

Analisis Tingkat Kesadaran Pengguna E-Commerce Terhadap Keamanan Dan Privasi Data Pribadi Menggunakan Algoritma C4.5

Hammam Zufar¹, Nabila Amelia¹, Leonny Setiawan¹, Syahramdan¹, Maisyaroh^{2,*}, Martua Hami Siregar²

¹Sistem Informasi, Universitas Bina Sarana Informatika; Jl. Raya Jatiwaringin No.18, RT.009/RW.005, Jaticempaka, Kec. Pd. Gede, Kota Bks, Jawa Barat 17411, Kota Bekasi, Indonesia, e-mail : hammamzufar@gmail.com, nabilaamelia600@gmail.com, leonnysetiawan9@gmail.com, syahramdan0202@gmail.com

²Teknologi Informasi, Universitas Bina Sarana Informatika; Jl. Kramat Raya no.98 Jakarta, Indonesia, e-mail: maysaroh.msy@bsi.ac.id, martua.mhe@bsi.ac.id

* Korespondensi: e-mail: maysaroh.msy@bsi.ac.id

Diterima: 29 November 2025 ; Review: 07 Desember 2025; Disetujui: 31 Desember 2025

Cara sitasi: Zufar H, Amelia N, Setiawan L, Syahramdan, Maisyaroh, Siregar HM. 2025. Analisis Tingkat Kesadaran Pengguna E-Commerce Terhadap Keamanan Dan Privasi Data Pribadi Menggunakan Algoritma C4.5. Bina Insani ICT Journal. Vol 12 (2): 181-189.

Abstrak: Perkembangan teknologi digital yang pesat serta meningkatnya pemanfaatan platform *E-commerce* mendorong pengguna untuk semakin sering membagikan data pribadi mereka secara mudah dan cepat, sehingga berpotensi meningkatkan risiko terhadap keamanan dan privasi informasi. Penelitian ini bertujuan untuk mengkaji tingkat kesadaran pengguna *e-commerce* dalam menjaga keamanan dan privasi data pribadi, sekaligus mengidentifikasi faktor-faktor yang memengaruhinya. Pendekatan penelitian yang diterapkan adalah kuantitatif dengan memanfaatkan teknik data mining melalui penggunaan algoritma C4.5. Data dikumpulkan menggunakan kuesioner yang dibagikan kepada 67 responden pengguna *e-commerce*, kemudian dianalisis menggunakan perangkat lunak RapidMiner untuk menghasilkan model klasifikasi dalam bentuk pohon keputusan. Hasil pengolahan data menunjukkan bahwa atribut P16 menjadi variabel yang paling dominan dalam memengaruhi tingkat kesadaran pengguna terkait keamanan dan privasi data pribadi. Pengujian model dengan confusion matrix memperoleh nilai akurasi sebesar 75%, dengan tingkat *recall* pada kategori kesadaran tinggi sebesar 91,67% dan kategori kesadaran sedang sebesar 50%. Temuan tersebut mengindikasikan bahwa algoritma C4.5 memiliki performa yang cukup efektif dalam mengklasifikasikan tingkat kesadaran pengguna, serta menegaskan perlunya peningkatan pemahaman mengenai keamanan digital guna meminimalkan risiko kebocoran data pribadi pada layanan *e-commerce*.

Kata Kunci: Keamanan Akun, E-Commerce, Data Pribadi, Data Mining, Algoritma C4.5

Abstract: *The accelerated advancement of digital technology along with the growing adoption of e-commerce platforms has led users to increasingly disclose personal data online, thereby potentially heightening risks related to information security and privacy. This study employs a quantitative approach using data mining techniques through the implementation of the C4.5 algorithm. The research data were collected via questionnaires distributed to 67 e-commerce users and subsequently processed using RapidMiner software to construct a decision tree-based classification model. The analysis results indicate that attribute P16 is the most influential factor in determining users' awareness of personal data security and privacy. Model evaluation using a confusion matrix achieved an accuracy of 75%, with recall values of 91.67% for the high-awareness category and 50% for the moderate-awareness category. These findings demonstrate that the C4.5 algorithm performs adequately in classifying users' awareness levels and highlight*

the importance of enhancing digital security literacy to reduce the risk of personal data breaches on e-commerce platforms.

Keywords: *Account Security, E-Commerce, Personal Data, Data Mining, C4.5 Algorithm*

1. Pendahuluan

Pesatnya perkembangan teknologi informasi, khususnya pada media komunikasi berbasis internet, telah membuat proses pertukaran informasi di masyarakat menjadi semakin cepat dan mudah. Pertumbuhan penggunaan internet tersebut turut mendorong meningkatnya pemanfaatan media sosial sebagai sarana untuk berbagi serta memperoleh informasi secara real-time. Dalam kehidupan masyarakat modern, media sosial tidak lagi hanya berfungsi sebagai alat komunikasi, tetapi juga telah menjadi bagian dari aktivitas sehari-hari dalam mencari, menyimpan, dan menyebarkan informasi [1].

Di sisi lain, berbagai manfaat yang ditawarkan media sosial juga diiringi dengan risiko yang berkaitan dengan kebocoran informasi pribadi. Pengguna kerap kali tidak menyadari bahwa data yang mereka bagikan, seperti nama lengkap, nomor telepon, alamat, hingga foto pribadi, berpotensi dimanfaatkan oleh pihak-pihak yang tidak bertanggung jawab. Kondisi ini menjadikan aspek keamanan informasi dan perlindungan privasi sebagai isu yang penting dalam pemanfaatan teknologi digital [2].

Kasus pelanggaran privasi dan kebocoran data pribadi menunjukkan kecenderungan peningkatan dari tahun ke tahun. Sejumlah insiden kebocoran data berskala besar yang terjadi pada platform e-commerce di Indonesia mengindikasikan bahwa tingkat kewaspadaan pengguna terhadap perlindungan data pribadi masih relatif rendah. Selain itu, kebiasaan penggunaan kata sandi yang lemah, perilaku membagikan informasi pribadi secara berlebihan di media sosial, serta rendahnya pemahaman mengenai pengaturan privasi pada aplikasi semakin meningkatkan kerentanan terhadap penyalahgunaan data [3].

Meskipun pemerintah telah menetapkan Undang-Undang Perlindungan Data Pribadi sebagai bentuk perlindungan hukum, penerapannya di masyarakat belum sepenuhnya optimal. Pengguna masih belum memahami secara menyeluruh hak-hak mereka dalam pengelolaan data pribadi serta cara menerapkan langkah-langkah mitigasi risiko digital secara mandiri [4]. Hal ini menegaskan bahwa perlindungan data pribadi tidak hanya menjadi tanggung jawab pemerintah dan penyedia layanan digital, tetapi juga membutuhkan kesadaran dan pemahaman dari setiap individu sebagai pengguna layanan [5].

Selain itu, peningkatan transaksi digital melalui layanan e-commerce terus berlangsung seiring dengan kemudahan akses terhadap teknologi. Kondisi ini turut memperbesar potensi ancaman terhadap keamanan data pribadi yang bersumber dari aktivitas daring pengguna [6]. Oleh karena itu, diperlukan penelitian yang mampu menggambarkan tingkat kesadaran pengguna e-commerce terhadap keamanan dan privasi data pribadi agar dapat memberikan rekomendasi yang lebih tepat dalam upaya peningkatan literasi keamanan digital.

Algoritma C4.5 telah banyak dimanfaatkan dalam penelitian data mining karena kemampuannya menghasilkan model klasifikasi berupa pohon keputusan yang mudah dipahami serta efektif dalam mengidentifikasi atribut-atribut yang berpengaruh. Penelitian yang dilakukan oleh Tukino [7] dan Wijaya serta Bakti [8] menunjukkan bahwa algoritma C4.5 mampu menghasilkan klasifikasi yang akurat pada studi prediksi keuntungan dan penjualan, sementara Amalina et al. [9] menerapkannya dalam klasifikasi keamanan data pengguna. Namun demikian, penelitian-penelitian tersebut masih berfokus pada prediksi dan klasifikasi data secara umum, tanpa secara khusus mengkaji tingkat kesadaran pengguna e-commerce terhadap keamanan dan privasi data pribadi. Oleh sebab itu, penelitian ini diarahkan untuk mengisi celah tersebut dengan menerapkan algoritma C4.5 dalam menganalisis tingkat kesadaran pengguna e-commerce serta mengidentifikasi faktor-faktor dominan yang memengaruhinya, sejalan dengan tujuan untuk mendukung peningkatan literasi keamanan digital.

Pesatnya pertumbuhan transaksi digital pada platform e-commerce turut meningkatkan risiko kebocoran dan penyalahgunaan data pribadi, sementara tingkat kesadaran pengguna terhadap keamanan akun dan privasi informasi masih belum sepenuhnya optimal. Berdasarkan

kondisi tersebut, penelitian ini bertujuan untuk menganalisis tingkat kesadaran pengguna e-commerce terhadap keamanan dan privasi data pribadi menggunakan algoritma C4.5, guna mengidentifikasi faktor-faktor dominan yang memengaruhinya serta menghasilkan model klasifikasi yang dapat dijadikan dasar rekomendasi dalam upaya peningkatan literasi keamanan digital. Adapun keterbatasan penelitian ini terletak pada jumlah responden yang relatif terbatas serta ruang lingkup data yang hanya mencakup pengguna e-commerce.

2. Metode Penelitian

2.1. Tahapan Penelitian

Tahapan penelitian dilaksanakan melalui beberapa langkah sebagai berikut:

1. Identifikasi Masalah

Tahapan awal penelitian dimulai dengan penelusuran dan penentuan permasalahan, yaitu meningkatnya potensi kebocoran serta penyalahgunaan data pribadi pada platform *e-commerce* yang belum diiringi oleh kesadaran pengguna terhadap aspek keamanan dan privasi data. Pada tahap ini juga ditetapkan tujuan penelitian, yaitu untuk mengkaji tingkat kesadaran pengguna *e-commerce* beserta faktor-faktor yang memengaruhinya sebagai landasan dalam upaya meningkatkan literasi keamanan digital.

2. Pengumpulan Data

Proses pengambilan data dilakukan melalui pendistribusian kuesioner kepada pengguna platform *e-commerce*. Kuesioner tersebut disusun untuk menilai tingkat kesadaran pengguna terhadap aspek keamanan dan privasi data pribadi yang mencakup dimensi pengetahuan, sikap, serta perilaku. Hasil dari tahap ini berupa data primer yang dikumpulkan dari 67 responden dan selanjutnya direkap dalam bentuk file Microsoft Excel.

3. Preprocessing Data

Tahap preprocessing bertujuan untuk mempersiapkan data agar layak digunakan dalam proses pemodelan. Kegiatan yang dilakukan meliputi pengecekan kelengkapan data, penghapusan data yang tidak valid, serta penyesuaian format data. Hasil dari tahap ini adalah dataset yang telah dibersihkan dan siap untuk diolah lebih lanjut.

4. Penerapan Algoritma C4.5

Pada tahapan ini dilakukan pembangunan model klasifikasi menggunakan algoritma C4.5 dengan memanfaatkan perangkat lunak RapidMiner. Model yang dihasilkan berbentuk pohon keputusan yang dibangun berdasarkan perhitungan nilai entropy dan information gain.

5. Evaluasi Model

Evaluasi terhadap model klasifikasi dilakukan menggunakan confusion matrix guna memperoleh nilai akurasi, precision, dan recall sebagai indikator kinerja model..

6. Interpretasi Hasil

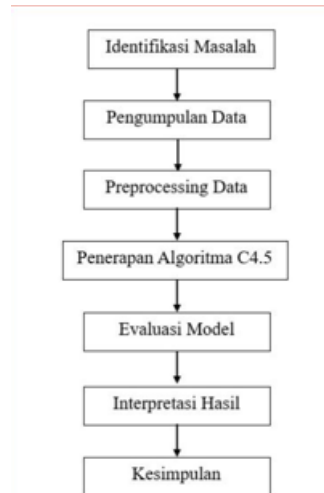
Tahap interpretasi hasil dilakukan dengan menganalisis struktur pohon keputusan yang dihasilkan serta mengidentifikasi atribut-atribut dominan yang memengaruhi tingkat kesadaran pengguna *e-commerce*.

7. Kesimpulan

Tahapan terakhir penelitian difokuskan pada penyusunan kesimpulan yang merujuk pada hasil pemodelan dan evaluasi model, serta perancangan rekomendasi yang dapat dimanfaatkan sebagai langkah untuk meningkatkan literasi keamanan digital.

2.2. Diagram Alir Tahapan Penelitian

Untuk memperjelas alur pelaksanaan penelitian, disajikan diagram alur yang menggambarkan tahapan-tahapan yang dilakukan dalam penelitian ini.



Sumber: Hasil Penelitian (2025)

Gambar 1. Diagram Alir Tahapan Penelitian

Gambar tersebut menunjukkan diagram alir tahapan penelitian, yang dimulai dari identifikasi masalah, dilanjutkan dengan pengumpulan data dan preprocessing data, kemudian penerapan algoritma C4.5, evaluasi model, interpretasi hasil, dan diakhiri dengan penarikan kesimpulan.

2.3 Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif karena data yang dianalisis berupa hasil pengukuran tingkat kesadaran pengguna e-commerce terhadap keamanan dan privasi data pribadi yang diperoleh melalui kuesioner dengan skala numerik. Pendekatan ini bertujuan untuk menggambarkan serta memprediksi pola kesadaran pengguna secara objektif berdasarkan temuan empiris [10].

Dalam proses analisis data, algoritma C4.5 dipilih sebagai metode klasifikasi utama. Algoritma ini merupakan pengembangan dari ID3 yang membangun model pohon keputusan dengan menggunakan perhitungan gain ratio, sehingga mampu menghasilkan model yang lebih stabil dan memiliki tingkat akurasi yang lebih baik. Berdasarkan literatur, algoritma C4.5 bekerja melalui tahapan perhitungan nilai entropy dan information gain, pemilihan atribut terbaik sebagai node, serta penerapan proses pruning untuk menyederhanakan struktur pohon agar lebih optimal dan mudah diinterpretasikan [7] [11] [8]. Oleh karena itu, algoritma ini dinilai sesuai untuk mengidentifikasi faktor-faktor yang memengaruhi tingkat kesadaran pengguna terhadap keamanan dan privasi data pribadi..

2.4 Sumber dan Jenis Data

Penelitian ini menggunakan dua jenis data, yaitu:

1. Data Primer

Data primer dalam penelitian ini diperoleh melalui penyebaran kuesioner kepada pengguna e-commerce. Instrumen kuesioner dirancang menggunakan skala Likert 1–5, dengan keterangan 1 = sangat tidak setuju, 2 = tidak setuju, 3 = netral, 4 = setuju, dan 5 = sangat setuju. Responden diminta untuk memberikan penilaian terhadap sejumlah pernyataan yang mengukur aspek pemahaman, sikap, dan perilaku yang berkaitan dengan keamanan dan privasi data pribadi. Pemilihan ketiga aspek tersebut didasarkan pada konsep *information security awareness* yang mencakup dimensi kognitif (pemahaman), afektif (sikap), dan konatif (perilaku). Konsep ini telah banyak digunakan dalam penelitian kesadaran keamanan informasi untuk menilai tingkat pemahaman serta tindakan pengguna dalam melindungi data pribadi [1] [2] [9].

2. Data Sekunder

Data sekunder pada penelitian ini dikumpulkan dari beragam referensi, antara lain jurnal ilmiah, buku, serta literatur yang berkaitan dengan keamanan informasi, privasi digital, e-commerce, dan algoritma C4.5 yang dijadikan sebagai dasar teoritis dalam penelitian. [7] [8].

2.5 Populasi dan Sampel

Populasi dalam penelitian ini mencakup pengguna e-commerce yang secara aktif melakukan transaksi digital. Platform e-commerce yang digunakan oleh responden antara lain Tokopedia, Shopee, Lazada, Bukalapak, Blibli, dan platform sejenis lainnya..

Teknik pengambilan sampel yang digunakan adalah *non-probability sampling*, dengan metode penyebaran kuesioner secara acak kepada pengguna e-commerce. Kuesioner disebarkan kepada mahasiswa, masyarakat umum, serta pengguna media sosial yang memanfaatkan aplikasi e-commerce dalam aktivitas belanja daring.

Berdasarkan hasil pengumpulan data, diperoleh sebanyak 67 responden yang dijadikan sebagai sampel penelitian. Jumlah sampel tersebut telah memenuhi batas minimal yang direkomendasikan dalam analisis klasifikasi data mining, di mana jumlah data di atas 30–50 responden dianggap memadai untuk membangun model klasifikasi dan melakukan evaluasi kinerja awal. Hal ini khususnya berlaku pada penelitian berskala kecil hingga menengah yang bersifat eksploratif dengan penerapan algoritma C4.5 [10] [12].

2.6 Teknik Pengumpulan data

Proses pengumpulan data pada penelitian ini dilaksanakan dengan mendistribusikan kuesioner secara online melalui *Google Form* kepada responden pengguna *e-commerce*, yang mencakup 20 item pertanyaan.

Kuesioner tersebut disusun berdasarkan beberapa indikator yang diukur melalui total 20 item pernyataan (P1–P20), dengan rincian sebagai berikut::

1. Kebiasaan berbagi informasi pribadi, terdiri dari 4 pertanyaan (P1–P4).
2. Penggunaan kata sandi, terdiri dari 4 pertanyaan (P5–P8).
3. Pemahaman risiko keamanan digital, terdiri dari 4 pertanyaan (P9–P12).
4. Pengetahuan pengaturan privasi akun, terdiri dari 4 pertanyaan (P13–P16).
5. Persepsi terhadap ancaman kebocoran data, terdiri dari 4 pertanyaan (P17–P20).

Data respon kemudian direkap dalam format *Microsoft Excel* sebagai dataset utama penelitian.

2.7 Teknik Analisa Data

Tahapan analisis data dilakukan melalui beberapa langkah berikut:

1. *Data Cleaning*
Memeriksa konsistensi data, menghapus nilai kosong, dan memastikan seluruh respon layak digunakan untuk pemodelan.
2. Klasifikasi Menggunakan Algoritma C4.5
Algoritma C4.5 digunakan karena mampu menghasilkan pohon keputusan yang jelas dan mudah dipahami. Algoritma ini menentukan node utama berdasarkan nilai *information gain* tertinggi dari masing-masing atribut dan kemudian membangun model klasifikasi [9].
Proses perhitungan C4.5 melibatkan:
 1. *Entropy* untuk tingkat ketidakteraturan data.
 2. *Information Gain* untuk menentukan atribut paling berpengaruh.
 3. Pembentukan struktur pohon keputusan yang hierarkis.
3. Pemodelan Menggunakan RapidMiner
Pemodelan dilakukan menggunakan software RapidMiner, dengan langkah-langkah:
 1. Mengimpor dataset hasil kuesioner.
 2. Menentukan *atribut class* yaitu Tingkat Kesadaran.
 3. Menggunakan operator *Decision Tree (C4.5)*.
 4. Menjalankan proses *training model*.
 5. Mengevaluasi model menggunakan metode *Accuracy*.

RapidMiner dipilih karena mampu melakukan pemodelan klasifikasi secara visual dan menghasilkan pohon keputusan yang mudah diinterpretasikan [8].

3. Hasil dan Pembahasan

3.1. Hasil Pengolahan Data

Data kuesioner yang diperoleh dari 67 responden dikumpulkan dalam bentuk data mentah menggunakan file Microsoft Excel yang memuat jawaban untuk setiap butir pertanyaan P1–P20, sebagaimana ditunjukkan pada Tabel 1 Daftar Butir Pertanyaan Kuesioner. Data tersebut selanjutnya diimpor ke dalam perangkat lunak RapidMiner untuk dilakukan preprocessing dan pemodelan menggunakan algoritma C4.5, di mana atribut dengan nilai *information gain* tertinggi dipilih sebagai akar pohon keputusan dan percabangan dibentuk berdasarkan atribut lainnya. Hasil pemodelan menunjukkan bahwa atribut P16 menjadi node utama atau akar pada struktur pohon keputusan yang dihasilkan, yang merepresentasikan perilaku pengguna dalam melakukan pembaruan aplikasi e-commerce secara berkala sebagai upaya menjaga keamanan sistem. Temuan ini mengindikasikan bahwa kesadaran pengguna dalam memperbarui aplikasi, yang berkaitan langsung dengan perlindungan sistem dari celah keamanan dan ancaman siber, merupakan faktor paling dominan dalam menentukan tingkat kesadaran pengguna terhadap keamanan dan privasi data pribadi.

Tabel 1. Kuesioner Analisis Tingkat Kesadaran

Kode	Pernyataan	1	2	3	4	5
P1	Saya memahami pentingnya menjaga kerahasiaan data pribadi saat berbelanja online.					
P2	Saya mengetahui jenis data pribadi yang dikumpulkan aplikasi e-commerce (misalnya nomor telepon, alamat, dan data pembayaran).					
P3	Saya tahu bahwa data pribadi dapat disalahgunakan jika jatuh ke pihak yang tidak bertanggung jawab.					
P4	Saya membaca kebijakan privasi aplikasi e-commerce sebelum menggunakan aplikasinya.					
P5	Saya merasa aman menyimpan informasi kartu debit/kredit di aplikasi e-commerce.					
P6	Saya merasa aplikasi e-commerce telah melindungi data pribadi penggunanya dengan baik.					
P7	Saya setuju jika aplikasi e-commerce menggunakan data saya untuk meningkatkan layanan.					
P8	Saya khawatir data pribadi saya dapat dibocorkan atau dicuri oleh pihak lain.					
P9	Saya menggunakan kata sandi yang kuat dan tidak mudah ditebak.					
P10	Saya rutin mengganti kata sandi akun aplikasi e-commerce saya.					
P11	Saya tidak membagikan kode OTP atau informasi login kepada orang lain.					
P12	Saya selalu memeriksa keaslian situs atau aplikasi sebelum login ke aplikasi e-commerce.					
P13	Saya mengetahui bahwa aplikasi e-commerce memiliki kebijakan privasi resmi.					
P14	Saya memahami bagaimana aplikasi e-commerce menggunakan data pengguna untuk tujuan komersial.					
P15	Saya mengetahui cara mengatur izin akses data (lokasi, kontak, kamera) di aplikasi e-commerce.					
P16	Saya selalu memperbarui aplikasi e-commerce untuk memastikan sistem keamanannya tetap terjaga.					
P17	Saya membaca notifikasi atau pemberitahuan dari aplikasi e-commerce yang berkaitan dengan keamanan akun.					
P18	Saya memahami risiko menggunakan jaringan Wi-Fi publik untuk mengakses akun aplikasi e-commerce.					
P19	Saya mengetahui cara melaporkan kejadian jika akun aplikasi e-commerce saya diretas atau disalahgunakan.					
P20	Saya menyadari bahwa memberikan data pribadi di luar aplikasi resmi e-commerce dapat membahayakan keamanan saya.					

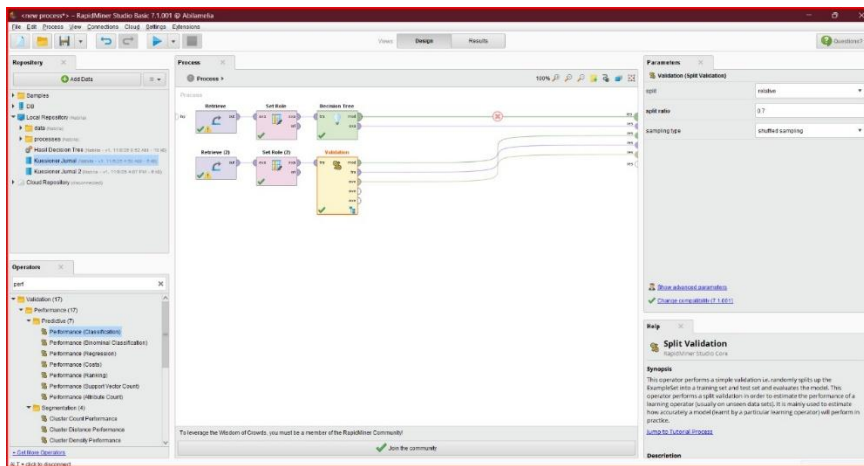
Sumber: Hasil Penelitian (2025)

Kemudian hasil dari kuesioner tersebut diolah dengan *Microsoft Excel*, kemudian dilanjutkan dengan Aplikasi *Rapid Miner*

Jenis Kelamin	Usia	Pendidikan	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	Tingkat Kesadaran	
Perempuan	15-25	SMA/SMK	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	Tinggi	
Perempuan	36-45	SMA/SMK	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	Tinggi	
Perempuan	15-25	SMA/SMK	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	Tinggi	
Laki-Laki	15-25	SMA/SMK	5	4	5	4	4	3	3	3	5	3	5	4	5	4	4	4	5	4	3	5	Tinggi	
Perempuan	15-25	SMA/SMK	4	3	4	4	3	4	4	4	3	4	4	4	4	4	4	4	4	3	4	4	Tinggi	
Laki-Laki	26-35	SI	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	Tinggi	
Perempuan	15-25	SMA/SMK	5	4	4	4	3	4	4	5	4	5	4	5	4	4	4	4	4	4	3	5	Tinggi	
Perempuan	15-25	SMA/SMK	3	3	4	4	4	4	3	4	4	2	3	2	3	4	4	3	3	4	4	3	Sedang	
Laki-Laki	36-45	SMA/SMK	5	4	4	4	2	2	4	4	5	4	5	4	4	3	4	4	4	4	4	4	5	Tinggi
Laki-Laki	15-25	SMA/SMK	4	3	4	3	3	4	4	4	3	5	5	5	5	5	5	4	1	3	3	5	Tinggi	
Laki-Laki	15-25	SMA/SMK	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	Tinggi	
Perempuan	15-25	SMA/SMK	5	4	5	4	3	3	4	4	4	5	4	4	4	3	4	4	4	3	4	5	Tinggi	
Perempuan	15-25	SMA/SMK	4	4	5	4	4	3	4	4	5	2	4	4	5	3	4	3	4	3	4	5	Tinggi	
Perempuan	26-35	SMA/SMK	1	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	Tinggi	
Perempuan	15-25	SMA/SMK	5	4	3	4	4	3	4	4	4	5	4	4	4	3	3	4	3	4	3	3	Tinggi	
Perempuan	36-45	SMA/SMK	5	3	5	4	3	5	2	5	5	5	5	5	5	4	4	4	4	5	3	5	Tinggi	
Laki-Laki	15-25	SI	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	4	Tinggi	
Perempuan	15-25	SI	5	5	5	2	1	2	4	5	2	1	5	2	5	4	5	1	1	5	1	5	Sedang	
Perempuan	15-25	SMA/SMK	1	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	Tinggi	
Laki-Laki	15-25	SMA/SMK	5	5	5	3	3	3	5	5	5	5	5	5	5	5	5	5	5	5	5	5	Tinggi	
Perempuan	15-25	SMA/SMK	5	5	4	3	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	4	Tinggi	
Laki-Laki	36-45	SMA/SMK	5	4	3	4	2	2	4	3	4	4	4	4	4	4	4	4	4	3	2	4	Sedang	
Perempuan	15-25	SMA/SMK	4	4	5	4	3	3	4	5	5	3	5	4	4	4	5	4	5	4	5	4	Tinggi	
Perempuan	15-25	SMA/SMK	5	5	5	4	3	3	2	4	4	2	4	3	4	3	4	3	4	2	4	4	Sedang	
Perempuan	15-25	SMA/SMK	4	3	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	Sedang	
Perempuan	15-25	SMA/SMK	4	3	5	2	5	5	3	4	4	3	3	3	5	4	3	4	4	4	4	5	Tinggi	
Laki-Laki	26-35	SMA/SMK	5	4	5	4	3	3	3	5	5	5	5	5	4	3	4	4	4	5	3	5	Tinggi	
Laki-Laki	15-25	SMA/SMK	5	5	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	Tinggi	
Laki-Laki	15-25	SMA/SMK	5	5	5	4	2	3	4	4	4	2	5	4	4	4	4	3	5	4	2	4	Tinggi	

Sumber:Hasil Penelitian (2025)

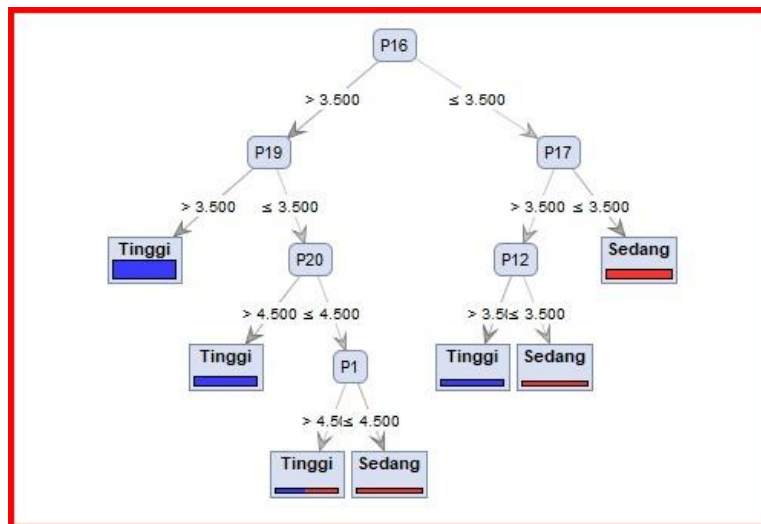
Gambar 2. Data Mentah Kuesioner (Excel)



Sumber:Hasil Penelitian (2025)

Gambar 3. Tampilan Workspace RapidMiner dan Proses C4.5

Gambar tersebut menampilkan proses klasifikasi di RapidMiner menggunakan algoritma Decision Tree dengan metode Split Validation, di mana data dibagi menjadi data latih (70%) dan data uji (30%), kemudian model dilatih dan dievaluasi untuk mengetahui kinerja klasifikasinya.



Sumber:Hasil Penelitian (2025)

Gambar 4. Hasil Pohon Keputusan

Berdasarkan model pohon keputusan yang dihasilkan, atribut P16 merupakan faktor yang paling dominan dalam membedakan tingkat kesadaran pengguna e-commerce terhadap keamanan dan privasi data pribadi. Selanjutnya, struktur percabangan mengarah pada atribut P19, P17, P20, P12, dan P1 yang berperan dalam menentukan kategori akhir tingkat kesadaran pengguna.

Secara umum, hasil klasifikasi menunjukkan:

1. Jika nilai P16 > 3.500 , maka kecenderungan kesadaran pengguna berada pada kategori Tinggi, terutama ketika didukung oleh tingginya nilai variabel P19 dan P20.
2. Jika nilai P16 ≤ 3.500 , maka kategori kesadaran lebih sering jatuh pada Sedang, terutama jika didukung oleh nilai variabel P17 dan P12 yang juga berada pada rentang menengah.

Dengan demikian, P16 (misalnya persepsi risiko atau pengetahuan privasi) menjadi indikator utama dalam menilai tingkat kesadaran pengguna.

3.2. Evaluasi Model

Evaluasi dilakukan menggunakan *confusion matrix* sehingga diperoleh hasil berikut:

Tabel 2. *Confusion Matrix*

	True Tinggi	True Sedang	Precision
Prediksi Tinggi	11	4	73.33%
Prediksi Sedang	1	4	80.00%

Sumber: Hasil Penelitian (2025)

1. *Accuracy* / Akurasi Model : 75%
2. *Recall* kelas Tinggi : 91.67%
3. *Recall* kelas Sedang : 50%

Interpretasi Akurasi

Nilai akurasi 75% menunjukkan bahwa model yang dibangun memiliki kemampuan klasifikasi yang baik dalam memprediksi tingkat kesadaran pengguna. Dengan tingkat *recall* pada kelas Tinggi yang mencapai 91.67%, model ini cukup baik dalam mendeteksi pengguna dengan kesadaran tinggi. Namun, *recall* pada kelas Sedang masih perlu ditingkatkan.

3.3. Pembahasan

Hasil penelitian menunjukkan bahwa mayoritas pengguna e-commerce memiliki tingkat kesadaran yang relatif baik terhadap pentingnya menjaga keamanan dan privasi data pribadi. Faktor utama yang memengaruhi tingkat kesadaran tersebut adalah pengetahuan pengguna mengenai risiko kebocoran data serta pemahaman mereka terhadap praktik perlindungan privasi, yang dalam penelitian ini direpresentasikan oleh atribut P16 sebagai akar pada pohon keputusan. Temuan ini sejalan dengan penelitian terdahulu yang menyatakan bahwa perilaku preventif pengguna, seperti melakukan pembaruan aplikasi dan menjaga sistem keamanan, merupakan indikator penting dalam membentuk kesadaran terhadap keamanan dan privasi data pribadi di lingkungan digital [9]. Selain itu, atribut P19, P20, P12, dan P1 menunjukkan bahwa perilaku pengguna dalam mengelola akun, termasuk penggunaan kata sandi yang kuat, kehati-hatian dalam membagikan informasi pribadi, serta pemahaman terhadap fitur keamanan pada platform e-commerce, turut berperan signifikan dalam meningkatkan tingkat kesadaran pengguna. Namun demikian, hasil klasifikasi juga mengindikasikan masih adanya sebagian pengguna yang berada pada kategori kesadaran sedang, yang umumnya disebabkan oleh rendahnya pemahaman terhadap pengaturan privasi, kurangnya kewaspadaan terhadap ancaman digital seperti phishing atau penyalahgunaan data, serta kebiasaan membagikan informasi pribadi secara terbuka. Oleh karena itu, diperlukan upaya peningkatan edukasi digital dan penguatan kampanye literasi keamanan data agar pengguna menjadi lebih proaktif dalam melindungi data pribadi saat melakukan transaksi melalui platform e-commerce.

4. Kesimpulan

Penelitian ini bertujuan untuk menganalisis tingkat kesadaran pengguna e-commerce terhadap keamanan dan privasi data pribadi serta mengidentifikasi faktor-faktor yang memengaruhinya melalui penerapan algoritma C4.5. Hasil penelitian menunjukkan bahwa tingkat kesadaran pengguna berada pada kategori tinggi dan sedang, dengan atribut P16 sebagai faktor yang paling dominan dalam menentukan tingkat kesadaran tersebut. Model klasifikasi yang

dihasilkan memperoleh tingkat akurasi sebesar 75%, yang mengindikasikan bahwa algoritma C4.5 memiliki kinerja yang cukup baik dalam mengklasifikasikan tingkat kesadaran pengguna e-commerce. Temuan ini menegaskan bahwa perilaku pengguna dalam menerapkan praktik keamanan digital, khususnya dalam melakukan pembaruan aplikasi secara berkala, memiliki peran penting dalam membentuk kesadaran terhadap keamanan dan privasi data pribadi. Sebagai rekomendasi, penelitian selanjutnya disarankan untuk memperluas jumlah responden, menambah serta mengembangkan variabel penelitian, serta melakukan perbandingan dengan algoritma klasifikasi lainnya guna meningkatkan tingkat akurasi dan kedalaman analisis.

Referensi

- [1] S. T. Zahwani and M. I. P. Nasution, "Analisis Kesadaran Masyarakat Terhadap Perlindungan Data Pribadi di Era Digital," vol. 2, no. 2, pp. 105–109, 2024.
- [2] K. Rima, A. Suari, and I. M. Sarjana, "Menjaga Privasi di Era Digital : Perlindungan Data Pribadi di Indonesia," no. 1, pp. 132–146, 2023, doi: 10.38043/jah.v6i1.4484.
- [3] P. Masyarakat *et al.*, "Jurnal dunia pendidikan," vol. 5, pp. 2507–2524, 2025.
- [4] D. E. Mahameru, A. Nurhalizah, A. Wildan, M. Haikal, and M. H. Rahmadia, "IMPLEMENTASI UU PERLINDUNGAN DATA," vol. 5, no. 2, pp. 115–131, 2023.
- [5] F. Sasi and S. Ramadhani, "Keterkaitan Pemberian Discount Pada Produk Kosmetik Dengan Konsep Predatory Pricing Dalam Perspektif Hukum Persaingan Usaha," vol. 1, no. 2, 2023.
- [6] T. W. Ramadhan, I. D. Florina, D. Permadi, I. Komunikasi, and U. P. Tegal, "Analisis Framing Pemberitaan Peretasan Pusat Data Nasional (PDN) di Media Online Tempo . co," vol. 5, no. 3, pp. 3368–3379, 2024.
- [7] J. S. Informasi and U. P. Batam, "Penerapan Algoritma C4 . 5 Untuk Memprediksi Keuntungan," vol. 01, pp. 39–46, 2019.
- [8] T. Dan, I. T. Information, and T. Pada, "Implementasi Data Mining Menggunakan Algoritma C4.5 Untuk Memprediksi Penjualan Alat – Alat Telekomunikasi Dan IT (," pp. 597–606.
- [9] J. Ilmiah and W. Pendidikan, "Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food Tifani Amalina 1 , Danendra Bima Adhi Pramana 2 , Betha Nurina Sari 3 Teknik Informatika, Universitas Singaperbangsa Karawang," vol. 8, no. September, pp. 574–583, 2022.
- [10] M. Soleh, S. R. Andani, and H. Qurniawan, "Algoritma C4 . 5 Dalam Mengklasifikasi Tindak Kejahatan Yang Sering Terjadi Di Kecamatan Siantar Barat Pematangsiantar," vol. 2, no. 5, pp. 335–340, 2021.
- [11] R. P. Aji, W. Mustofiyah, and S. A. Rizki, "Penerapan Data Mining Menggunakan Algoritma C45 Untuk Prediksi Penjualan Carica Di Dieng," vol. 11, no. 4, pp. 407–417, 2024.
- [12] D. F. Putri, I. Agama, I. Negeri, and I. Kediri, "ANALISIS PERLINDUNGAN NASABAH BSI TERHADAP KEBOCORAN DATA DALAM MENGGUNAKAN DIGITAL BANKING," vol. 1, no. 4, pp. 173–181, 2023.