

ANALISA PENJUALAN TERHADAP CAFFE WARUNG CERBON MENGGUNAKAN KLASIFIKASI *DECISION TREE*

Faozan Mubarak¹, Sari Susanti²

¹Universitas Adhirajasa Reswara Sanjaya
E-mail: mubarakfaozan@gmail.com

²Universitas Adhirajasa Reswara Sanjaya
E-mail: sarisusanti@ars.ac.id

Abstrak

Manusia sebagai makhluk hidup memiliki kebutuhan yang paling mendasar untuk bertahan hidup yaitu makan, Sehingga kian marak munculnya usaha-usaha di bidang kuliner. Oleh karena itu peluang bisnis dibidang kuliner memiliki peluang yang sangat besar, hal ini dapat kita lihat dari banyaknya café, resto maupun gerai makanan yang bermunculan. Tujuan penelitian ini ialah untuk menerapkan teknik data mining dengan menggunakan metode Decision tree algoritma C4.5 pada kasus di café dan diharapkan mampu memberikan informasi berupa klasifikasi penjualan produk apa saja yang paling sering diminati oleh konsumen dan produk yang kurang diminati. Pada penelitian ini, menggunakan data yaitu data penjualan. Data ini dapat diolah memanfaatkan lebih dari satu metode data mining supaya diperoleh metode yang mampu digunakan sebagai rules dalam memprediksi penjualan barang yang laris dan kurang laris. Setelah di lakukan penelitian untuk memprediksi penjualan produk laris dan kurang laris adalah dengan metode Algoritma C4.5 Decision Tree. Dan hasil yang didapatkan dari pengujian Pred. Laris dan Pred. Kurang Laris adalah penggunaan metode Algoritma C4.5 Decision Tree saja pada klasifikasi menghasilkan hasil nilai akurasi yang paling optimal sebesar 58,89%

Keywords: Algoritma C4.5, Decision Tree, Café, Data Mining,

Abstract

Humans as living beings have the most basic needs for survival, namely food, so that the emergence of businesses in the culinary field is increasingly prevalent. Therefore, business opportunities in the culinary field have a very big opportunity, we can see this from the many cafes, restaurants and food outlets that have sprung up. The purpose of this study is to apply data mining techniques using the C4.5 algorithm Decision Tree method in cases in cafes and is expected to be able to provide information in the form of a classification of what products are most often in demand by consumers and products that are less desirable. In this study, using data, namely sales data. This data can be processed using more than one data mining method in order to obtain a method that can be used as a rule in predicting sales of in-demand and less-selling goods. After doing research to predict sales of best-selling and less-selling products is the C4.5 Decision Tree Algorithm method. And the results obtained from Pred testing. Laris and Pred. Less well-selling is the use of the C4.5 Decision Tree Algorithm method alone in the classification yields the most optimal accuracy value of 58.89%.

Keywords: C4.5 Algorithm, Decision Tree, Café, Data Mining.

1. Pendahuluan

Manusia sebagai makhluk hidup memiliki kebutuhan yang paling mendasar untuk bertahan hidup yaitu makan, Sehingga kian marak munculnya usaha-usaha di bidang kuliner. Dewasa ini kebutuhan kita untuk makan bukan lagi sekedar kebutuhan bertahan hidup saja namun sudah menjadi gaya hidup. Oleh karena itu peluang bisnis dibidang kuliner memiliki peluang yang sangat besar, hal ini dapat kita lihat dari banyaknya café, resto maupun gerai makanan yang bermunculan. Bahkan gaya hidup di zaman sekarang yang hampir setiap aktifitas kita bagikan ke sosial media, karena itu minat konsumen untuk datang berkunjung ke gerai, café, maupun resto bukan hanya tertarik akan menu atau rasa makanan melainkan konsep bangunan dan juga dekorasi yang “instagramable”. Dengan kemajuan teknologi dan informasi saat ini, sangat di perlukan dalam berbisnis baik di bidang kuliner atau di bidang apapun. salah satu teknologi yang dapat kita gunakan adalah Data Mining. (Fardani, 2013)

Data Mining adalah suatu proses data dimana dapat dicari dengan pola atau informasi berharga menggunakan metode-metode tertentu. Salah satu metode yang dipakai ialah metode Decision tree dengan menggunakan algoritma C4.5. (Lukhayu Pritalia, 2018) Dalam usaha penjualan khususnya dibidang kuliner pastinya mempunyai data-data penjualan yang begitu banyak sehingga agar dapat memantau data penjualan diperlukannya sebuah metode untuk mengklasifikasi penjualan produk apa saja yang paling digemari pelanggan dan yang sedikit digemari. Metodologi yang digunakan dalam mengklasifikasi penjualan barang ini menggunakan Decision Tree untuk mengetahui apa saja yang paling sering dipesan oleh konsumen. Dengan adanya data mining diharapkan mampu untuk mengolah data-data yang banyak menjadi informasi yang bermanfaat sehingga dapat diolah menjadi data atau informasi yang dibutuhkan dengan demikian bisa menggali informasi yang berguna dan mampu mengambil kesimpulan dari permasalahan yang ada pada masalah penggalian data menjadi sebuah informasi agar dapat membantu atau mempermudah menu apa saja yang paling sering di beli oleh konsumen.

Proses transaksi penjualan pada Caffe ini memiliki banyak data-data transaksi

yang begitu banyak sehingga menyulitkan untuk mendapatkan informasi data produk apa saja yang paling laku dan produk apa saja yang kurang diminati. Tujuan penelitian ini ialah untuk menerapkan teknik data mining dengan menggunakan metode Decision tree algoritma C4.5 pada kasus di caffe dan diharapkan mampu memberikan informasi berupa klasifikasi penjualan produk apa saja yang paling sering diminati oleh konsumen dan produk yang kurang diminati. (Lukhayu Pritalia, 2018) Sehingga kedepannya pemilik bisnis ini dapat melakukan analisa produk apa saja yang paling sering dipesan oleh konsumen.

2. Metode Penelitian

Pada penelitian ini, menggunakan data yaitu data penjualan. Data ini dapat diolah memanfaatkan lebih dari satu metode data mining supaya diperoleh metode yang mampu digunakan sebagai rules dalam memprediksi penjualan barang yang laris dan tidak cukup laris. Dalam penelitian ini dapat dilaksanakan lebih dari satu tahapan penelitian seperti yang nampak terhadap gambar di bawah ini : **Tabel 2.1**

Pengujian Data

JUMLAH DATA	610	100%
DATA TRAINING	488	80%
DATA TESTING	122	20%

Data yang telah disiapkan di bagi menjadi dua untuk proses klasifikasi seperti pada **Tabel 2.1** data Training berjumlah (80%) dan data Testing berjumlah (20%). (Cynthia & Ismanto, 2018) Dalam pembagian data ini menggunakan tools Split Validation untuk menghasilkan data training dan data testing. Split Validation merupakan teknik validasi untuk membagi data menjadi dua bagian secara acak, sebagian menjadi data training dan sebagian lainnya menjadi data testing. Dengan menggunakan split validation maka dilakukan percobaan training berdasarkan split rasio yang sudah ditentukan sebelumnya, untuk kemudian sisa dari split rasio data training akan dianggap sebagai data testing.

Data training adalah data yang akan digunakan dalam proses pembelajaran sedangkan data testing adalah data yang belum pernah digunakan sebagai pembelajaran dan akan berguna menjadi

data dalam pengujian kebenaran atau tingkat akurasi hasil pembelajaran.(Rohman et al., 2017)

3. Hasil dan Pembahasan

Hasil pengujian penelitian untuk mengetahui hasil dari suatu proses penelitian. Berikut adalah beberapa hasil pengujian yang telah dilakukan :

3.1. Pengujian Decision Tree

Tahapan awal yang dilakukan dalam penelitian ini adalah menyiapkan dataset penjualan produk yang didapatkan dari sumber yaitu Café Warung Cerbon. Data yang diambil adalah dataset penjualan produk yang merupakan data yang berjumlah 610 data dari 30 hari.

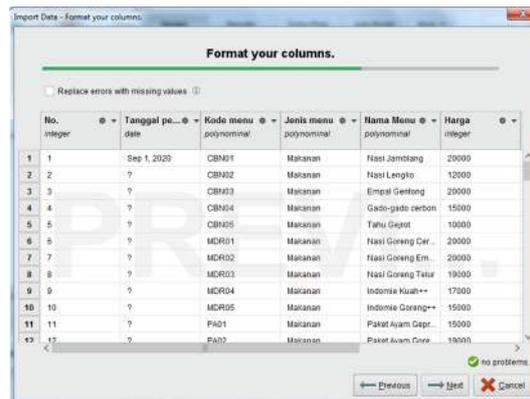
Tabel 3.1 Data Penjualan

No.	Tanggal penjualan	Kode menu	Jenis menu	Nama Menu	Harga	Jumlah Terjual	Kategori
1	9/1/2020	CBN01	Makanan	Nasi Jambalang	20000	13	Laris
2		CBN02	Makanan	Nasi Lengko	12000	14	Kurang Laris
3		CBN03	Makanan	Empal Gentong	20000	23	Laris
4		CBN04	Makanan	Gado-gado cerbon	15000	9	Laris
5		CBN05	Makanan	Tahu Gejrot	10000	17	Kurang Laris
6		MDR01	Makanan	Nasi Goreng Cerbon	20000	7	Kurang Laris
7		MDR02	Makanan	Nasi Goreng Empal	20000	12	Kurang Laris
8		MDR03	Makanan	Nasi Goreng Telur	19000	12	Kurang Laris
9		MDR04	Makanan	Indomie Kuah++	17000	12	Laris
10		MDR05	Makanan	Indomie Goreng++	15000	20	Kurang Laris
11		PA01	Makanan	Paket Ayam Geprek	15000	15	Kurang Laris
12		PA02	Makanan	Paket Ayam Goreng	19000	17	Laris
13		CMLO1	Camilan	Bala-Bala Cerbon	5000	20	Kurang Laris
14		CMLO2	Camilan	Rujak Cireng	5000	14	Kurang Laris
15		CMLO3	Camilan	Sosis Bakar	10000	21	Laris
16		CMLO4	Camilan	Tahu Lada Garam	10000	16	Laris
17		CMLO6	Camilan	Roti Bakar	8000	23	Laris
18		MNMO1	Minuman	Kopi Tubruk	5000	25	Laris
19		MNMO2	Minuman	Kopi Gula Aren	5000	24	Kurang Laris
20		MNMO3	Minuman	Es Teh Manis	5000	15	Kurang Laris
21		MNMO4	Minuman	Es Teh Tarik	10000	12	Kurang Laris
22		MNMO5	Minuman	Aneka Jus	12000	10	Laris
23		MNMO6	Minuman	Milkshake	15000	9	Laris

Data dari Tabel 3.1 Data penjualan tersebut lalu diklasifikasikan menjadi 7 atribut, yaitu di klasifikasi berdasarkan Tanggal Penjualan, Kode Menu, Jenis Menu, Nama Menu, Harga, Jumlah Terjual, Kategori. Pengujian ini menggunakan operator yang terdapat pada aplikasi RapidMiner. Setelah data dianggap sudah cukup maka langkah selanjutnya yaitu pengujian pada aplikasi RapidMiner. Pengujian ini menggunakan Aplikasi

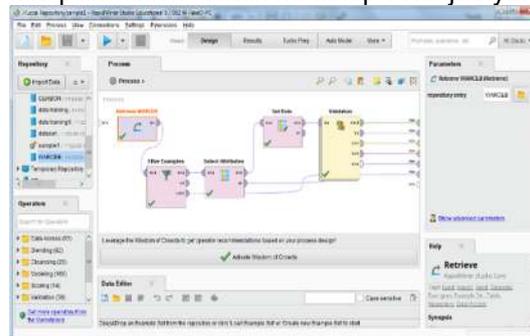
RapidMiner 9.7.(Shiddiq et al., 2018) Adapun operators yang digunakan yaitu sebagai berikut :

Pertama peneliti menggunakan operators Retrieve yang digunakan untuk melakukan import dataset dengan cara pilih Import data dari excel setelah itu pilih data yang akan diolah dan pilih atribut yang akan melalui tahapan proses pengolahan data, lalu akan tampil hasil seperti berikut :



Gambar 3.1 Import Dataset.

Pada gambar tersebut diatas menampilkan isi dari dataset yang akan diolah. Pilih "Next" lalu pilih "finish" untuk ke tahap selanjutnya.



Gambar 3.2 Tampilan Operator yang Digunakan.

Gambar 3.2 Tampilan Operator yang Digunakan.

Penggunaan Tool Examples digunakan untuk melengkapi data, atau data atribut yang hilang (data yang tidak terisi), kemudian pilih atribut bahwa data tidak dimuat dan kemudian "OK". Oleh karena itu, jika menggunakan data alat ini, itu akan diisi dengan nilai rata-rata secara otomatis pada RapidMiner. Penggunaan operator ini juga, jika data yang dimasukkan adalah kesalahan atau terjadi Human error terjadi ketika data awal sebelumnya diproses untuk diproses dengan aplikasi RapidMiner.

Set Role menentukan parameter atribut akan dijadikan hasil dari bentuk pohon keputusan. Setelah itu pilih "Set Role", maka akan terlihat tampilan dipojok atas kanan, pilih "Atribut Name". Atribut yang dipilih adalah atribut "Keterangan", lalu pilih "Target Role" dan di ubah menjadi "Label".

Split Validation adalah proses validasi yang memisahkan data menjadi dua bagian secara random, sebagian menjadi data training dan sebagian lainnya menjadi data testing. Dengan menggunakan Split Validation akan dilakukan percobaan training berdasarkan split ratio yang sudah ditentukan sebelumnya, untuk selanjutnya sisa dari split ratio data training akan dianggap sebagai data testing. Namun jika di klik 2 kali pada operator split validation maka akan memiliki dua bagian yaitu Training dan Testing, atau bisa dilihat seperti gambar berikut ini :



Gambar 3. 3 Split Validation.

Operator Split Validation ini berfungsi sebagai pembagi data menjadi 2 bagian berdasarkan split ratio yang sudah ditentukan sebelumnya, split ratio yang sudah dipilih lalu digunakan sebagai data training dan sisanya akan digunakan sebagai data testing. Pada parameter Split Validation pilih split ratio 0.8 karena peneliti ingin data yang ada dibagi menjadi 80% untuk data training dan 20% untuk data testing.

Dapat dilihat pada gambar Validation bagian Training digunakan untuk algoritma klasifikasi Decision Tree dan pada bagian Testing memakai operator Apply Model untuk mengaplikasikan model pada data testing dan operator Performance yang digunakan untuk mengetahui tampilan accuracy.

Langkah Selanjutnya yaitu menghubungkan seluruh operator yang akan digunakan atau bisa dilihat pada Gambar 3.2 dan Gambar 3.3. Dari proses yang sudah dijalankan dalam penjelasan berikut maka didapatkan hasil sebagai berikut:



Gambar 4. 4 Decision Tree (Pohon Keputusan).

Dari Gambar 3.4 maka dapat dilihat bentuk dari tampilan hasil dari pohon keputusan atau dapat dilihat penjelasannya sebagai berikut :
Tree

Tanggal penjualan = ?

| Jumlah Terjual > 0.500: Laris {Laris=227, Kurang Laris=153}

| Jumlah Terjual ≤ 0.500

| | Harga > 18000

| | | Harga > 19500

| | | | No. > 374.500: Kurang Laris {Laris=0, Kurang Laris=2}

| | | | No. ≤ 374.500: Laris {Laris=3, Kurang Laris=1}

| | | Harga ≤ 19500: Kurang Laris {Laris=1, Kurang Laris=6}

| | Harga ≤ 18000: Laris {Laris=19, Kurang Laris=19}

Tanggal penjualan > Sep 27, 2020: Laris {Laris=3, Kurang Laris=0}

Tanggal penjualan ≤ Sep 27, 2020

| No. > 449.500: Kurang Laris {Laris=0, Kurang Laris=5}

| No. ≤ 449.500

| | Jumlah Terjual > 0.500: Kurang Laris {Laris=4, Kurang Laris=6}

| | Jumlah Terjual ≤ 0.500: Laris {Laris=3, Kurang Laris=0}

Deskripsi diatas menunjukkan aturan-aturan yang berlaku untuk digunakan sebagai landasan memprediksi data penjualan barang laris dan kurang laris. Berikut adalah hasil accuracy, recall, precision.

Actual \ Predicted	Laris	Kurang Laris	Total
Laris	227	153	380
Kurang Laris	153	227	380
Total	380	380	760

Gambar 3. 5 Hasil Accuracy Decision Tree.

Dari Gambar 3.5 maka dapat dilihat nilai accuracy yang didapatkan dari pengujian algoritma Decision Tree C4.5 yaitu sebesar 58,89%. Dan nilai hasil dari recall, precision dapat dilihat pada gambar di atas.

3.2. Keterbatasan Penelitian

Berdasarkan pada pengalaman peneliti dalam melakukan penelitian ini ada beberapa keterbatasan yang dialami dan dapat menjadi beberapa faktor yang agar dapat untuk lebih diperhatikan bagi peneliti-peneliti selanjutnya yang akan datang agar lebih menyempurnakan penelitiannya. Keterbatasan penelitian dan kekurangan penelitian ini antara lain :

1. Keterbatasan waktu dalam pelaksanaan penelitian sehingga di anggap kurang maksimal dalam pengerjaan penelitian.
2. Dalam proses pengambilan data, informasi yang diberikan pada responden melalui kuisioner terkadang tidak menunjukkan pendapat responden yang sebenarnya, sehingga peneliti pernah mendapati responden yang tidak kooperatif. Hal ini terjadi karna perbedaan pemahaman, pemikiran, tanggapan yang berbeda pada setiap responden.

4. Kesimpulan

Dari proses penelitian yang sudah dilakukan sebelumnya maka dapat diambil sebuah kesimpulan :
Pengolahan data yang sudah dilakukan pada penelitian ini adalah Algoritma C4.5 Decision Tree. Dan hasil yang didapatkan dari pengujian Pred. Laris dan Pred. Kurang Laris adalah penggunaan metode Algoritma

C4.5 Decision Tree saja pada klasifikasi menghasilkan hasil nilai akurasi yang paling optimal sebesar 58,89%.

Dari hasil penghitungan menggunakan RapidMiner 9.7 dari data 690 total keseluruhan penjualan pred Laris adalah 366, dan pred Kurang Laris adalah 271.

Referensi

- Cynthia, E. P., & Ismanto, E. (2018). Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji. *Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika)*, 3(July), 1. <https://doi.org/10.30645/jurasik.v3i0.60>
- Fardani, A. N. (2013). Analisis Pengaruh Kualitas Pelayanan, Harga, Dan Promosi Terhadap Kepuasan Pelanggan Dyriana Bakery & Caf  Pandanaran Semarang. *Skripsi*, 1–75.
- Lukhayu Pritalia, G. (2018). Penerapan Algoritma C4.5 untuk Penentuan Ketersediaan Barang E-commerce. *Indonesian Journal of Information Systems*, 1(1), 47–56. <https://doi.org/10.24002/ijis.v1i1.1727>
- Rohman, A., Suhartono, V., & Supriyanto, C. (2017). Penerapan algoritma c4.5 berbasis. 13(1), 13–19. <http://research.pps.dinus.ac.id/index.php/Cyberku/article/view/4/4>
- Shiddiq, A., Niswatin, R. K., & Farida, I. N. (2018). Ahmad Shiddiq Analisa Kepuasan Konsumen Menggunakan Klasifikasi Decision Tree Di Restoran Dapur Solo (Cabang Kediri). *Generation Journal*, 2(1), 9. <https://doi.org/10.29407/gj.v2i1.12051>