

Penerapan Data Mining Algoritma Naive Bayes Clasifier Untuk Mengetahui Minat Beli Pelanggan Terhadap Kartu Telkomsel (Ricks Cell Simpang Candra)

Arya Atmanegara¹, Rts CiptaNingsi², Errissya Rasywir³, Despita Meisak⁴

^{1,2} Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Dinamika Bangsa, Kota Jambi, Indonesia

Email : yy4akz@gmail.com , rts.ciptaningsih@gmail.com

Email Penulis Korespondensi: yy4akz@gmail.com , rts.ciptaningsih@gmail.com ,errissya.rasywir@gmail.com, despitam88@gmail.com

Abstrak - Persaingan yang terjadi dalam dunia bisnis memaksa para pelakunya untuk selalumemikirkan strategi-strategi dan terobosan yang dapat menjamin kelangsungan dari bisnis yang dijalkannya. Hal ini akan memunculkan persaingan antar sesama provider kartu terhadap kartu internet. Para provider kartu internet berlomba-lomba menarik minat pelanggan dengan berbagai macam strategi pemasaran agar tidak kalah saing dan tetap eksis. Dan perusahaan ingin selalu meluncurkan kartu internet terbaru tanpa memikirkan kartu internet tersebut akan laku atau tidak dipasaran. Konsep data mining akan memudahkan cara menyelesaikan masalah yang terjadi di Ricks Cell Simpang Candra. Maka, metode klasifikasi mampu menemukan model yang membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Oleh sebab itu, algoritma naive bayes dapat memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya. Hasil dari penelitian ini agar dapat memprediksi atau memperkirakan laku atau tidak kartu internet yang baru, sehingga pemilik usaha dapat mengambil keputusan dan meningkatkan strategi pemasaran.

Kata Kunci : data mining, kartu internet, klasifikasi, algoritma naive bayes

Abstract - Competition is happening in the business world to force the perpetrators to always think of the strategy - and a breakthrough strategy that can guarantee the continuity of the business being operated. This leads to competition among card providers on the internet card. The provider of internet card race-the race to attract customers with a variety of marketing strategies in order not to lose competitiveness and still exist. And the company wants to always launch the latest internet card without thinking of internet cards will be sold in the market or not. The concept of data mining will ease the way to solve the problem that occurred in the Ricks Cell Intersection Candra. Thus, the classification method is able to find a model that distinguishes the concept or class of data, in order to be able to predict the class of an unknown object label. Therefore, Naive Bayes algorithm can predict future opportunities based on the experience of earlier. Results from this study in order to predict or estimate the behavior or the new Internet card, so that the bussines owner can make decisions and improve marketing strategy.

Keywords: data mining, internet cards, classification, naive Bayes algorithm

1. PENDAHULUAN

Persaingan disuatu usaha membuat para pemilik usaha harus selalumemikirkan strategi-strategi terobosanyang dapat menjamin kelangsungan bisnismereka. Pada suatu usaha memiliki ketersediaan data yang melimpah. Ini melahirkan kebutuhan akan adanyateknologi yang dapat memanfaatkannya untuk membangkitkan pengetahuan-pengetahuan baru, yang dapat membantudalam pengaturan strategi dalam menjalankan bisnis. Prediksi minat konsumen akan sangat penting bagi perusahaan, dimana dengan adanya prediksi minat konsumen perusahaan dapat mengambil suatu keputusan atau strategi yang benar dan tepat bagikonsumennya.

Database dalam jumlah yang besar merupakan salah satu aset berharga yang dimiliki sebuah perusahaan. Sebagai salah satu perusahaan yang bergerak di bidang bisnis telekomunikasi, pihak marketing Ricks Cell Simpang Candra haruslah memikirkan strategi dalam pemasaran untuk mampu menghasilkan suatu informasi yang siap digunakan untuk membantu pihak marketing dalam mengambil keputusan strategispemasaran. Untuk memenuhi kebutuhan-kebutuhan di atas, banyak cara yang dapat ditempuh. Salah satunya adalah dengan melakukan pemanfaatan *database* perusahaan menggunakan teknik *Data Mining*. Saat ini konter banyak sekali menjual kartu internet yang berbagai macam jenis tanpa memperkirakan kartu tersebut laku atau tidaknya dipasarkan dimasyarakat. Hal ini tentu akan membuat konter merugi jika kartu internet tersebut tidak laku terjual dimasyarakat.

Data Mining dikenal sebagai *Knowledge Discovery in Database* (KDD), *data mining* merupakan sebuah proses untuk menemukan hubungan atau pola yang memiliki makna dengan cara mengamati data yang besar, dimana datanya disimpan dalam sebuah penyimpanan dan mengelolanya dengan teknik statistic dan atau besar/banyak, dimensi data yang sangat tinggi dan kumpulan data yang beragam [1].

Metode *Klasifikasi* adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Dalam mencapai tujuan tersebut, proses *klasifikasi* membentuk suatu model yang mampu membedakan data kedalam kelas-kelas yang berbeda berdasarkan aturan atau fungsi tertentu. Model itu sendiri bisa berupa aturan “jika-maka”, berupa pohon keputusan, atau formula matematis.

Algoritma *Naive Bayes* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes*, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. *Teorema* tersebut dikombinasikan dengan *Naive* dimana diasumsikan kondisi antar atribut saling bebas.

2. METODOLOGI PENELITIAN

Metodologi penelitian yang digunakan dalam pembuatan Penelitian Penerapan Data Mining ini adalah dengan Algoritma *Naive Bayes Classifier*

2.1 Naive Bayes

Naive Bayes Classifier adalah metode klasifikasi atau teknik *machine learning* yang populer/umum digunakan dalam klasifikasi teks, memiliki performa yang baik pada banyak domain, sederhana dan efisien. Namun *Naive Bayes* sangat sensitive dalam pemilihan fitur, terlalu banyak jumlah fitur dapat mengakibatkan meningkatnya waktu perhitungan dan menurunkan akurasi klasifikasi [2].

2.1.1 Akurasi

Pengujian akurasi merupakan pengujian yang dilakukan untuk memperkirakan seberapa tepat hasil klasifikasi terhadap data yang ada, di mana persamaan 4 digunakan untuk menghitung akurasi [3].

2.2 Pengumpulan Data

Pada tahap ini yang dikumpulkan berupa data-data transaksi penjualan kartu *internet* TELKOMSEL sebagai sampel pada tahun 2022 pada Ricks Cell Simpang Candra.

a. Studi Literatur

Studi literatur yang digunakan dalam penulisan Penelitian ini adalah mengumpulkan bahan referensi mengenai metode *klasifikasi* dari berbagai buku, jurnal, dan beberapa referensi lainnya.

b. Observasi

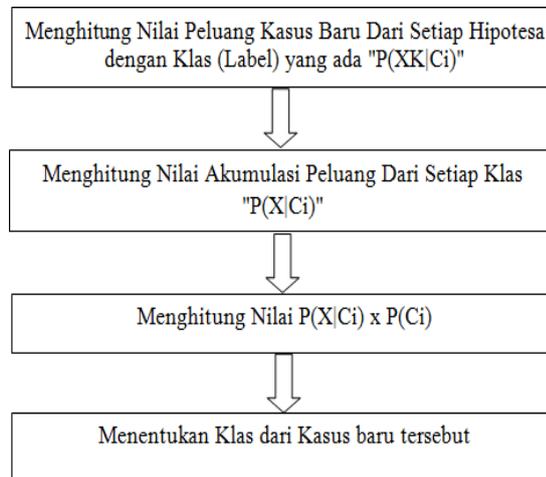
Metode pengumpulan data yang dilakukan dengan mengadakan pengamatan atau peninjauan langsung terhadap sumber permasalahan serta berkomunikasi langsung dengan pihak yang bersangkutan yaitu Pemilik dari Ricks Cell Simpang Candra tersebut.

c. Wawancara

Metode pengumpulan data yang dilakukan melalui tanya jawab yang diajukan secara langsung kepada narasumber untuk mendapatkan data atau informasi yang berkaitan dengan objek penelitian.

3. HASIL DAN PEMBAHASAN

3.1 Diagram Penyelesaian Masalah Naive Bayes



Gambar 1. Tahapan Algoritma Naive Bayes

Tabel 1. Data Kasus

No	Kuota	Masa Aktif	Harga	Klasifikasi
1	Dibawah 4gb	3hari	Rendah	Laris
2	Dibawah 4gb	3hari	Rendah	Laris
3	Dibawah 4gb	3hari	Rendah	Laris
4	Diatas 4gb	1minggu	Sedang	Kurang
5	Diatas 4gb	1minggu	Sedang	Kurang
6	Diatas 4gb	1minggu	Sedang	Kurang
7	Dibawah 4gb	1minggu	Rendah	Laris
8	Dibawah 4gb	1minggu	Rendah	Laris
9	Dibawah 4gb	1minggu	Rendah	Laris
10	Dibawah 4gb	1minggu	Rendah	Laris
11	Diatas 4gb	1bulan	Tinggi	Laris
12	Diatas 4gb	1bulan	Tinggi	Laris
13	Diatas 4gb	1bulan	Tinggi	Kurang
14	Diatas 4gb	1bulan	Tinggi	Kurang
15	Diatas 4gb	1bulan	Tinggi	Laris
16	Dibawah 4gb	1minggu	Rendah	Laris
17	Dibawah 4gb	1minggu	Rendah	Laris
18	Dibawah 4gb	1minggu	Rendah	Laris
19	Diatas 4gb	1bulan	Tinggi	Laris
20	Diatas 4gb	1bulan	Tinggi	Laris

Keterangan:

- Kriteria 1 = Menjelaskan tentang Kriteria “Kuota”
- Kriteria 2 = Menjelaskan tentang Kriteria “Masa Aktif”

- Kriteria 3 = Menjelaskan tentang Kriteria “Harga”

Tahap awal cara kerja dari proses perhitungan *Naive Bayes* adalah dengan cara melakukan pengambilan data training dari penjualan kartu *internet TELKOMSEL Ricks Cell*. Adapun variabel penentu yang digunakan dalam mengklasifikasikan data kartu *internet* yaitu:

A. Kuota

Merupakan variabel yang dikelompokkan dalam dua kategori yaitu kuota diatas 4GB dan kuota dibawah 4GB. Kuota diatas 4GB merupakan batas kuota dari kuota 4GB, 5GB, 17GB dan 21GB. Kuota dibawah 4GB merupakan batas kuota dari kuota 3,5GB, 2GB, 1,5GB,.

Tabel 2. Kuota

Kriteria	Kategori	Keterangan
Kuota	Diatas 4GB	4GB – 21 GB
	Dibawah 4GB	1,5GB – 3,5GB

B. Masa Aktif

Merupakan variabel lamanya masapakai kartu internet yang dikelompokkan dalam 3 kategori yaitu 3hari, 1minggu, dan 1bulan. Kriteria masa aktif dapat dilihat dari tabel 3.

Tabel 3 Kriteria Masa Aktif

Kriteria	Kategori	Keterangan
Masa Aktif	3hari	3 hari di mulai dari mengaktifkan kuota
	1minggu	7 hari di mulai dari mengaktifkan kuota
	1bulan	30 hari di mulai dari mengaktifkan kuota

C. Harga

Merupakan variable jenis harga kartu internet yang dikelompokkan dalam tiga kategori yaitu Rendah, Sedang, dan Tinggi.

Adapun kategori tersebut mempunyai nilai batas ambang yaitu:

- Kategori “rendah” mulai dari harga Rp8.000,00 s/d Rp 20.000,00
- Kategori “sedang” mulai dari harga Rp28.000,00 s/d Rp 33.000,00
- Kategori “tinggi” mulai dari harga Rp4000,00 s/d 85.000,00.

Kriteria harga dilihat dari tabel 4

Tabel 4 Kriteria Harga

Kriteria	Kategori	Keterangan
Harga	Rendah	- harga Rp 8.000,00 s/d Rp 20.000,00
	Sedang	- harga Rp 28.000,00 s/d Rp 33.000,00
	Tinggi	- harga Rp 40.000,00 s/d 85.000,00.

4. KESIMPULAN

Berdasarkan perancangan, analisis, implementasi dan pengujian pada Penerapan Data Mining dengan Algoritma *Naive Bayes Classifier* Untuk Mengetahui Minat Beli Pelanggan Terhadap Kartu *Internet* TELKOMSEL dengan Menggunakan Metode Klasifikasi, maka dapat ditarik beberapa kesimpulan sebagai berikut:

1. Sistem klasifikasi data kartu internet ini digunakan untuk menampilkan informasi klasifikasi laris atau kurang larisnya minat masyarakat pada kartu internet TELKOMSEL ini dengan menggunakan algoritma *Naive Bayes*.
2. Algoritma *Naive Bayes* sangat cocok diterapkan dalam memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya sehingga memudahkan perusahaan untuk memprediksi peminatan masyarakat terhadap kartu internet jika ingin meluncurkan kartu internet jenis yang baru.
3. Dengan mengetahui Laris atau kurang larisnya kartu internet TELKOMSEL ini, akan meminimalisir kerugian pada perusahaan.

UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada Ibu Errissya Rasywir, S.Kom, MT selaku dosen Mata Kuliah Data Mining, dan selaku pembimbing atas pengerjaan penelitian ini, Terima kasih kepada pihak Ricks Cell yang telah memberikan data pembelian kartu internet TELKOMSEL serta pihak-pihak yang telah mendukung terlaksananya penelitian ini.

REFERENCES

- [1] M. A. Maricar, P. Widiadnyana and W. A. Wijaya, "Analysis of Data Mining for Forecasting Total Goods Delivery with Moving Average Method (Case Study: Agent "X" Expedition "Z")," International Journal of Engineering and Emerging Technology, Vol. 2, No. 1, pp. -10, 2017.
- [2] D. A. Muthia, "Opinion Mining pada Review Buku Menggunakan Algoritma Naive Bayes," Jurnal Teknik Komputer Amik BSI, Vol.II, No.1, pp. 1-8, 2016.
- [3] M. A. Maricar, I. N. S. Kumara and M. Sudarma, "Opinion Mining on Twitter Social Media to Classify Racism Using Combination of POS Tagging, Naive Bayes Classifier, and K-Nearest Neighbor," in International Conference on Smart-Green Technology in Electrical and Information System, Bali, 2018.