

Implementasi Data Mining Metode C.45 untuk Memprediksi Peminat Kuota Internet pada Masa Covid 19

¹⁾Heliyanti Susana

Program Studi Teknik Informatika, STMIK IKMI Cirebon
email: heliyanti00@gmail.com

²⁾Shanti Khofidoh

Program Studi Teknik Informatika, STMIK IKMI Cirebon
email: shanti23@gmail.com

³⁾Azhar Al Afghani

Program Studi Teknik Informatika, STMIK IKMI Cirebon
email: azhar222@gmail.com

ABSTRACT

During the Covid 19 pandemic, Teling experienced an increase in sales and there were several internet SIM cards that consumers were less interested in. Various kinds of internet cards are sold and offered to attract consumers' attention, but it cannot be predicted which internet cards the consumers will be interested in. The C4.5 algorithm is a data classification algorithm with a decision tree technique that has advantages. These advantages, for example, can process numeric (continuous) and discrete data, can handle missing attribute values, produce easy rules. The results of the study explain that all providers including Telkomsel, Indosat, and axis are in demand by consumers. In analyzing using the C.45 algorithm with the Rapidminer Tools, the first is that the decision tree results in determining consumer interest are seen from the price, while the accuracy using the C.45 algorithm or decision tree produces an accuracy of 94.67%.

Keywords: Internet, Data Analyst, Big Data, C.45 Algorithm, Covid 19

PENDAHULUAN

Perkembangan yang sangat pesat dalam bidang teknologi sekarang ini membuat banyak peluang usaha, salah satunya dibidang penjualan kartu internet. Teling Ponsel merupakan salah satu unit bisnis yang bergerak di bidang penjualan kartu kuota internet yang bertahan dalam persaingan. Teling mengklaim diri sebagai penyumbang terbesar dalam pasar kartu internet di kota Cirebon. Berbagai macam produk kartu internet dijual dan ditawarkan untuk menarik perhatian konsumen, tetapi tidak dapat diprediksi kartu internet apa saja yang diminati konsumen sehingga sering terjadi ketidaksesuaian penjualan kartu kuota internet. Manajemen telering Ponsel juga tidak dapat mengambil kebijakan dalam pengadaan kartu kuota internet Sehingga merugikan perusahaan. Banyak kartu kuota internet yang tidak terjual sama sekali dalam kurun waktu satu bulan. Kartu kuota internet yang tidak terjual sebagian dapat diretur kembali kepada supplier dan sebagian lagi terbuang. Sehingga pada permasalahan ini tidak hanya perusahaan Teling Ponsel yang mengalami kerugian tetapi pihak supplier juga merasakan hal yang sama.

Data mining bertujuan dalam penemuan pola secara otomatis atau semi otomatis dari sebuah data yang dimanfaatkan untuk menyelesaikan suatu masalah melalui berbagai aturan proses [1][2]. Prediksi atau klasifikasi adalah salah satu teknik data mining dalam memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi dimasa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Klasifikasi membantu memprediksi laku atau tidaknya produk sehingga perusahaan dapat mengambil keputusan dan meningkatkan strategi pemasaran[3]. Algoritma Decision Tree digunakan untuk mendapatkan informasi dengan tujuan pengambilan sebuah keputusan. Decision tree dimulai dengan sebuah titik awal yang digunakan oleh user untuk mengambil tindakan[4].

METODE PENELITIAN

2.1 Survei

Pada penelitian ini dilakukan survei yang bertujuan untuk mengamati transaksi yang dilakukan di

Teling serta meminta izin kepada pemiliknya untuk melakukan penelitian di tempat tersebut.

2.2 Studi Dokumentasi

Studi dokumentasi tersebut diantaranya adalah Data Motif batik Cirebon, Data Pertanyaan dan Jawaban Survei, dan Data Kuesioner.

2.3 Tahapan Penelitian

Berikut tahapan-tahapan yang akan dilakukan dalam penelitian ini :



Gambar 1. Alur Penelitian

A. Pengumpulan data

Tahap ini merupakan tahap awal penelitian dimana peneliti menghimpun data sekunder berupa data peminatan kuota konsumen Teling.

B. Data Cleaning

Proses data cleaning merupakan proses Pembersihan data yang bertujuan untuk menghilangkan data yang tidak memiliki nilai (null), data yang salah input, data yang tidak relevan, duplikat data dan data yang tidak konsisten [5] karena keberadaannya bisa mengurangi mutu atau akurasi dari hasil data mining nantinya. Pembersihan data juga akan mempengaruhi performansi dari sistem data mining karena data yang akan ditangani akan berkurang jumlah dan kompleksitasnya.

C. Transformasi data

Tahapan ini dilakukan dengan memberikan inisialisasi terhadap data yang memiliki nilai nominal menjadi bernilai numerik.

D. Klasifikasi

Pada tahap ini yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan. Ini adalah langkah penting di mana teknik kecerdasan diterapkan untuk mengekstrak pola informasi yang berpotensi berguna dari data yang dipilih.

E. Evaluasi

Pada tahap evaluasi, akan diketahui apakah hasil daripada tahap data mining dapat menjawab tujuan yang telah ditetapkan.

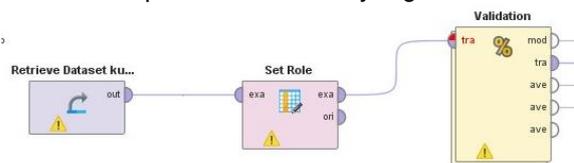
HASIL DAN PEMBAHASAN

Data yang digunakan untuk penelitian ini merupakan data peminatan kuota konsumen Teling. Adapun rincian detail dari Kriteria Pertama yaitu Kartu Internet dengan sub kriteria telkomsel, indosat, dan axis. Kriteria Kedua yaitu Kuota dengan sub kriteria adalah lebih dari 4 Gb dan Kurang dari 4 Gb. Kriteria ketiga adalah Masa Aktif dengan sub kriteria yaitu 1 minggu, 1 bulan dan 6 bulan. Kriteria Keempat adalah Harga dengan sub kriteria yaitu Tinggi dengan rate harga dari Harga Tinggi di atas Rp 100.000, Harga Sedang Rp 50.000 sd Rp 70.000 dan Harga Rendah Rp 10.000 sd Rp 49.000.

Tabel 1. Dataset kuota yang diminati konsumen Teling

No	Kartu	Kuota	Masa Aktif	Harga	Klasifikasi
1	Telkomsel	Lebih dari 4 Gb	1 Minggu	Tinggi	Minat
2	Indosat	Kurang dari 4 Gb	1 Bulan	Sedang	Tidak
3	Telkomsel	Kurang dari 4 Gb	6 Bulan	Rendah	Minat
4	Indosat	Lebih dari 4 Gb	1 Bulan	Rendah	Tidak
5	Telkomsel	Kurang dari 4 Gb	1 Bulan	Tinggi	Minat
6	Indosat	Lebih dari 4 Gb	6 Bulan	Tinggi	Minat
7	Indosat	Kurang dari 4 Gb	1 Bulan	Rendah	Tidak
8	Telkomsel	Kurang dari 4 Gb	6 Bulan	Rendah	Minat
9	Indosat	Lebih dari 4 Gb	1 Bulan	Rendah	Tidak
10	Telkomsel	Kurang dari 4 Gb	1 Bulan	Tinggi	Minat

Dari data tersebut peneliti melakukan cleaning atau pembersihan data kepada data yang belum lengkap agar data dapat digunakan untuk tahap selanjutnya pada rapidminer. Setelah itu pada data yang sudah bersih dilakukan transformasi data yang bertujuan untuk melakukan normalisasi pada data kriteria yang memiliki nilai.



Gambar 1. Proses data mining menggunakan rapidminer

Tahapan selanjutnya yaitu dilakukan proses data mining menggunakan bantuan tools rapidminer. Rapidminer merupakan sebuah alat yang berguna dalam membantu proses data mining. Pada rapidminer terdapat operator-operator untuk modifikasi dan mengolah data. Dalam penelitian ini tahapan pertama yang dilakukan menggunakan rapidminer adalah menarik data dari repository menggunakan operator retrieve. Selanjutnya gunakan operator set

role untuk menentukan fungsi dari setiap atribut dataset. Setelah itu peneliti menggunakan operator validation untuk membuat model data training dan data testing. Pada bagian training dimasukan operator decision tree sebagai model data mining yang akan digunakan dalam penelitian ini. Pada bagian testing dimasukan operator apply model dan performance untuk menguji keakuratan model tersebut.

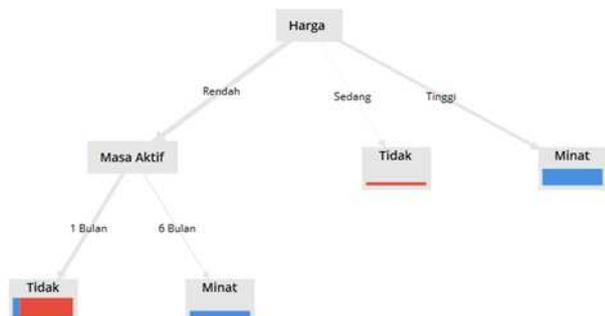
Dari hasil pemodelan data mining menggunakan algoritma Decision Tree didapatkan hasil akurasi senilai 94.67% dengan rincian sebagai berikut :

Tabel 2. Hasil klasifikasi menggunakan Decision Tree

	true Minat	true Tidak	class precision
pred. Minat	38	0	100%
pred. Tidak	4	33	89.19%
class recall	90.48%	100%	

Hasil Prediksi minat dan ternyata minat sebanyak 38. Hasil Prediksi minat dan ternyata tidak sebanyak 0. Hasil Prediksi tidak dan ternyata minat sebanyak 4. Hasil Prediksi tidak dan ternyata tidak sebanyak 33.

Faktor yang mempengaruhi konsumen dalam memilih kartu internet dilihat berdasarkan harga. Faktor tersebut dapat terlihat pada algoritma C.45 yang merupakan pohon keputusan pada kasus ini adapun pohon keputusan tersebut ada di bawah ini :



Gambar 2. Hasil decision tree

Berdasarkan gambar diatas tentang hasil pohon keputusan maka untuk menentukan minat dari keinginan konsumen dapat dilihat dari harga, penjelasannya sebagai berikut :

- Hasil pohon keputusan dalam menentukan minat konsumen dilihat dari Harga kategori rendah memiliki banyak peminat, akan tetapi harga rendah dilihat dari masa aktif 1 bulan memiliki peminat 13 dan tidak minat sebanyak 98. Sedangkan masa aktif 6 bulan memiliki peminat sebanyak 33 dan tidak minat 0.
- Hasil pohon keputusan dalam menentukan minat konsumen dilihat dari Harga kategori sedang tidak memiliki peminat.

- Hasil pohon keputusan dalam menentukan minat konsumen dilihat dari Harga kategori tinggi. Memiliki banyak peminat sebanyak 94 peminat.

KESIMPULAN

Dari hasil penelitian dapat ditarik kesimpulan bahwa harga kategori rendah adalah yang memiliki peminat paling banyak. Dengan demikian disarankan bagi Teling untuk memperbanyak stok kartu kuota yang berharga rendah atau murah.

REFERENSI

- [1] A. Bahtiar, "Prediksi Tingkat Kelulusan Mahasiswa Menggunakan Metode Hybrid," *J. Informatics, Sci. Technol.*, vol. 10, no. 1, pp. 1–6, 2019, [Online]. Available: <http://jurnal.stmik.banisaleh.ac.id/index.php/JIST/article/download/94/59>.
- [2] G. Dwilestari, Mulyawan, Martanto, and I. Ali, "Analisis Clustering menggunakan K-Medoid pada Data Penduduk Miskin Indonesia," *J. Sist. Inf. dan Manaj.*, vol. 9, no. 3, 2021.
- [3] Y. Yusuf, "Perbandingan Performansi Algoritma Decision Tree C5.0, Cart," *Seminar*, vol. 2007, no. Snati, pp. 0–3, 2007.
- [4] 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.
- [5] Mulyawan, G. Dwilestari, A. Bahtiar, F. M. Basysyar, and N. Suarna, "Classification of human development index using particle swarm optimization based on support vector machine algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1088, no. 1, p. 012033, 2021, doi: 10.1088/1757-899x/1088/1/012033.