



KLASIFIKASI PRODUK ROKOK PADA PT WICAKSANA OVERSEAS INTERNASIONAL KOTA LUBUKLINGGAU DENGAN METODE *DECISION TREE*

Ice Pransiska¹, Joni Karman², Hardi Mulyono³, Nelly Khairani Daulay⁴

^{1,2}Program Studi Sistem Informasi, ³Program Studi Manajemen, ⁴Program Rekayasa Sistem
Komputer, Universitas Bina Insan, Lubuklinggau

e-mail: ¹icefransiska53@gmail.com, ²joni_karman@univbinainsan.ac.id,

³hardimulyono@univbinainsan.ac.id, ⁴nellydaulay@univbinainsan.ac.id

*Corresponding Author

Abstrak

Dilihat dari beragamnya produk rokok dan juga banyaknya permintaan konsumen yang berbeda-beda, maka dibutuhkan klasifikasi produk rokok pada PT wicaksana Overseas Internasional yang bertujuan untuk mengklasifikasi jenis rokok yang termasuk kategori laris dan kurang laris dengan harapan dapat mempermudah bagian penyedia stok barang pada perusahaan. Proses klasifikasi pada penelitian ini menggunakan metode *Decision tree*, Keunggulan dari metode *Decision Tree* ini yaitu jika daerah pengambilan keputusan yang sebelumnya kompleks dan sangat global, dapat diubah menjadi lebih simpel, Sampel yang digunakan dalam penelitian ini adalah data yang diolah menggunakan data penjualan pada tahun 2020 (januari- desember) dan data tahun 2021 (januari - Agustus), dengan hasil klasifikasi produk rokok Laris dan Kurang. Uji data menggunakan *tools rapidminer* dengan performa akurasi 81.48% dan klasifikasi eror 18.52% dari performa akurasi menggambarkan model klasifikasi ini memiliki performa yang cukup baik untuk digunakan pada proses pengklasifikasian.

Kata kunci— Klasifikasi; Decision Tree; Rapidminer

Abstract

Judging from the variety of cigarette products and also the many different consumer demands, it requires the classification of cigarette products at PT Wicaksana Overseas International which aims to classify the types of cigarettes which are included in the best-selling category and are not in demand in the hope of making it easier for the stock provider of goods in the company. The classification process in this study uses the Decision Tree method, the advantage of this Decision Tree method is that if the previously complex and very global decision -making area, can be changed to be more simple, the sample used in this study is the data processed using sales data in 2020 (January -December) and data in 2021 (January - August), with the results of the classification of best -selling cigarette products and less. The data test uses a rapidminer tool with an accuracy performance of 81.48% and an error classification of 18.52% of the accuracy performance illustrates this classification model has a good enough performance for use in the classification process.

Keywords : Classification; Decision Tree; Rapidminer



I. PENDAHULUAN

Perusahaan-perusahaan terkemuka saat ini tidak lepas dari kemajuan teknologi terutama internet. Terapan teknologi yang digunakan juga berkaitan dengan sistem yang canggih. Perusahaan terkemuka tentunya sudah memiliki sebuah sistem yang baik. Salah satunya adalah sistem yang digunakan dalam pengolahan untuk menghasilkan informasi. Ada banyak sistem pengolahan data yang dapat menghasilkan informasi salah satunya adalah penerapan *data mining*[1].

Klasifikasi adalah sebuah proses untuk menemukan model yang membedakan konsep dan kelas data, yang bertujuan untuk dapat memperkirakan kelas dari suatu objek yang kelasnya tidak diketahui. Salah satu metode yang paling populer adalah Decision tree karena mudah untuk diinterpretasi oleh manusia. Model klasifikasi ini terdiri dari dua yaitu deskripsi dan prediksi[2].

Salah satu kegunaan dari klasifikasi adalah untuk membantu pemilik perusahaan dalam perencanaan penyediaan stok, karena klasifikasi ini dapat memberikan *output* terbaik, sehingga perusahaan dapat dengan mudah mengatasi persediaan stok, klasifikasi digunakan untuk menemukan informasi dari sejumlah data yang besar.

PT. Wicaksana Overseas Internasional merupakan perusahaan multinasional. Salah satu cabang perusahaan ini beralamat di Sumatera Selatan, Kota Lubuklinggau, Kec. Lubuklinggau Selatan II, Kel. Simpang Periuk, RT.04 yang merupakan distributor Produk Rokok dengan mengelola penjualan produk Rokok yang terdiri dari 9 macam, beberapa diantaranya :*Dunhill Filter 12, Dunhill Filter 16, Dunhill Mild 20, Dunhill Mild 16, Lucky Strike Filter 20, Lucky Strike Mild 16, Bentoel Sejati 12, Comodore 20, Sky mild 16*.

Dengan pengelolaan data yang bersifat manual, membuat stok pada gudang sulit untuk diklasifikasikan, mana barang yang laris dan kurang laris sehingga sering terjadi penumpukan produk pada gudang. Dilihat dari banyaknya macam produk rokok dan banyaknya permintaan konsumen

yang berbeda-beda dengan permasalahan yang terjadi, maka dibutuhkan klasifikasi produk rokok pada PT wicaksana Overseas Internasional yang bertujuan untuk mengklasifikasi jenis rokok yang termasuk kategori laris dan kurang laris.

Tujuan dibuat klasifikasi ini untuk dapat mempermudah bagian penyedia stok barang pada perusahaan, agar dikemudian hari tidak lagi mengalami kesulitan dalam penyediaan stok barang. Sehingga stok barang yang tersedia dan barang keluar dapat di tangani dengan baik dan tidak terjadi penumpukan. Pada penelitian sebelumnya yang dilakukan oleh Budanis Dwi Meilani dkk pada tahun 2012, dimana penelitian tersebut membahas tentang penerapan klasifikasi pada data-data karyawan dan jadwal kerja, dimana diharapkan nantinya dapat menghasilkan suatu pola-pola tertentu.

Dari pola-pola tersebut dapat dibuat jadwal kerja yang baru menggunakan *decision tree* dan algoritma yang digunakan untuk membentuk pohon keputusan adalah algoritma C4.5. *Rule* yang digunakan adalah *rule* januari 2011 karena *rule* tersebut memiliki nilai akurasi paling tinggi yaitu sebesar 87%, atribut hasil yang digunakan adalah atribut jadwal, jika atribut lain yang digunakan, maka pohon keputusan tidak akan terbentuk karena kurangnya jumlah data[3].

Kemudian Penelitian kedua yang dilakukan oleh Nurhidayati dan Alimudin pada tahun 2019 dimana tujuan penerapan klasifikasi adalah untuk mengklasifikasi jenis obat-obatan yang termasuk kategori laris dan kurang laris. Berdasarkan hasil pengujian yang telah dilakukan dengan menerapkan 10 *cross validation* didapatkan hasil akurasi tertinggi yaitu pada *K-Fold validation* 8 dengan nilai 97,43% dan *Area Under Curve* yaitu 0,985[4].

Penelitian ketiga yang dilakukan oleh Iwan Santosa dan kawan-kawan, pada tahun 2018 menjelaskan tentang penyakit *tuberculosis* dimana penyakit ini merupakan penyakit menular yang menyerang saluran pernapasan pada manusia yang disebabkan oleh bakteri *mycobacterium*. Semakin cepat seseorang sadar terdiagnosa penyakit *Tuberculosis*



maka proses penyembuhan dimungkinkan akan semakin cepat.

Salah satu cara untuk membantu mendeteksi penyakit tersebut adalah dengan menggunakan *data mining*. Setelah dilakukan pengujian dengan 100 data pasien diperoleh hasil dengan nilai akurasi sebesar 90% menggunakan *Confusion Matrix*[5].

II. METODOLOGI PENELITIAN

Pada penelitian ini digunakan metode *Decision Tree*. Untuk melakukan pemisahan objek (*split*) di dalam *Decision Tree* dilakukan tes terhadap atribut dengan mengukur tingkat ketidak murnian pada sebuah simpul (*node*). Pada algoritma ID3 hanya dapat digunakan untuk menyelesaikan permasalahan *classification*. Sebelum menghitung *Entropy* perolehan perlu menghitung dulu nilai informasi dalam satuan *bits* dari suatu kumpulan objek. Beberapa langkah-langkah Metode *Decision tree* secara garis besar dalam membangun sebuah pohon keputusan sebagai berikut [6]:

1. Menyusun *decision system* yang terdiri dari atribut, kondisi, lalu atribut keputusan. Dapat dilihat contoh *decision system* yang pada penelitian ini, yang terdiri dari n objek, $E1, E2, E3, E4, \dots, En$ lalu atribut kondisi penjualan, stok gudang, beban usaha, pembelian, sedangkan profit (*decision attribute*).
2. Perhitungan jumlah kolom, jumlah data menurut anggota atribut hasil dengan syarat tertentu. Untuk tahap pertama masih kosong.
3. Pilih atribut sebagai *node*.
4. Selanjutnya buat cabang untuk setiap anggota sebagai *node*.
5. Selanjutnya buat cabang untuk setiap anggota sebagai *node*.

6. Lalu periksa apakah nilai *entropy* dari bagian *node* ada yang bernilai nol jika ada, maka tentukan daun yang terbentuk, dan jika nilai *entropy* pada bagian seluruh *node* adalah nol, maka proses perhitungan pun berhenti.

7. Bila masih terdapat anggota *node* memiliki nilai nol, maka ulang lagi proses dari awal dengan *node* sebagai syarat sampai semua anggota *node* memiliki nilai nol.

Node merupakan atribut yang memiliki nilai *gain* tertinggi dari semua atribut, untuk menghitung *gain* digunakan rumus yang dapat dilihat pada gambar dibawah ini :

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

Sumber: penelitian Ice pransiska.2021

Keterangan :

S : Himpunan Kasus

A : Atribut

n : Jumlah Partisi A

S_i : Proporsi dari S_i terhadap S

S : Jumlah kasus dalam S

Sementara itu untuk menghitung nilai *Entropy* dapat dilihat pada persamaan di bawah ini :

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

Sumber: penelitian Ice pransiska.2021

Keterangan :

S : Himpunan Kasus

n : Jumlah partisi S

p_i : Proporsi S_i terhadap S

Berikut adalah alur penerapan teknik klasifikasi menggunakan metode *decision tree* :



Sumber: penelitian Ice pransiska.2021

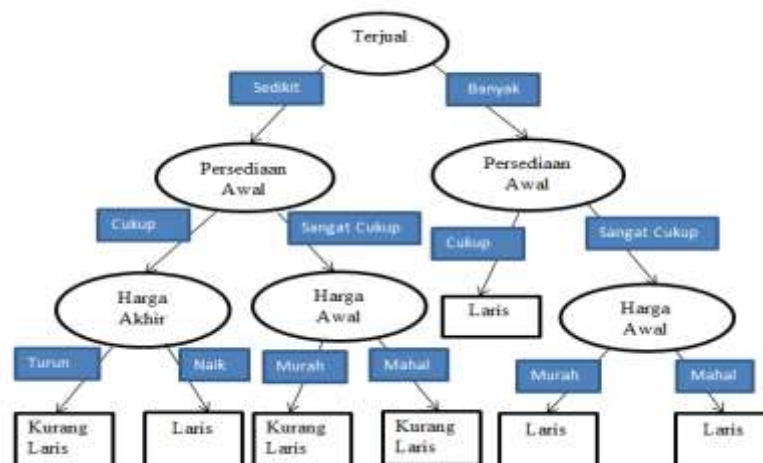
Untuk Perhitungan Entropy dan Information gain dapat dilihat pada tabel dibawah :

Atribut		jumlah	Laris	kurang laris	Entropy	Gain
Total		153	107	46	0.882088382	
Persediaan Awal	Cukup	67	50	17	0.817138776	0.005938792
	Sangat Cukup	86	57	29	0.922123131	
Harga Awal	Murah	68	43	25	0.948848294	0.012276305
	Mahal	85	64	21	0.806583102	
Harga Akhir	Turun	68	43	25	0.948848294	0.012276305
	Naik	85	64	21	0.806583102	
Terjual	Sedikit	71	34	37	0.998711751	0.131899611
	Banyak	82	73	9	0.499179836	

Sumber: penelitian Ice pransiska.2021

Untuk pohon keputusan dan Information gain dapat dilihat pada gambar dibawah :

III. HASIL DAN PEMBAHASAN



Sumber: penelitian Ice pransiska.2021

Berdasarkan pohon keputusan di atas maka dapat disimpulkan yang menjadi *rule* keputusan :

1. Jika produk Terjual = Sedikit dengan Persediaan Awal = Cukup dan Harga Akhir Turun maka produk = Kurang Laris

2. Jika produk Terjual = Sedikit dengan Persediaan Awal = Cukup dan Harga Akhir = Naik maka produk = Laris

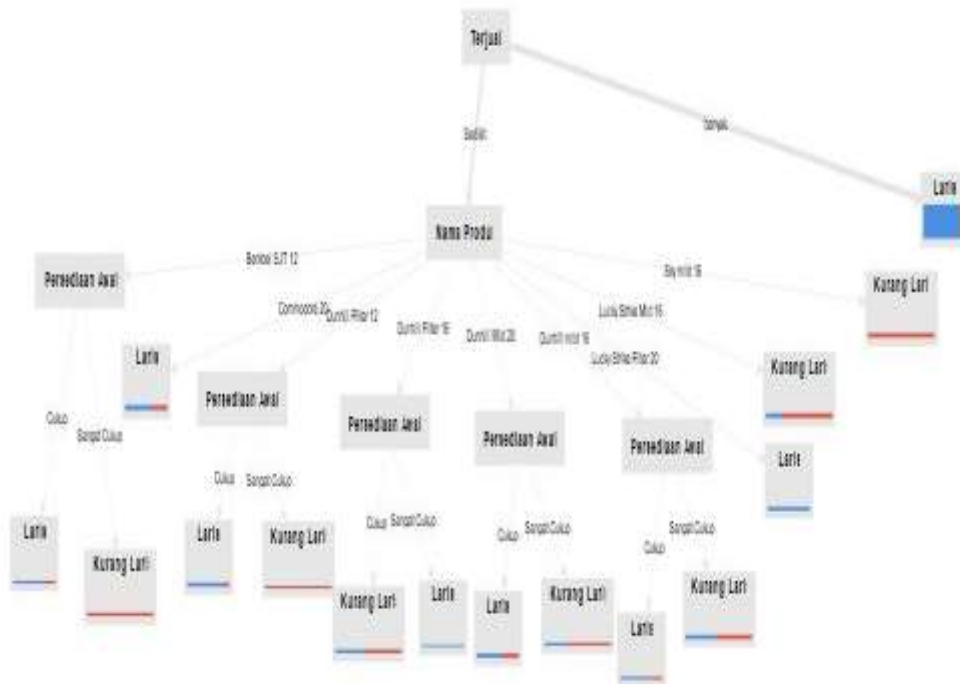
3. Jika Produk Terjual = Sedikit dengan Persediaan Awal = Sangat Cukup dan Harga Awal = Murah maka produk = Kurang Laris



4. Jika produk Terjual =Sedikit dengan Persediaan Awal =Sangat Cukup dan Harga Awal =Mahal maka produk =Kurang Laris
5. Jika produk Terjual =Banyak dengan Persediaan Awal =Cukup maka produk =Laris
6. Jika Produk Terjual = Banyak dengan Persediaan Awal =Sangat Cukup dan Harga Awal =Murah maka produk =Laris

7. Jika produk Terjual = Banyak dengan Persediaan Awal =Sangat Cukup dan Harga Awal =Mahal maka produk =Laris

Setelah melakukan perhitungan dan menghasilkan sebuah pohon keputusan berikutnya akan dilanjutkan dengan pengujian algoritma metode *decision tree* pada *Rapidminer* untuk mengetahui tingkat akurasi, apakah metode ini baik atau sesuai untuk digunakan dalam sebuah pengkalsifikasian sebuah produk seperti pada penelitian ini.

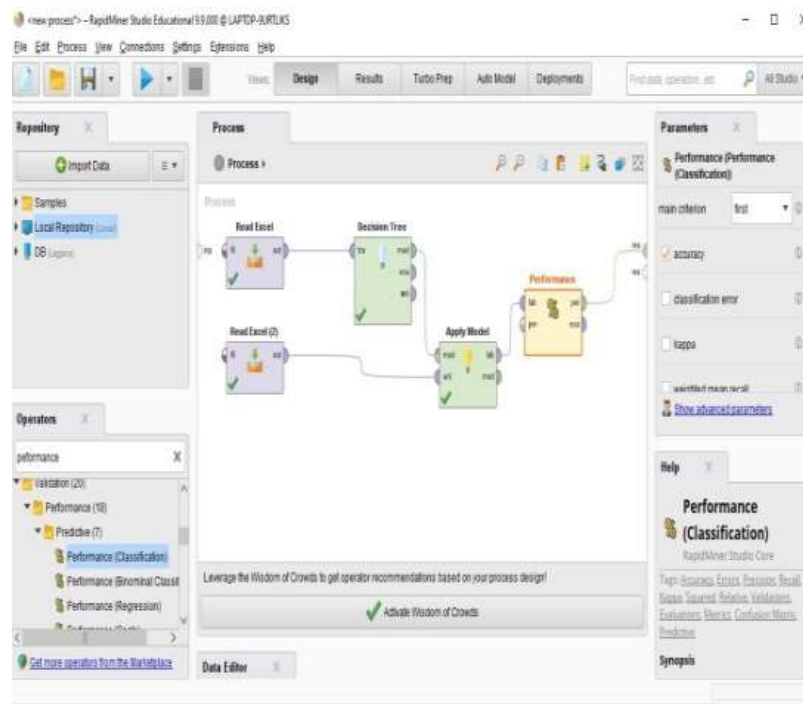


Sumber: penelitian Ice pransiska.2021



Lalu tampilan akan menunjukkan hasil model dari operator yang di input pada

halaman proses, seperti **tree Decision tree** dapat dilihat pada gambar di bawah ini



Sumber: penelitian Ice pransiska.2021

Selanjutnya untuk mengetahui performance dari pengolahan data ini maka tahap selanjutnya **klik performance pada operator** → **drag ke halaman**

proses → **sambungkan semua operator** → **klik play** dan akan menampilkan sebuah tabel dengan nilai akurasi dan klasifikasi eror beserta keterangan seperti gambar di bawah:

accuracy: 81.48%

	true Kurang Laris	true Laris	class precision
pred. Kurang Laris	6	2	75.00%
pred. Laris	3	16	84.21%
class recall	66.67%	88.89%	

Sumber: penelitian Ice pransiska.2021



IV. KESIMPULAN

Berdasarkan hasil dan pembahasan yang dilakukan, Simpulan yang diperoleh pada penelitian ini adalah : Hasil klasifikasi dalam pengolahan data ini menghasilkan Produk rokok yang dalam kategori status penjualan Laris yaitu dengan Produk rokok *Dunhill filter 12*, *Dunhill Mild 20*, *Dunhill Mild 16*, *Lucky Strike Filter 20*, *Bentoel SJT 12*, *Sky mild 16* sedangkan dengan kategori Kurang Laris yaitu dengan produk rokok *Dunhill filter 16*, *LuckyStrike Mild 16*, *Commodore 20*. Hasil uji data menggunakan tools *rapidminer* dengan performa akursi 81.48% dan eror 18.52%.

**V. DAFTAR PUSTAKA**

- [1] P. Meilina, “Penerapan Data Mining dengan Metode Klasifikasi Menggunakan Decision Tree dan Regresi,” *J. Teknol. Univ. Muhammadiyah Jakarta*, vol. 7, no. 1, pp. 11–20, 2015, [Online]. Available: jurnal.ftumj.ac.id/index.php/jurtek.
- [2] I. Budiman and R. Ramadina, “Penerapan Fungsi Data Mining Klasifikasi untuk Prediksi Masa Studi Mahasiswa Tepat Waktu pada Sistem Informasi Akademik Perguruan Tinggi,” *Ijccs*, vol. x, No.x, no. 1, pp. 1–5, 2015.
- [3] F. Dwi Meliani Achmad, Budanis, Slamet, “Klasifikasi Data Karyawan Untuk Menentukan Jadwal Kerja Menggunakan Metode Decision Tree,” *J. IPTEK*, vol. 16, no. 1, pp. 18–23, 2012, [Online]. Available: <http://jurnal.itats.ac.id/wp-content/uploads/2013/06/3.-BUDANIS-FINAL-hal-17-23.pdf>.
- [4] “PENERAPAN METODE DECISION TREE UNTUK KLASIFIKASI PENJUALAN OBAT PERTANIAN LARIS DAN KURANG LARIS PADA UD. CAHAYA TANI - Repository Hamzanwadi.” .
- [5] I. Santosa, H. Rosiyah, and E. Rahmanita, “Implementasi Algoritma Decision Tree C . 45 Untuk Diagnosa Penyakit Tuberculosis (Tb),” *J. Ilm. NERO*, vol. 3, no. 3, pp. 169–176, 2018.
- [6] s. s. M. s. Retno Tri Vlandari, *Data mining*. Yogyakarta: yogyakarta: Gava media, 2017.