

PREDICTION TINGKAT POTENSI PEMBELI PADA HARI SPECIAL DI TOKO MITRA SHOPEE

Nur Artisya¹, Saut Pintubipar Saragih²

¹Mahasiswa Program Studi Sistem Informasi, Universitas Putera Batam

²Dosen Program Studi Sistem Informasi, Universitas Putera Batam

email: pb181510050@upbatam.ac.id

ABSTRACT

Shopee is a Marketplace Platform that stepped on the Indonesian market at the peak of May 2015 and started running at the end of June 2015 in Indonesia. An item is whatever can be proposed to a market to fulfill a need or need, including actual merchandise, administrations, encounters, occasions, individuals, places, properties, associations, data and thoughts. The shopee partner of the Bagan Wholesale store has been established for 3 years with a total of more than 700 products, selling household goods, fashion, decorations etc. Implementing Data Mining, especially classification using the C4.5 algorithm in predicting buyers at Bagan Wholesale shopee partner stores. The choice tree strategy is a flowchart structure that is like a tree structure, each tree point is a quality that has been tried, each branch is an experimental outcome, and the end point is the subsequent class division. In the analysis process, the researcher uses data mining techniques using the C4.5 classification algorithm which is processed using the Rapidminer application version 9.10.001 in order to find buyer predictions at shopee partner stores.

Keywords: Prediction, C4.5 Algoritma, Rapidminer, Buyer's prediction.

PENDAHULUAN

Kemajuan teknologi yang semakin *modern* mengakibatkan kemajuan pada perbelanjaan *online*, terutama *e-commerce* yang ada di Indonesia. Cepatnya pertambahan dagang *online* telah membuat kesempatan dan tantangan bagi *situs* belanja *online*. Karena membeli barang *online* sangat mudah, menyebabkan kenaikan pola pemakaian masyarakat secara cepat. Situasi ini menyebabkan peralihan pola hidup masyarakat menjadi konsumtif.

Karena membeli barang *online* sangat mudah, menyebabkan kenaikan pola pemakaian masyarakat secara cepat. Keadaan saat ini membuat perubahan cara hidup individu rusak.

Gaya penggunaan individu sangat tunduk pada kemudahan dan kelangsungan hidup pengenalan belanja. Keringanan membeli barang *online* yang berlangsung di masyarakat sekarang, adalah dari *desktop* berubah menjadi *mobile*.

Masyarakat yang telah menikmati kemudahan saat membeli barang *online*, biasanya akan melakukan pembelian lagi dan kemungkinan yang besar untuk berbelanja. *Shopee* adalah *Platform Marketplace* yang menginjak pasar Indonesia pada puncak bulan Mei 2015 dan mulai berjalan akhir Juni 2015 di Indonesia. Sekarang, *shopee* telah di unduh sebanyak 100 juta lebih di *Google play store*. *Mitra shopee* toko Bagan Grosir sudah berdiri selama 3 tahun dengan total produk 700 lebih, menjual

barang barang kebutuhan rumah tangga, *fashion*, dekorasi dll. *Shopee* memberikan layanan kepada *seller* agar bisa berjualan dengan mudah dan pembeli dengan metode bayar yang terjamin dan kontrol logistik yang tertata.

Dengan total pengikut sudah mencapai 108,7 ribu, dan memiliki rating toko 4.8 dari 68,8 ribu penilaian dengan rating tertinggi 5.0, yang berarti 96 persen dari nilai tertinggi. Toko *Mitra Shopee* Bagan Grosir juga sudah menyanggah status *Star Plus* dan pembeli bisa menikmati *program gratis* ongkir pada toko yang diberikan oleh pihak *shopee*.

KAJIAN TEORI

2.1 Data

Data adalah kenyataan pahit tentang individu, kesempatan dan sesuatu yang penting untuk dikoordinasikan. Informasi terdiri dari titik data statistik yang ditangani ke dalam data. Dapat dianggap bahwa informasi adalah sekelompok realitas atau angka dan dapat ditangani menjadi data yang bermanfaat. Dapat disimpulkan bahwa data adalah sekelompok realitas atau angka dan dapat ditangani menjadi data berharga.

2.2 Data Mining

Data mining adalah salah satu bidang pengembangan tercepat karena tuntutan yang sangat besar untuk harga tambahan set data skala besar yang sejalan dengan pengembangan inovasi data (Wahyuni, 2018). Koleksi data mining tidak hanya mengumpulkan informasi tetapi mencakup penelitian dan perkiraan data yang ingin Anda

tunjukkan. Informasi yang dikumpulkan disimpan dalam kumpulan data dan kemudian ditangani sehingga dapat digunakan untuk pengambilan keputusan dengan mempertimbangkan informasi yang akan digunakan (Sunge, 2018).

Data Mining adalah strategi atau teknik tertentu yang melacak bentuk atau data menarik berkaitan dengan informasi yang dipilih (Harman, 2020). Data mining sendiri memiliki beberapa metode, salah satunya adalah pengelompokan. Pengaturan adalah cara paling umum untuk melacak properti serupa dalam banyak artikel dalam kumpulan data dan mencirikannya ke dalam berbagai kelas sesuai model pesanan yang dicirikan (Harapan & Rismayanti, 2018).

2.3 Algoritma C4.5

Algoritma merupakan sistem kerja komputer yang digunakan untuk menyusun langkah yang tepat untuk digunakan untuk mengatasi masalah (Cholil et al., 2020). Algoritma C4.5 untuk melakukan proses klasifikasi data dengan teknik pohon keputusan. Perhitungan C4.5 adalah perluasan perhitungan ID3 dan memanfaatkan aturan pohon pilihan komparatif (Iswati, 2022). *Algoritma* C4.5 adalah kemajuan perhitungan ID3 yang dibuat oleh J. Rose Quinlan.

Sebagai aturan, *Algoritma* C4.5 untuk membangun pohon pilihan adalah sebagai berikut:

- a. pilih karakteristik sebagai akar
- b. membuat cabang untuk setiap nilai
- c. bagi kasus dalam cabang
- d. ulangi siklus untuk setiap cabang sampai semua kasus di cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari

atribut-atribut yang ada. Untuk menghitung *gain* digunakan persamaan 1.

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

Rumus 2.1 Perhitungan *gain*

Keterangan :

S = himpunan kasus

A = atribut

n = jumlah partisi atribut A

|S_i| = jumlah kasus pada partisi ke-i

|S| = jumlah kasus dalam S

Sementara itu, perhitungan nilai *entropy* dapat dilihat pada persamaan. Entropy adalah ukuran dari teori informasi yang dapat mengetahui karakteristik (Arif & Risal, 2016).

2.

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

Rumus 2.2 Perhitungan *entropy*

2.4 Decision Tree

Metode *decision tree* adalah Struktur flowchart menyerupai struktur pohon, setiap titik pohon adalah merek dagang yang dicoba, setiap cabang adalah hasil dari tes, dan titik akhir adalah pembagian kelas berikut. Pohon keputusan adalah penggambaran dasar dari prosedur pengaturan untuk berbagai kelas tak terbatas, di mana hub ke dalam dan hub akar dipisahkan oleh nama-nama sifat, tulang rusuk diberi nama kualitas yang mungkin dan ikatan daun dipisahkan dengan berbagai kelas (Eska, 2016).

Aturan dapat berhasil dilihat dengan bahasa sederhana dan mereka juga dapat disampaikan dalam desain bahasa indeks informasi seperti SQL (Structured Query Language) untuk mencari catatan dalam pengelompokan tertentu. Pohon pilihan adalah pohon yang digunakan sebagai strategi mencurigai untuk mengamati jawaban untuk masalah yang disertakan.

2.5 Prediksi

Prediksi adalah proses menilai secara metodis apa yang mungkin akan terjadi di kemudian hari mengingat lebih dari data rentang waktu yang signifikan, sehingga kesalahan (perbedaan antara sesuatu yang terjadi dan hasil perkiraan) dapat dikurangi. Ramalan penjualan adalah metode yang terlibat dengan menilai beberapa persyaratan masa depan yang mengingat kebutuhan untuk proporsi jumlah, kualitas, pengaturan keseluruhan yang diperlukan untuk memenuhi kebutuhan akan tenaga kerja dan produk (Rosita Dewi & Farouq Mauladi, 2020). Pengertian Prediksi setara dengan angka atau dugaan.

Sesuai dengan referensi kata besar bahasa Indonesia, prediksi adalah konsekuensi dari tindakan meramalkan atau mengantisipasi atau menilai nilai-nilai di kemudian hari dengan memanfaatkan informasi masa lalu. Perkiraan menunjukkan apa yang akan menimpa situasi tertentu dan berkontribusi pada persiapan dan siklus dinamis.

Peramalan (*forecasting*) adalah Teknik untuk membuat data yang dapat diverifikasi tentang keadaan sosial di masa depan berdasarkan data yang ada tentang masalah pengaturan. Dugaan memiliki tiga struktur prinsip: proyeksi, prediksi, dan perkiraan.

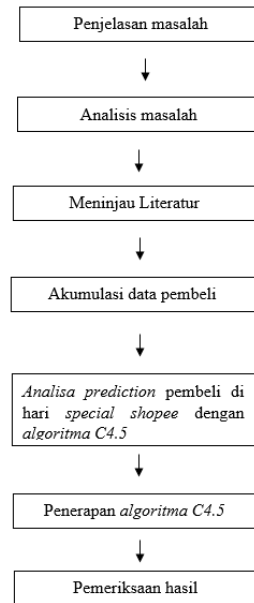
2.5 Penjualan

Penjualan adalah pekerjaan terkoordinasi untuk mendorong rencana penting yang dipandu di tempat kerja untuk memenuhi kebutuhan dan kebutuhan pembeli, untuk mendapatkan penawaran yang menciptakan manfaat. Penawaran juga dapat diuraikan sebagai upaya yang dilakukan oleh orang-orang untuk meneruskan barang dagangan yang dibutuhkan oleh pelanggan. Dengan tawaran orang dapat bertemu dari dekat dan pribadi untuk membuat, meningkatkan, ace atau mengikuti koneksi perdagangan sehingga keuntungan umum.

Pengertian penjualan lainnya, Penjualan adalah tindakan yang menunjuk ke arah melacak pembeli, mempengaruhi dan mendorong pembeli untuk mengubah kebutuhan mereka untuk penciptaan yang ditawarkan dan masuk ke penyelesaian pada biaya yang bagus. Penjualan dipahami sebagai suatu kegiatan yang secara sistematis merupakan adalah suatu rangkaian usaha yang secara terpadu dilakukan dalam menjalankan rencana organisasi untuk melakukan pemenuhan kebutuhan dan keinginan konsumen dalam rangka mendapatkan selisih antara harga jual dan harga beli yang disebut sebagai laba atau keuntungan (Izyuddin & Wibisono, 2020).

METODE PENELITIAN

Pada sub bab ini menjelaskan mengenai prediksi tingkat potensi pembeli pada hari *special shopee* di toko bagan grosir . berikut gambar desain penelitiannya



Gambar 1. Desain penelitian (sumber : penulis, 2021)

Menurut gambar diatas berikut pengertian dari penjelasan diatas :

1. Penjelasan masalah

Menguraikan masalah yang ada pada penelitian , tentang bagaimana prediksi pembeli. Disini penulis mencari tau masalah yang terjadi, masalah yang didapatkan adalah mitra shopee ingin memprediksi pembeli pada hari special shopee agar tidak kekurangan stok nantinya.

2. Analisis masalah

Di tahap ini peneliti menganalisis masalah yang sudah di jelaskan dan dapat memahami masalah tersebut, yaitu jika terjadi kekurangan produk dan promosi masih berlanjut dan masih banyak pembeli yang menginginkan produk maka penjual tidak bisa menjual karena kehabisan stok.

3. Meninjau Literatur

Tabel 3 Data Training (80% dari Total Data)

Total Pembeli	Bintang 5	Bintang 4	Bintang 3	Bintang 2	Bintang 1	Minat
200	86	8	3	0	0	TURUN
204	95	17	6	1	1	TURUN
217	86	19	0	0	0	NAIK
120	56	10	0	0	1	TURUN
183	105	5	2	0	1	NAIK
71	27	5	1	0	0	TURUN
94	30	6	2	0	0	NAIK
63	21	3	2	0	0	NAIK
80	38	1	1	0	0	NAIK
152	64	6	1	0	0	NAIK
296	115	10	2	0	0	NAIK
227	115	16	1	0	0	NAIK
181	90	12	4	0	0	TURUN
304	144	9	4	0	1	TURUN
319	160	17	2	1	0	NAIK
152	84	11	3	0	0	TURUN
272	131	25	2	0	2	NAIK
320	159	19	5	1	0	TURUN
436	212	24	10	0	1	TURUN
471	241	19	3	0	0	TURUN
124	60	6	4	0	0	TURUN
333	172	23	3	0	0	TURUN
292	155	14	3	1	1	TURUN
100	62	7	1	1	0	NAIK
57	33	4	3	0	0	TURUN
191	64	10	4	1	0	TURUN
417	199	22	4	3	3	TURUN
281	121	14	2	2	1	NAIK
79	39	7	0	0	0	NAIK

Tabel 4 Data Training (80% dari Total Data)

55	18	3	0	0	0	NAIK
87	37	6	0	0	0	NAIK
33	16	1	0	0	0	NAIK
125	63	10	1	0	1	NAIK
33	22	2	0	0	0	NAIK
27	14	0	1	0	0	NAIK
6	1	0	0	0	0	NAIK
77	56	2	2	0	0	NAIK
180	83	7	2	1	0	NAIK
139	76	13	1	0	0	NAIK
0	3	0	0	0	0	NAIK
138	62	15	12	0	2	TURUN
150	63	13	8	1	1	TURUN
182	71	19	7	1	0	NAIK
84	27	16	8	0	0	NAIK
58	36	1	0	0	0	NAIK
123	46	2	1	0	0	NAIK
79	34	5	0	0	0	NAIK
9	3	2	1	0	0	NAIK
100	54	10	0	0	0	NAIK
105	66	7	2	1	1	NAIK
51	44	7	0	0	0	NAIK
31	28	2	0	0	0	NAIK
68	14	3	2	0	0	NAIK
47	11	2	2	0	0	NAIK
41	28	4	3	0	0	TURUN
35	35	2	2	0	1	NAIK
49	30	6	3	0	0	TURUN
33	17	3	0	0	0	NAIK

Informasi yang disiapkan di atas adalah 80% dari informasi lengkap, menyiapkan informasi diselesaikan sembarangan dan kredit item, item absolut, kode item, dan bulan dilarang dengan alasan bahwa mereka tidak mempengaruhi saat mengawasi informasi tentang RapidMiner.

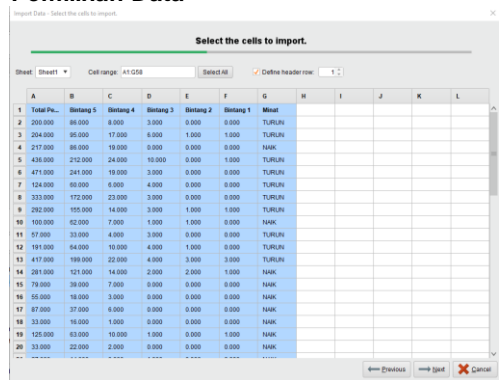
Hanya semua properti pembeli, 5 bintang, 4 bintang, 3 bintang, 2 bintang, dan 1 bintang yang dikenakan.

Tabel 5 Data Testing (20% dari Total Data)

Total Pembeli	Bintang 5	Bintang 4	Bintang 3	Bintang 2	Bintang 1	Minat
120	56	10	0	0	1	TURUN
183	105	5	2	0	1	NAIK
71	27	5	1	0	0	TURUN
94	30	6	2	0	0	NAIK
63	21	3	2	0	0	NAIK
80	38	1	1	0	0	NAIK
152	64	6	1	0	0	NAIK
296	115	10	2	0	0	NAIK
227	115	16	1	0	0	NAIK
181	90	12	4	0	0	TURUN
304	144	9	4	0	1	TURUN
319	160	17	2	1	0	NAIK
152	84	11	3	0	0	TURUN
272	131	25	2	0	2	NAIK
320	159	19	5	1	0	TURUN

Pengujian informasi di atas terdiri dari 20% dari informasi all out, informasi dipilih sembarangan, sama seperti informasi yang menyiapkan karakteristik semua item, item, kode item, dan bulan dikecualikan dengan alasan bahwa mereka tidak mempengaruhi saat mengawasi informasi di RapidMiner. Hanya semua properti pembeli, 5 bintang, 4 bintang, 3 bintang, 2 bintang, dan 1 bintang yang dikenakan.

Pemilihan Data



Gambar 1. Pemilihan Data (sumber : penulis, 2021)

Penulis mengimpor informasi yang akan diawasi ke RapidMiner. Input informasi adalah persiapan informasi atau pengujian informasi. Dalam interaksi ini tidak ada perkembangan informasi dan dapat memilih

dengan mudah informasi mana yang harus dibuat.

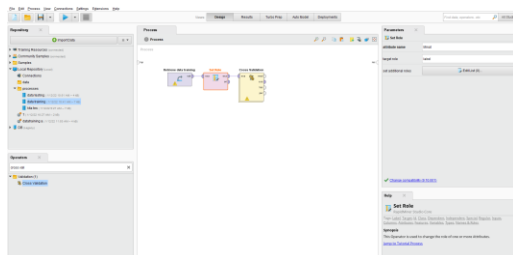
Pemilihan Atribut

Total Pembeli	Bintang 5	Bintang 4	Bintang 3	Bintang 2	Bintang 1	Minat
integer	integer	integer	integer	integer	integer	nominal Label
1 200	86	8	3	0	0	TURUN
2 204	85	17	6	1	1	TURUN
3 217	86	19	0	0	0	NAM
4 436	212	24	10	0	1	TURUN
5 471	241	19	3	0	0	TURUN
6 124	60	6	4	0	0	TURUN
7 333	172	23	3	0	0	TURUN
8 292	155	14	3	1	1	TURUN
9 100	42	7	1	1	0	NAM
10 57	33	3	0	0	0	TURUN
11 191	64	10	4	1	0	TURUN
12 417	199	22	4	3	3	TURUN
13 281	121	14	2	2	1	NAM
14 79	38	3	0	0	0	NAM
15 16	18	3	0	0	0	NAM
16 87	37	6	0	0	0	NAM
17 33	16	1	0	0	0	NAM
18 125	63	10	1	0	1	NAM
19 11	59	7	0	0	0	NAM

Gambar 2. Pemilihan Atribut
Sumber: Penulis

Dalam kualitas minat penulis merubah menjadi binominal mengingat fakta bahwa untuk kepentingan hanya ada 2 discretionary sementara untuk 1 bintang, 2 bintang, 3 bintang, bintang 4, bintang 5, menggunakan properti nomor dengan alasan bahwa informasi sebagai angka dan sifat dapat menyesuaikan diri dengan informasi yang ada, untuk kepentingan dalam jenis pekerjaan dalam substansi nama. Nama itu sendiri adalah kelas yang signifikan, tanda harus diberikan kepada 1 pekerjaan informasi tidak sama dengan kelas yang berbeda yang dapat digunakan dalam banyak bagian informasi.

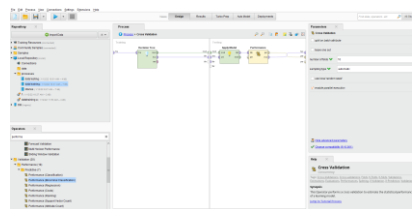
Cross Validation



Gambar 3 Cross Validation
(sumber : penulis, 2021)

Di sini pencipta menambahkan administrator operator yang ditetapkan yang terkait dengan informasi yang mendominasi sehingga program dapat merasakan pekerjaan dalam informasi, dan memerlukan administrator persetujuan silang sehingga informasi dapat dicoba dan dapat mengukur eksekusi informasi.

Main Procs Decision Tree

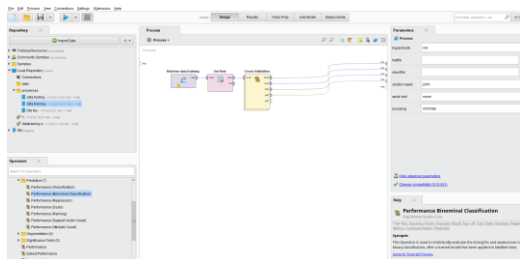


Gambar 4. Main Procs Decision Tree (Data Training)

(sumber : penulis, 2021)

Pada tahap ini penulis memasukkan operator decision tree agar dapat di klasifikasi dan juga dapat menentukan data dengan membuat pohon keputusan. Penulis juga menambahkan operator performance yang ada pada main proses untuk mendapatkan seberapa baik hasil prediction data nantinya, dan operator apply model juga dibutuhkan untuk memprediksi atau mengubah data yang tidak terlihat dan harus kompatibel dengan atribut model.

Main Procs Validation

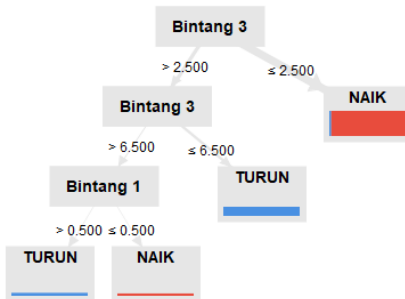


Gambar 5. Main Procs Validation
(sumber : penulis, 2021)

Penulis menghubungkan link Input ke Read Excell satu link, Baca Excell ke Validasi satu link dan Validasi ke Res empat link, Untuk melihat hasil informasi, misalnya, pohon pilihan, eksekusi, model dan sebagainya berakhir dengan informasi bermain dan hasil informasi dapat dilihat.

Pembahasan

Decision Tree



Gambar 6. *Decision Tree* (sumber : penulis, 2021)

Pengaruh terbesar dari keputusan pembeli adalah bintang 3 karena bintang 5 dan 4 sudah di pastikan naik, jika bintang 3 kurang dari 2.5 maka pembeli naik tapi jika besar dari 2.5 dan kurang dari 6.5 maka pembeli turun. Bila lebih besar dari 6.5 dan bintang 1 pada produk tersebut kurang dari 0.5 maka naik dan sebaliknya lebih besar dari 0.5 turun.

Confusion Matrix

Confusion matrix adalah tabel yang mengungkapkan ukuran yang tepat dari informasi tes yang ditandai dan berapa banyak informasi tes yang tidak akurat diatur untuk melihat harga ketepatan berikutnya. (Handayani et al., 2021).

Accuracy (Data Training)

Informasi ketepatan dari menyiapkan informasi mencapai 91,33%. Seperti yang ditunjukkan oleh tabel atas ekspektasi naik dengan berlaku oleh 38 pembeli dan berlaku di suatu tempat di dekat 3 pembeli, dan memiliki akurasi harga hasil 92,68%. Untuk ekspektasi pembeli barang turun untuk berlaku oleh 2 pembeli dan untuk valid down ada 14 pembeli yang memiliki akurasi harga hasil 87,50%. Dengan ulasan informasi pembeli yang sangat tinggi di kelas ulasan asli naik sebesar 95,00% dan untuk kelas ulasan asli turun tidak terlalu rendah, yaitu 82,35%.

Recall Data (Data Training)

Hasil pengukuran *recall* data yang diperoleh dari data *training* nilainya mencapai 95.00%. Dari tabel tersebut diketahui prediksi naik dengan *true* naik mencapai 38 pembeli dan *true* turun sebanyak 3 pembeli, dengan hasil pencapaian nilai presisi sebesar 92.68%. Sedangkan untuk prediksi penjualan barang turun untuk *true* naik mencapai 2 pembeli dan untuk *true* turun terdapat 14 pembeli dengan hasil pencapaian nilai presisi sebesar 87.50%. Untuk *recall* data pembeli yang sangat tinggi terdapat pada *class recall true* naik mencapai 95.00% sedangkan untuk *class recall true* turun mencapai 82.35%.

Accuracy (Data Testing)

Hasil *accuracy data* Nilai *data testing* mencapai 70,00%. Seperti yang ditunjukkan oleh tabel atas perkiraan naik dengan berlaku oleh 7 pembeli dan berlaku di suatu tempat di dekat 3 pembeli, dan memiliki akurasi harga hasil 70,00%. Untuk ekspektasi pembeli turun berlaku oleh 2 pembeli dan untuk valid down ada 3 pembeli yang memiliki akurasi harga hasil 60,00%. Dengan ulasan informasi pembeli yang benar-benar tinggi di kelas ulasan asli naik 77,78% dan untuk kelas ulasan asli turun tidak terlalu rendah, yaitu 50,00%.

Recall Data (Data Testing)

Efek samping dari estimasi tinjauan informasi yang diperoleh dari harga pengujian informasi mencapai 77,78%. Dari tabel, diketahui bahwa perkiraan naik dengan pendakian asli yang tiba di 7 pembeli dan berlaku di atas 3 pembeli, dengan konsekuensi mencapai nilai akurasi 70,00%. Mengenai harapan penawaran produk turun untuk berlaku hingga 2 pembeli dan untuk valid down ada 3 pembeli dengan efek samping mencapai nilai akurasi 60,00%. Untuk ulasan, informasi pembeli yang sangat tinggi di kelas ulasan asli naik 77,78% sementara untuk kelas ulasan asli sekitar 50,00%.

ROC/AUC (Area Under Cover)

Eksekusi presisi AUC (Gorunescu, 2010) dapat dicirikan menjadi lima pertemuan:

- 0,90 - 1,00 = Klasifikasi Yang Sangat Baik.
- 0,80 - 0,90 = Klasifikasi Baik.
- 0,70 - 0,80 = Klasifikasi Adil.
- 0,60 - 0,70 = Klasifikasi Miskin.
- 0,50 - 0,60 = Kegagalan.

ROC/AUC Data Training

Hasilnya didapat dari penanganan ROC dari strategi Algoritma C4.5 yang ditambahkan hingga 0,838 menggunakan informasi yang disiapkan di atas dengan tingkat ketepatan *Good Classification*.

ROC/AUC Data Testing

Dan untuk pengolahan ROC menggunakan data testing sebesar 0.750 dapat dilihat pada gambar diatas dengan tingkat akurasi *Fair Classification*.

Dengan akurasi mencapai 91.33% maka penulis menyimpulkan bahwa produk yang diteliti sebagai berikut:

Prediksi pembeli naik : cd pria 13, sepatu wanita 06, tas jelly 213, tas anak 226, lampu tumblr kotak, botol minum 03, lampu tumblr plastik, tas pria 198.

Prediksi pembeli turun : dompet wanita 44, sepatu sandal, tirai benang, tas anak 221, tas anak 76, bunga hias 09, tas olahraga 16, tas wanita 192, bunga hias 17.

SIMPULAN



Menurut penilaian hasil dan pendahuluan pada penelitian penambangan informasi harapan pembeli menggunakan perhitungan C4.5, spesialis dapat menyimpulkan bahwa dengan menggunakan perhitungan C4.5 maka pembeli dapat diantisipasi dengan ketepatan yang cukup tinggi, efek samping dari memperkirakan presisi informasi yang diperoleh dari menyiapkan informasi dengan nilai tambah Confusion Matrix mencapai 91,33%. selanjutnya informasi pengujian tingkat presisi mencapai 70,00%. Ketika dilihat dengan tikungan ROC, informasi persiapan memiliki presisi Klasifikasi yang Baik 0,838 dan pengujian informasi dengan ketepatan Klasifikasi Yang Adil 0,750.

DAFTAR PUSTAKA

- Arif, A., & Risal, A. (2016). Risal, Implementasi Metode Data Mining Untuk Memprediksi Pertandingan Sepak Bola 48. In *Jurnal SISTEMASI* (Vol. 5, Issue 2).
- Cholil, S. R., Dwijayanto, A. F., Ardianita, T., Studi, P., Informasi, S., & Komunikasi, D. (2020). Prediksi penyakit demam berdarah di puskesmas ngemplak simongan menggunakan algoritma c4.5. *SISTEMASI: JURNAL SISTEM INFORMASI*, 9(3), 529–542.
- effendi, M. M., & Rahmawati, D. (2018). Prediksi penjualan produk roti menggunakan algoritma c4.5 pada PT.Prima top boga. *SIGMA*

- Information Technology Journal*, 9(1).
- Eska, J. (2016). Penerapan data mining untuk prediksi penjualan wallpaper menggunakan algoritma c4.5. *JURTEKSI*, 2(2), 9–3.
- Handayani, K., Latif, A., Rifqi Firdaus, M., & Nur Hasan, F. (2021). Komparasi Algoritma C4.5 dan Naïve Bayes dalam Penentuan Status Kelayakan Donor Darah. *SISTEMASI: Jurnal Sistem Informasi*, 10(3), 676–687. <http://sistemasi.ftik.unisi.ac.id>
- Harapan, U., & Rismayanti, M.). (2018). Decision Tree Penentuan Masa Studi Mahasiswa Prodi Teknik Informatika (Studi Kasus: Fakultas Teknik dan Komputer. *Jurnal Sistem Informasi*, 1.
- Harman, R. (2020). Penerapan data mining untuk prediksi penjualan readymix menggunakan metode algoritma c4.5 pada PT Remicon widyaprima. *JURNAL COMASIE*, 3(3).
- Iswati, W. (2022). Penerapan data mining dengan algoritma c4.5 dalam prediksi penjualan pada toko deshop cikarang. *JURNAL MANDIRI IT*, 10(2).
- Rosita Dewi, K., & Farouq Mauladi, K. (2020). Analisa Algoritma C4.5 untuk Prediksi Penjualan Obat Pertanian di Toko Dewi Sri. In *Seminar Nasional Inovasi Teknologi* (Vol. 25, Issue 2020).
- Sunge, A. (2018). *Prediksi kompetensi karyawan menggunakan algoritma c4.5(studi kasus : PT hankook tire indonesia)*. <https://www.researchgate.net/publication/333853261>
- Wahyuni, S. (2018). Implementation of Data Mining to Analyze Drug Cases

Using C4.5 Decision Tree. *Journal of Physics: Conference Series*, 970(1). <https://doi.org/10.1088/1742-6596/970/1/012030>

	<p>Biodata, Penulis pertama, Nur Artisya, merupakan mahasiswa Prodi Sistem Informasi Universitas Putera Batam.</p>
	<p>Biodata, Penulis kedua, Saut Pintubipar Saragih, S.Kom., M.MSI. ,merupakan Dosen Prodi Sistem Informasi Universitas Putera Batam. Penulis banyak berkecimpung di bidang sistem informasi.</p>