan kelas[11].

2.8. RapidMiner

Menurut Şeker dan Şadi, Rapid miner adalah perangkat lunak ilmu data yang dikembangkan oleh perusahaan bernama sama dengan yang menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam, penambahan teks, dan analisis prediktif[12].

2.9. Data Training dan Testing

Data Training adalah data training adalah data yang digunakan sebagai acuan untuk membangun model klasifikasi. Data Testing adalah data yang digunakan untuk menguji performa dari model klasifikasi tersebut[13].

2.10. Confusion Matrix

Tahap ini dilakukan uji ketepatan dengan confusion matrix kedua algoritma, yaitu menghitung accuracy, Class recall, precision dengan Confusion Matrix yang dilakukan pada tools RapidMiner[14]

Tabel 1. Confusion Matrix

Kelas	Terklarifikasi Positif	Terklarifikasi Negatif
Positif	TP (True Positif)	FN (False Negative)
Negatif	FP (False Positif)	TN (True Negative)

Pada kinerja Confusion Matrix dapat diukur dengan TP, FN, FP dan TN.

TP (True Positive) adalah jumlah data true yang memang memiliki nilai kebenaran true.

FN (False Negative) adalah jumlah data false yang memang memiliki nilai kebenaran false.

FP (False Negative) adalah jumlah data true yang dianggap sistem memiliki nilai kebenaran false.

TN (True Negative) adalah jumlah data false yang dianggap sistem memiliki nilai kebenaran true.

Menghitung *Accuracy*:

$$\frac{TP+TN}{TP+FN+FP+TN} \dots\dots\dots(1)$$

Menghitung *Class Recall*:

$$\frac{TP}{TP+FN} \dots\dots\dots(2)$$

Menghitung *Precision*:

$$\frac{TP}{TP+FP} \dots\dots\dots(3)$$

2.11. Penelitian Terkait

Terdapat beberapa penelitian analisa kepuasan yang serupa dengan penelitian ini, seperti:

Penerapan Algoritma C4.5 Untuk Menganalisis Kepuasan Pelanggan E-Commerce Shopee. aHasil penelitian ini memprediksi kepuasan pelanggan Shopee dengan tingkat akurasi sebesar 94,17%. Maka dapat disimpulkan pengujian melalui pengumpulan data dengan menerapkan Algoritma C4.5 dapat mengetahui kepuasan pelanggan pengguna Shopee[15].

Analisis Keputusan Pembelian Produk Terhadap Promosi Dan Viral Marketing Di Tiktok Shop Menggunakan Algoritma C4.5. Penelitian ini dilakukan untuk menganalisis keputusan pembelian produk di platform e-commerce TikTok dan dampaknya terhadap strategi promosi dan pemasaran viral. Hasil dari penelitian ini yaitu algoritma C4.5 dapat mengidentifikasi keputusan pembeli dengan nilai akurasi yang diperoleh yaitu sebesar 91.73%. Dengan demikian dapat disimpulkan bahwa persepsi dari tingkat keputusan pembeli produk di TikTok Shop terhadap promosi dan viral marketing mendapatkan hasil “Ya”[5].

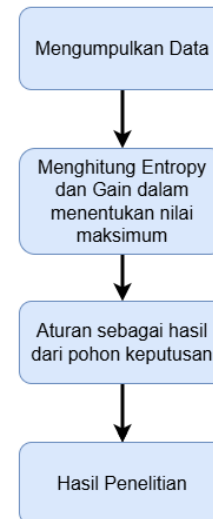
Analisis Tingkat Kepuasan Masyarakat Terhadap Kinerja Ketua Rw 007 Menggunakan Algoritma C4.5. Penelitian ini untuk mencari tahu seberapa puas masyarakat terhadap kinerja ketua RW 007 selama ini dan hasil penelitian ini sebagai bahan evaluasi untuk memajukan wilayah RW 007. Hasil dari penelitian ini menunjukkan bahwa faktor paling penting dalam kepuasan masyarakat adalah faktor kinerja dan

sebagian warga sudah merasa puas dengan kinerja saat ini dengan tingkat akurasi yang diraih pun baik yaitu di angka 96,15%[14].

3. METODE PENELITIAN

3.1. Metode Penelitian

Metode penelitian dilakukan dengan tahap-tahap sebagai berikut:



Gambar 1. Tahapan Penelitian

Untuk tahapan penelitian dimulai dari pengumpulan data dengan menyebarkan kuesioner melalui sosial media. Pengolahan data dimulai dari pemilahan, klasifikasi dengan menggunakan *tool* Algoritma C4.5 sebagai *tool* pengujian data yang telah direkap dengan mencari *entropy* dan *gain* guna menentukan akar dari pohon keputusan serta aturan yang terbentuk dari hasil pohon keputusan tersebut. Lalu di dapat hasil berupa *accuracy*, *class recall* dan *precision* yang didapat sebagai beberapa kesimpulan terhadap faktor yang mempengaruhi.

Penelitian yang digunakan dalam penelitian ini yaitu melalui pendekatan penelitian kuantitatif. Pendekatan kuantitatif merupakan metode pendekatan yang biasanya menekankan analisisnya pada data - data numerik atau angka yang digunakan untuk meneliti pada populasi atau sampel tertentu sesuai dengan banyaknya populasi dan sampel yang akan diuji[5].

3.2. Sumber Data

Data primer adalah data yang diperoleh langsung dari sumber data yang disimpulkan secara khusus dan berhubungan langsung dengan masalah yang diteliti[5]. Dalam penelitian ini data primer berupa hasil jawaban responden atas kuesioner yang diajukan. Dalam penelitian ini data primer bersumber dari penyebaran kuesioner kepada pengguna media sosial dan marketplace Shopee.

3.3. Populasi dan Sampel

Populasi adalah objek penelitian itu sendiri, pusat perhatian dan sumber data penelitian[5]. Populasi pada

penelitian ini adalah 100 masyarakat pengguna aplikasi marketplace Shopee.

Sampel adalah sekelompok elemen yang dipilih dari kelompok yang lebih besar dengan harapan mempelajari kelompok yang lebih kecil ini (sampel) akan mengungkapkan informasi penting tentang kelompok yang lebih besar (populasi)[5]. Kriteria sampel yang kami ambil adalah jenis kelamin dan usia.

Penentuan besar sampel dalam penelitian ini menggunakan rumus Lemeshow (1997), karena jumlah populasi tidak diketahui. Berikut rumus Lemeshow.

$$n = \frac{z^2 p (1-p)}{d^2} \dots \dots \dots (4)$$

Keterangan:

n = Jumlah sampel

z = Nilai standart = 1.96

p = Maksimal estimasi = 50% = 0.5

d = alpha (0,10) atau sampling error = 10%

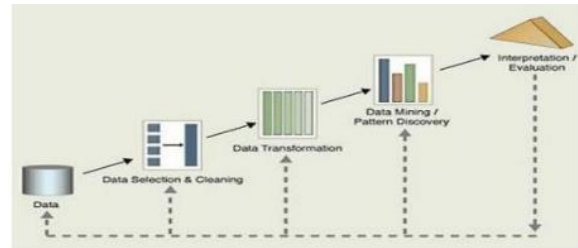
Kemudian diperoleh jumlah sampel minimal yang dibutuhkan dalam penelitian ini adalah 96 responden dan akan dibulatkan oleh peneliti menjadi 100 responden. Alasan peneliti menggunakan rumus Lemeshow (1997) adalah karena populasi sasaran terlalu besar dengan jumlah yang bervariasi[16].

3.4. Teknik Pengumpulan Data

Teknik pengumpulan data pada penelitian ini yaitu menggunakan kuesioner dengan media google formulir. Dalam tahap ini peneliti sudah menyiapkan kuesioner dengan media google formulir. Kuesioner yang digunakan dalam penelitian ini dirancang secara cermat dengan memperhatikan pertanyaan - pertanyaan yang relevan dan mampu menggali informasi yang diperlukan untuk menjawab tujuan penelitian[5].

3.5. Teknik Analisis Data

Teknik analisis yang digunakan pada penelitian ini adalah menggunakan Knowledge Discovery in Databases (KDD). Knowledge Discovery in Databases (KDD) adalah sekumpulan proses untuk menggali dan menganalisis sejumlah besar himpunan data dan mengekstrak informasi dan pengetahuan yang berguna. Langkah-langkah proses KDD ditampilkan pada gambar berikut:



Gambar 2. Tahapan Teknik Analisis Data (Sumber: Amanda Febriyani dkk, 2021)

Cara untuk mendapatkan atribut sebagai akar adalah dengan melibatkan perhitungan jumlah kasus dan jumlah target atribut. Setelah itu, perlu dihitung nilai entropy untuk menilai seberapa informatif suatu atribut input dalam menghasilkan atribut output. Rumus entropy sebagai berikut:

$$Entropy(S) = - \sum_{i=1}^n p_i * \log_2 P_i \dots \dots \dots (5)$$

Keterangan :

S : himpunan kasus

n : jumlah partisi

S pi : proporsi dari Si terhadap S

Adapun rumus gain dapat ditulis dalam persamaan berikut :

$$Gain S.A = Entropy(S) \sum_{i=1}^n \frac{|S_i|}{|S|} *$$

$$Entropy (S_i) \dots \dots \dots (6)$$

Keterangan :

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

|Si| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S [5]

4. HASIL DAN PEMBAHASAN

Pengumpulan data dengan menyebar kuesioner kepada 100 masyarakat pengguna aplikasi marketplace Shopee sebanyak 24 laki-laki dan 76 perempuan dengan berdasarkan umur.

Tabel 2. Data Responden

Usia	Jumlah
Kurang dari 18 Tahun	11
18 – 25 Tahun	87
26 – 35 Tahun	1
36 – 45 Tahun	1

Tabel 3. Entropy dan Gain

NODE	Atribut	Nilai	Jumlah Kasus	Pengaruh	Tidak Pengaruh	Entropy	Gain
1	Total		100	99	1	0,0807931	
	Jenis Kelamin	Laki-laki	24	24	0	0	0,00398224
		Perempuan	76	75	1	0,101067	
	Usia	kurang dari 18 Tahun	11	11	0	0	0,00201998
		18 - 25 Tahun	87	86	1	0,0905439	
		26 - 35 Tahun	1	1	0	0	
		36 - 45 Tahun	1	1	0	0	
	Pengalaman Pribadi	Pengaruh	94	93	1	0,0849956	0,00509973
		Tidak Pengaruh	6	6	0	0	
	Kepercayaan Sumber	Periksa	98	97	1	0,0821431	0,00029295
		Tidak Periksa	2	2	0	0	
	Kualitas Informasi	Penting	99	98	1	0,081462	0,00014573
		Tidak Penting	1	1	0	0	
	Kuantitas Ulasan	Pengaruh	91	91	0	0	0,03549989
		Tidak Pengaruh	9	8	1	0,5032583	
	Jenis Platform	Mencari	87	87	0	0	0,02993147
		Tidak Mencari	13	12	1	0,3912436	
	Pengaruh Sosial	Pernah	90	90	0	0	0,03389358
		Tidak Pernah	10	9	1	0,4689956	

Hasil perhitungan tabel 3 dapat dijabarkan sebagai berikut:

Rumus Entropy: $Entropy(S) = -\sum_{i=1}^n p_i * \log_2 p_i$

Berikut perhitungan untuk Entropy Total dengan node pertama pada pohon keputusan.

$$Entropy (Total) = (- [99/100] * [\log]_2 (99/100)) + (- [1/100] * [\log]_2 (1/100)) = 0,0807931$$

Selanjutnya adalah menghitung Entropy dari masing-masing atribut yang hasil perhitungannya dapat dilihat pada tabel 3. Sebagai contoh penjabaran perhitungan khusus dimuat untuk atribut Kuantitas Ulasan dengan nilai Pengaruh dan Tidak Pengaruh untuk perhitungan Entropy dan Gainnya.

$$Entropy (Atribut Kuantitas Ulasan, Pengaruh) = (- [91/91] * [\log]_2 (91/91)) + (- [0/91] * [\log]_2 (0/91)) = 0$$

$$Entropy (Atribut Kuantitas Ulasan, Tidak Pengaruh) = (- [8/9] * [\log]_2 (8/9)) + (- [1/9] * [\log]_2 (1/9)) = 0,5032583$$

Lalu perhitungan untuk Gain :

Rumus Gain :

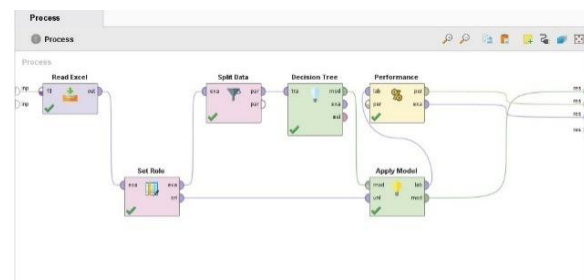
$$Gain S.A = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

$$Gain (Atribut Kuantitas Ulasan) = 0,0807931 - ((91/100 * 0) - (9/100 * 0,5032583)) = 0,0354999$$

Dapat dilihat dari tabel 3 diketahui gain terbesar dari masing-masing atribut adalah atribut Kuantitas Ulasan dengan nilai 0,0354999. Maka atribut Kuantitas Ulasan menjadi akar (*root*) seperti terlihat pada gambar 4 yang merupakan hasil dari tools RapidMiner.

Langkah terakhir penerapan Algoritma C4.5 yaitu menguji dengan tools RapidMiner yang

membutuhkan penyesuaian hasil perhitungan secara manual. Gambaran proses pengoperasian sistem pada data ditampilkan di sini.

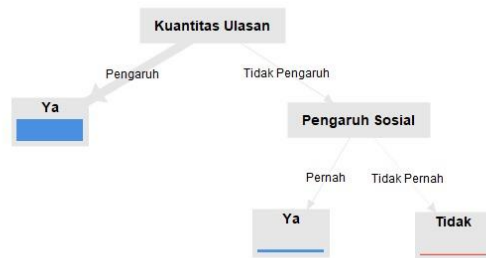


Gambar 3. Proses RapidMiner

Proses menguji tools RapidMiner dimulai dengan menggunakan operator Read Excel lalu memasukkan data excel, dimana tipe data dapat diatur sesuai kebutuhan. Setelah itu, hubungkan INP pada papan set dengan FIL pada Read Excel untuk memuat data. Kemudian, gunakan operator Set Role, atur attribute name dan target sesuai kebutuhan. Selanjutnya hubungkan OUT pada Read Excel dengan EXA pada Set Role, lalu klik apply. Kemudian, gunakan operator Split Data dengan rasio 0.7 melalui pengaturan partisi lalu klik OK. Lalu hubungkan EXA pada Set Role dengan EXA pada Split Data. Kemudian, gunakan operator Decision Tree, ubah minimal leaf size menjadi 1, lalu hubungkan PAR pada Split Data ke TRA pada Decision Tree. Selanjutnya, gunakan operator Apply Model dengan menghubungkan ORI pada Set Role ke UNL pada Apply Model, lalu hubungkan MOD pada Decision Tree ke MOD pada Apply Model. Setelah itu, hubungkan MOD pada Apply Model ke RES pada papan set, lalu pilih Play untuk melihat hasil nya. Terakhir, gunakan operator Performance dan hubungan LAB Performance dengan LAB pada Apply Model, lalu hubungkan PER pada Performance ke

RES pada papan set, klik Play untuk melihat hasil nya. Kemudian, hubungan EXA pada Performance ke RES pada papan set, klik Play untuk melihat hasil nya. Dari langkah-langkah sebelumnya terdapat 3 hasil yaitu Example Set, Decision Tree dan Performance Vector.

Selanjutnya proses pada RapidMiner dilakukan untuk membentuk pohon keputusan dan *Confusion Matrix*.



Gambar 4. Pohon Keputusan

Berikut adalah hasil dari pengolahan pada Rapidminer berupa bentuk pohon keputusan dari data yang sudah diolah dimana data Kuantitas Ulasan menjadi akar dalam pohon keputusan. Dari pohon keputusan diatas maka terbentuklah suatu aturan sebagai berikut :

Jika Kuantitas Ulasan = Pengaruh, maka hasilnya Ya (Ya = 91, Tidak = 0)

Jika Kuantitas Ulasan = Tidak Pengaruh, Pengaruh Sosial = Pernah, maka hasilnya Ya (Ya = 8, Tidak = 0)

Jika Kuantitas Ulasan = Tidak Pengaruh, Pengaruh Sosial = Tidak Pernah, maka hasilnya Tidak (Ya = 0 dan Tidak = 1)

Berikut adalah hasil perhitungan *Confusion Matrix* untuk mencari nilai accuracy, precision dan recall dari metode Algoritma C4.5.

accuracy: 100.00%			
	true Ya	true Tidak	class precision
pred Ya	99	0	100.00%
pred Tidak	0	1	100.00%
class recall	100.00%	100.00%	

Gambar 5. Confusion Matrix

Berdasarkan gambar diatas yang merupakan perhitungan dari *tools* RapidMiner maka diketahui:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} = \frac{99 + 1}{99 + 0 + 0 + 1} = 1 = 100\%$$

$$TPRate(Recall) = \frac{TP}{TP + FN} = \frac{99}{99 + 0} = 1 = 100\%$$

$$Precision = \frac{TP}{TP + FP} = \frac{99}{99 + 0} = 1 = 100\%$$

Dalam hal ini tingkat Accuracy sebesar 100% menunjukkan prediction tingkat pengaruh yang sangat baik dalam semua variable. Untuk Precision diperoleh 100% dikatakan tidak adanya kasus *true negative* yang artinya semakin baik model dalam menghindari

kesalahan memprediksi kelas positif yang sebenarnya tidak. Sedangkan untuk Recall sebesar 100% untuk kelas *true* bernilai sempurna.

Semua parameter yang digunakan untuk membuat pohon keputusan mematuhi standar pohon keputusan RapidMiner. Temuan ini memuatkan temuan perhitungan manual yang dilakukan di Microsoft Excel dan diuji menggunakan RapidMiner untuk membangun model faktor penentu pengguna terhadap keputusan pembelian di Shopee.

5. KESIMPULAN DAN SARAN

Berdasarkan data di atas dapat diketahui bahwa dari enam atribut yang digunakan hanya dua atribut saja yang muncul dalam pohon keputusan. Atribut Kuantitas Ulasan menjadi *node* akar (*root*) lalu diikuti dengan atribut Pengaruh Sosial. Faktor dengan gain terbesar diraih oleh atribut Kuantitas Ulasan dengan nilai gain 0,0354999. Tingkat akurasi yang diraih pun sempurna yaitu di angka 100%. Secara keseluruhan, hasil penelitian menunjukkan bahwa seluruh responden menjadikan faktor Kuantitas Ulasan sebagai faktor penentu terhadap keputusan pembelian pada Marketplace Shopee.

DAFTAR PUSTAKA

- [1] T. Simorangkir, "Pengaruh Ulasan Konsumen Online, Kualitas Pelayanan Elektronik dan Harga Terhadap Minat Beli," *Univ. HKBP Nommensen*, no. 2504, pp. 1–9, 2022.
- [2] A. I. Rahmawati, "Pengaruh Online Customer Review, Online Customer Rating Dan Kepercayaan Terhadap Keputusan Pembelian Online (Studi Kasus Pada Mahasiswa Feb Universitas PGRI Semarang)," *J. Ilm. Manaj. Bisnis dan Ekon. Kreat.*, vol. 1, no. 1, pp. 18–23, 2021, doi: 10.26877/jibeka.v1i1.3.
- [3] D. Arif and D. A. Pramestie, "Pengaruh Promosi Dan Review Produk Pada Marketplace Shopee Terhadap Keputusan Pembelian Di Umaha," *Ecopreneur.12*, vol. 4, no. 2, p. 172, 2021, doi: 10.51804/econ12.v4i2.1013.
- [4] R. Hermiati, A. Asnawati, and I. Kanedi, "Pembuatan E-Commerce Pada Raja Komputer Menggunakan Bahasa Pemrograman Php Dan Database Mysql," *J. Media Infotama*, vol. 17, no. 1, pp. 54–66, 2021, doi: 10.37676/jmi.v17i1.1317.
- [5] N. Nurliana, N. Suarna, and W. Prihatono, "Analisis Keputusan Pembelian Produk Terhadap Promosi Dan Viral Marketing Di Tiktok Shop Menggunakan Algoritma C4.5," *J. Mhs. Tek. Inform.*, vol. 8, no. 3, pp. 2870–2876, 2024.
- [6] C. Zai, "Implementasi Data Mining Sebagai Pengolahan Data," *J. Portal Data*, vol. 2, no. 3, pp. 1–12, 2022, [Online]. Available: <http://portaldata.org/index.php/portaldata/article/view/107>
- [7] P. B. N. Setio, D. R. S. Saputro, and Bowo Winarno, "Klasifikasi Dengan Pohon Keputusan

- Berbasis Algoritme C4.5,” *Prism. Pros. Semin. Nas. Mat.*, vol. 3, pp. 64–71, 2020.
- [8] I. Romli and A. T. Zy, “Penentuan Jadwal Overtime Dengan Klasifikasi Data Karyawan Menggunakan Algoritma C4.5,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 4, no. 2, pp. 694–702, 2020.
- [9] K. Aidi Saputra, J. Tata Hardinata, M. Ridwan Lubis, S. Retno Andani, and I. Syahputra Saragih, “KLIK: Kajian Ilmiah Informatika dan Komputer Klasifikasi Algoritma C4.5 Dalam Penerapan Tingkat Kepuasan Siswa Terhadap Media Pembelajaran Online,” *Media Online*, vol. 1, no. 3, pp. 113–118, 2020, [Online]. Available: <https://djournals.com/klik>
- [10] F. Siska and S. Heni, “Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5,” *J. Teknol. dan Sist. Inf.*, vol. 2, no. 4, pp. 89–95, 2021.
- [11] Y. Partogi and A. Pasaribu, “Perancangan Metode Decision Tree Terhadap Sistem Perpustakaan STMIK Kuwera,” *J. Sist. Inf. dan Teknol.*, vol. 1, no. 2, pp. 20–25, 2022, doi: 10.56995/sintek.v1i2.4.
- [12] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, “Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner,” *JBASE - J. Bus. Audit Inf. Syst.*, vol. 4, no. 1, pp. 13–21, 2021, doi: 10.30813/jbase.v4i1.2729.
- [13] D. Darwis, N. Siskawati, and Z. Abidin, “Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional,” *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.
- [14] M. Andreas Rayando, I. Fajri, F. Kurniawan, M. Stefanus Rafael Sihotang, and S. Suharjanti, “Analisis Tingkat Kepuasan Masyarakat Terhadap Kinerja Ketua Rw 007 Menggunakan Algoritma C4.5,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 4, pp. 7463–7469, 2024, doi: 10.36040/jati.v8i4.10334.
- [15] A. Steffany, K. A. Aldzikri, M. A. R. Tricahya, and M. I. Arfiandono, “Penerapan Algoritma C4.5 Untuk Menganalisis Kepuasan E-Commerce Shopee,” *J. Nas. Komputasi dan Teknol. Inf.*, vol. 6, no. 6, pp. 700–706, 2023, doi: 10.32672/jnkti.v6i6.7112.
- [16] M. H. Setiawan, R. Komarudin, and D. N. Kholifah, “Pengaruh Kepercayaan, Tampilan Dan Promosi Terhadap Keputusan Pemilihan Aplikasi Marketplace,” *J. Infortech*, vol. 4, no. 2, p. 141, 2022.