

MEMPREDIKSI LOYALITAS NASABAH PADA BMT ABC DENGAN METODE INTERATIVE DICHOTOMISER THREE (ID 3)

Arief Soma Darmawan¹⁾, Devi Sugianti²⁾, Wim Hapsoro³⁾

STMIK Widya Pratama¹²³⁾

Ariefsoma24@gmail.com¹⁾, devi.sugianti9807@gmail.com²⁾, wimhapsoro06@gmail.com³⁾

Abstrak

Persaingan perusahaan semakin ketat menuntut perusahaan memberikan pelayanan yang memuaskan kepada konsumen. Dalam mempertahankan konsumen salah satu kuncinya adalah mengelola informasi pelanggan. Perusahaan akan terbantu jika mempunyai sistem informasi pelanggan dikelola dengan baik, data pelanggan yang baru ataupun data pelanggan yang lama. BMT ABC mempunyai mempunyai 4.882 nasabah, akan tetapi sampai tahun 2020 nasabah yang aktif melakukan penabungan adalah 2.392 nasabah. Dalam memprediksi nasabah yang aktif di BMT ABC menggunakan metode ID3. ID 3 adalah salah satu metode klasifikasi pada data mining decision tree (pohon keputusan). Langkah-langkah penelitian: identifikasi dan perumusan masalah, penentuan variabel penelitian, penentuan data latih, hitung gain dan entropy, membangun decision tree dengan algoritma ID3. Pada penelitian ini menggunakan atribut: jenis kelamin, usia, pekerjaan, pendidikan, status, dan setoran awal. Dari hasil perhitungan Gain dan Entropy dihasilkan 5 aturan dan menghasilkan pohon keputusan

Kata kunci: Loyalitas Nasabah, ID 3

1. Pendahuluan

Dengan berkembangnya teknologi pengusaha berlomba-lomba untuk mencari terobosan baru (Febtadianrano, vulandari dan laksito 2020). Persaingan antar perusahaan semakin ketat, menuntut perusahaan memberikan pelayanan yang memuaskan kepada konsumen agar perusahaan menjadi terdepan dan terbaik (izzaty 2016). Salah satu kegiatan yang dilakukan oleh perusahaan secara berkala adalah pengukuran tingkat kepuasan pelanggan (Tama 2015). Salah satu kunci untuk mempertahankan keunggulan kompetitif dengan mengelola informasi pelanggan (fauziah, hartama dan sudahri 2020). Perusahaan untuk dapat melihat pelanggan yang berpotensi atau tidak berpotensi belum diolah secara optimal, karena mengalami kesulitan dalam pengolahannya (Febtadianrano, vulandari dan laksito 2020). Perusahaan akan terbantu jika mempunyai sistem informasi pelanggan dikelola dengan baik, data pelanggan yang baru ataupun data pelanggan yang lama (Ulin dan Fatmawati 2019)

BMT ABC berdiri pada tahun 2000, dengan jumlah nasabah sampai tahun 2020 bulan januari adalah 4.882 nasabah. Akan tetapi yang aktif melakukan penabungan sampai di bulan januari 2020 adalah 2.392 nasabah yang aktif masih melakukan penabungan, nasabah yang tidak aktif 2.490 nasabah. Dengan rata-rata penabungan per bulan adalah 3.388 transaksi dengan nominal sekitar Rp 842.926.300 dalam satu bulan pada tahun 2019. BMT ABC belum dapat mengklasifikasikan pelanggan yang aktif dan tidak aktif. Maka dari itu diperlukan klasifikasi pelanggan

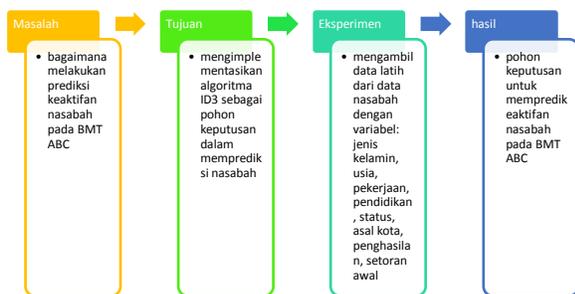
Untuk dapat memprediksi nasabah yang aktif dan tidak aktif pada BMT ABC dapat dikembangkan suatu sistem dengan memanfaatkan berbagai metode. Salah satu metode yang dapat digunakan adalah data mining yang dapat menggali informasi yang tersembunyi dari sebuah data, data mining dapat menemukan pola atau aturan dengan menganalisa data yang ada (Inda dan nasir 2019). Salah satu metode

klasifikasi pada data mining adalah *decision tree* (pohon keputusan). Metode tersebut dapat digunakan untuk prediksi penyakit diabetes (subhan dan wibawa 2018), pengukuran kinerja PNS (adiana, sediyono dan nurhayati 2014) dan dapat digunakan untuk menprediksi loyalitas pelanggan (santoso 2014). Algoritma *decision tree* yang digunakan adalah algoritma ID3, keuntungan menggunakan algoritma ID3 adalah membentuk pohon keputusan dengan tingkat variasi yang komplit dapat memangkas pohon menjadi lebih sederhana. (adiana, sediyono dan nurhayati 2014)

2. Metode Penelitian

Masalah yang dihadapi oleh BMT ABC adalah memprediksi loyalitas nasabah agar selalu aktif. Dari data nasabah yang tidak pernah di olah untuk mendapatkan informasi tentang keaktifan nasabah. Dalam melakukan prediksi menggunakan metode data mining dengan menghasilkan pohon keputusan menggunakan algoritma ID3. Dimana variabel yang digunakan adalah jenis kelamin, usia, pekerjaan, pendidikan, status, setoran awal dan yang menjadi variabel prediktor adalah keaktifan dari nasabah tersebut.

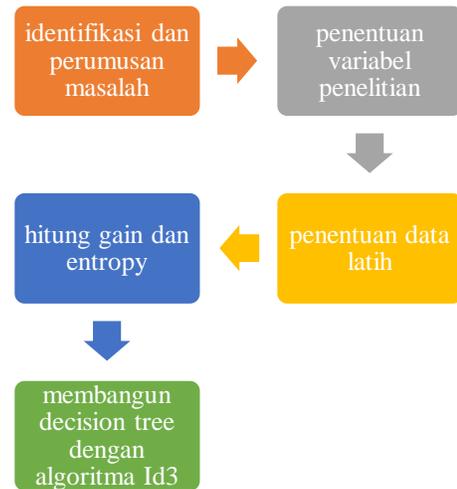
Kerangka pemikiran dari penelitian ini adalah



Gambar 1. Kerangka pemikiran penelitian

Dalam memprediksi keaktifan nasabah pada BMT ABC dengan tahapan yang harus dilaksanakan untuk memastikan mencapai hasil yang maskimal. Tahapan dari prosedur penelitian pertama dengan mengidentifikasi dan merumuskan masalah yang menjadi tujuan dari penelitian ini, dengan memprediksi keaktifan

nasabah. Adapun langkah penelitian adalah sebagai berikut:



Gambar 2 Langkah Penelitian.

Dalam perhitungan atribut atau fitur sebagai akar maka dilakukan perhitungan nilai entropy terlebih dahulu dengan menggunakan rumus perhitungan sebagai berikut: (Pudjo, Handayanto dan Helawati 2013)

$$Entropy (s) = \sum_{i=1}^c -P_i \log_2 p_i \dots\dots\dots(1)$$

Keterangan:

S= Himpunan Kasus

P_i= Proporsi dari S_i terhadap S

Sedangkan untuk memilih atribut sebagai akar didasarkan pada nilai Gain tertinggi dari atribut-atribut yang ada, gain dilakukan dengan rumus sebagai berikut:

$$Gain (S, A) = Entropy(s) - \sum \frac{|s_v|}{|s|} entropy (s) \dots\dots\dots(2)$$

Keterangan:

Gain (S,A) = information gain dari sebuah atribut A pada koleksi S

S = Himpunan Kasus

A = Atribut

|S_i| = Jumlah kasus pada Partisi ke I

|S| = Jumlah kasus dalam S

3. Hasil dan Pembahasan

Untuk penelitian ini menggunakan pengitungan manual dengan 9 nasabah. Dengan 6 atribut yang dijadikan sebagai acuan untuk proses perhitungan. Berikut ini langkah-langkah penelitian yang dirinci sebagai berikut:

3.1 Identifikasi dan Perumusan masalah

BMT ABC mengalami kesulitan dalam memprediksi loyalitas nasabah. Loyalitas nasabah diperlukan agar BMT dapat menjaga agar nasabah selalu aktif dalam penabung.

3.2 Penentuan variabel penelitian

BMT melakukan pendataan nasabah saat pembukaan rekening baru. Tidak semua data nasabah dijadikan variabel untuk proses perhitungan prediksi. Adapun yang dijadikan variabel untuk proses perhitungan adalah: Jenis kelamin, usia, pekerjaan, pendidikan, status, dan setoran awal. Untuk variabel prediktor adalah keaktifan dari nasabah. Adapun untuk pembagian kategori dari variabel seperti tabel 1 dibawah ini

Tabel 1 pembagian kategori variabel

Nomor	atribut	kategori	ket
1	Jenis kelamin	L	Laki-laki
		P	Perempuan
2	Usia	muda	>25
		Dewasa	25-59
		Tua	>59

Tabel 2. Klasifikasi Nasabah

jenis kelamin	usia	pekerjaan	pendidikan terakhir	status	setoran awal	keaktifan
L	MUDA	TIDAK BEKERJA	B	BELUM MENIKAH	SEDANG	YA
L	TUA	TIDAK BEKERJA	B	MENIKAH	TINGGI	YA
P	MUDA	BEKERJA	C	MENIKAH	SEDANG	YA
P	TUA	TIDAK BEKERJA	C	MENIKAH	TINGGI	TIDAK
L	DEWASA	BEKERJA	A	BELUM MENIKAH	SEDANG	YA
P	TUA	BEKERJA	A	MENIKAH	TINGGI	YA
P	TUA	BEKERJA	B	MENIKAH	SEDANG	TIDAK
P	DEWASA	TIDAK BEKERJA	B	BELUM MENIKAH	RENDAH	TIDAK
L	MUDA	TIDAK BEKERJA	B	MENIKAH	RENDAH	YA

Nomor	atribut	kategori	ket
3	Pekerjaan	BL bekerja Bekerja	Belum bekerja Bekerja SD,SMP
4.	Pendidikan Terakhir	A	
		B	SMA, D1,D2,D3
		C	S1,S2,S3
5	Status	Menikah	
		Blm Menikah	Belum menikah
6	Setoran awal	rendah	<500.000
		Sedang	500.000- 2.500.000
		Tinggi	>2.500.000

3.3 Penentuan data latih

Dari data nasabah yang ada yang akan dijadikan sebagai data latih terdapat 9 data nasabah yang dijadikan sebagai data latih atau dataset. Berikut ini adalah data latih yang disajikan pada tabel 2.

3.4 Hitung gain dan entropy

Dalam menghitung gain entropy dapat dilakukan sebagai berikut, dimulai dengan node akar.

Entropy untuk semua

$$E(\text{semua}) = -((p(\text{ya}|\text{semua}) \times \log_2 p(\text{ya}|\text{semua})) + (p(\text{tidak}|\text{semua}) \times \log_2 P(\text{tidak}|\text{semua})))$$

$$E(\text{semua}) = -\left(\left(\frac{6}{9}\right) \times \log_2 \left(\frac{6}{9}\right) + \left(\frac{3}{9}\right) \times \log_2 \left(\frac{3}{9}\right)\right)$$

=0.918

Entropy untuk jenis kelamin laki-laki

$$E(\text{semua}_{\text{laki-laki}}) = -((p(\text{ya}|\text{laki-laki}) \times \log_2 p(\text{ya}|\text{laki-laki})) + (p(\text{tidak}|\text{laki-laki}) \times \log_2 p(\text{tidak}|\text{laki-laki})))$$

$$E(\text{semua}_{\text{laki-laki}}) = -\left(\left(\frac{4}{4}\right) \times \log_2 \left(\frac{4}{4}\right) + \left(\frac{4}{0}\right) \times \log_2 \left(\frac{4}{0}\right)\right)$$

=0

Entropy untuk jenis kelamin perempuan

$$E(\text{semua}_{\text{perempuan}}) = -((p(\text{ya}|\text{perempuan}) \times \log_2 p(\text{ya}|\text{perempuan})) + (p(\text{tidak}|\text{perempuan}) \times \log_2 p(\text{tidak}|\text{Perempuan})))$$

$$E(\text{semua}_{\text{perempuan}}) = -\left(\left(\frac{2}{5}\right) \times \log_2 \left(\frac{2}{5}\right) + \left(\frac{3}{5}\right) \times \log_2 \left(\frac{3}{5}\right)\right)$$

=0.971

$G(\text{semua}, \text{Jenis kelamin})$

$$= E(\text{semua}) - ((p(\text{Laki-laki}|\text{semua}) \times E(\text{semua}_{\text{Laki-laki}})) + (p(\text{Perempuan}|\text{semua}) \times E(\text{semua}_{\text{perempuan}})))$$

$$= 0.918 - ((4/9) \times 0) + ((5/9) \times 0.971)$$

=0.379

Tabel 3 hasil perhitungan entropy dan gain untuk node akar

node		jumlah	ya	tidak	entropy	gain	
1 total		9	6	3	0,918		
	jenis kelamin	L	4	4	0	0,000	0,379
		P	5	2	3	0,971	
Usia	muda	3	3	0	0,000	0,252	
	dewasa	2	1	1	1,000		
	tua	4	2	2	1,000		
Pekerjaan	Bekerja	4	3	1	0,811	0,018	
	tidak bekerja	5	3	2	0,971		
pendidikan terakhir	A	2	2	0	0,000	0,157	
	B	5	3	2	0,971		
	C	2	1	1	1,000		
status	menikah	6	4	2	0,918	0,000	

	belum menikah	3	2	1	0,918	
setoran awal	rendah	2	1	1	1,000	0,029
	sedang	4	3	1	0,811	
	tinggi	3	2	1	0,918	

Tabel 3 hasil perhitungan entropy dan gain untuk node akar, menunjukkan nilai gain tertinggi adalah jenis kelamin. Maka dari itu jenis kelamin dijadikan node akar. Selanjutnya data akan dipecah menjadi 2 yaitu “ Laki-laki “ dan “ Perempuan”. Hasil pohon yang dibentuk disajikan pada gambar 3. Pemisah data Tabel 4 hasil pemisahan data oleh fitur “ jenis kelamin

didapatkan dari fitur jenis kelamin dapat dilihat pada tabel 4. Sebagai berikut:

jenis kelamin	usia	pekerjaan	pendidikan terakhir	status	setoran awal	keaktifan
L	MUDA	TIDAK BEKERJA	B	BELUM MENIKAH	SEDANG	YA
L	TUA	TIDAK BEKERJA	B	BELUM MENIKAH	TINGGI	YA
L	DEWASA	TIDAK BEKERJA	A	BELUM MENIKAH	SEDANG	YA
L	MUDA	BEKERJA	B	MENIKAH	RENDAH	YA
P	MUDA	BEKERJA	C	MENIKAH	SEDANG	YA
P	TUA	TIDAK BEKERJA	C	MENIKAH	TINGGI	TIDAK
P	TUA	BEKERJA	A	MENIKAH	TINGGI	YA
P	TUA	BEKERJA	B	MENIKAH	SEDANG	TIDAK
P	DEWASA	TIDAK BEKERJA	B	BELUM MENIKAH	RENDAH	TIDAK

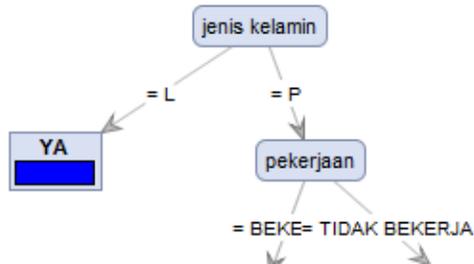
Dalam perhitungan entropy dan gain berikutnya, fitur “jenis kelamin” tidak dilibatkan. Berdasarkan pemisahan hasil dari jenis kelamin “Laki-laki” entropy node bernilai nol sehingga dapat dipastikan semua data di node cabang jenis kelamin laki-laki masuk dikelas yang sama dan

node yang dituju oleh cabang jenis kelamin laki-laki menjadi daun (node terminal) seperti pada gambar 3, dari pemisahan jenis kelamin “perempuan” maka didapatkan bahwa tidak bekerja keaktifannya adalah “tidak” seperti terlihat pada tabel 5.

Tabel 5. Hasil pemisahan data dari fitur pekerjaan

jenis kelamin	usia	pekerjaan	pendidikan terakhir	status	setoran awal	keaktifan
P	MUDA	BEKERJA	C	MENIKAH	SEDANG	YA

P	TUA	BEKERJA	A	MENIKAH	TINGGI	YA
P	TUA	BEKERJA	B	MENIKAH	SEDANG	TIDAK
P	TUA	TIDAK BEKERJA	C	MENIKAH	TINGGI	TIDAK
P	DEWASA	BEKERJA	B	BELUM MENIKAH	RENDAH	TIDAK



Dari tabel 5 hasil pemisahan data fitur pekerjaan maka didapatkan bahwa “ tidak bekerja” maka keaktifan tidak aktif. Untuk node 3. Dapat dihitung seperti pada tabel 6.

Gambar 3 decision tree untuk node 2

Tabel 6 Hasil perhitungan entropy dan gain untuk node 3

node		jumlah	ya	tidak	entropy	gain
3	total	3	2	1	0,918	
	Usia					
	muda	1	1	0	0,000	0,252
	dewasa	0	0	0	0,000	
	tua	2	1	1	1,000	
	pendidikan terakhir					
	A	1	1	0	0,000	0,918
	B	2	0	1	0,000	
	C	2	1	0	0,000	
	status					
	menikah	3	2	1	0,918	0,000
	belum menikah	0	0	1	0,000	
	setoran awal					
	rendah	0	0	0	0,000	0,252
	sedang	2	1	1	1,000	
	tinggi	1	1	0	0,000	

Hasil yang didapat di Tabel 6 menunjukkan bahwa gain tertinggi ada difitur “pendidikan” sehingga pendidikan dijadikan sebagai node

internal (node3). Untuk cabangnya digunakan 3 nilai yaitu A, B, C. Dari data pemisah dapat dilihat pada Tabel 7.

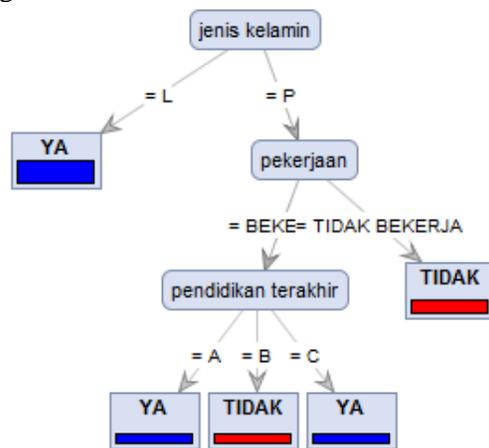
Tabel 7 Hasil pemisahan data dengan fitur “Pendidikan”

jenis kelamin	usia	pekerjaan	pendidikan terakhir	status	setoran awal	keaktifan
P	MUDA	BEKERJA	C	MENIKAH	SEDANG	YA
P	TUA	BEKERJA	A	MENIKAH	TINGGI	YA
P	TUA	BEKERJA	B	MENIKAH	SEDANG	TIDAK

Dari tabel 7 dapat dilihat bahwa fitur pendidikan terakhir “A” maka keaktifan “Ya”, untuk pendidikan “B” maka keaktifan “Tidak” dan untuk pendidikan terakhir “C” maka keaktifannya adalah “Ya”. Hasil akhir dari inuksi decision tree disajikan pada gambar 4.

3.5 Membangun decision tree dengan algoritma ID3

Setelah dilakukan proses perhitungan Gain dan Entropy dapat dihasilkan decision tree seperti pada gambar 4.



Gambar 4. Decision tree untuk kasus

Aturan yang diperoleh dari pohon keputusan pada gambar 5. Dengan berdasarkan atribut yang digunakan untuk keaktifan nasabah adalah sebagai berikut:

1. **IF** Jenis kelamin= L **Then** keaktifan =ya
2. **IF** Jenis kelamin= P **And** Pekerjaan=Tidak Bekerja **Then** keaktifan =tidak

3. **IF** Jenis Kelamin = P **And** Pekerjaan = Bekerja **And** Pendidikan= A **Then** = ya
4. **IF** Jenis Kelamin = P **And** Pekerjaan = Bekerja **And** Pendidikan= B **Then** = Tidak
5. **IF** Jenis Kelamin = P **And** Pekerjaan = Bekerja **And** Pendidikan= C **Then** = ya

Dari pohon keputusan yang diperoleh pada gambar 4. Nasabah yang aktif dalam penabungan dapat dilihat dari deskripsi gambar pohon keputusan seperti gambar 5 dibawah ini

Tree

```

jenis kelamin = L: YA {YA=4, TIDAK=0}
jenis kelamin = P
| pekerjaan = BEKERJA
| | pendidikan terakhir = A: YA {YA=1, TIDAK=0}
| | pendidikan terakhir = B: TIDAK {YA=0, TIDAK=1}
| | pendidikan terakhir = C: YA {YA=1, TIDAK=0}
| pekerjaan = TIDAK BEKERJA: TIDAK {YA=0, TIDAK=2}

```

Gambar 5 Deskripsi Pohon Keputusan

4. Kesimpulan dan Saran

Pohon keputusan yang dihasilkan oleh algoritma ID 3 dengan pohon yang lebih sederhana setelah dilakukan proses pemangkasan, variasi data yang digunakan bisa bersifat komplit. Dari penelitian ini dihasilkan 5 aturan dengan output keaktifan dari nasabah penabungan. Node akar yang dihasilkan adalah jenis kelamin. Pada penelitian selanjutnya, perlu dikaji dengan kondisi data yang lebih banyak dan atribut yang lebih banyak pula. Selain itu metode seleksi atribut lain seperti *forward selection*, *backward elimination*.

DAFTAR PUSTAKA

- adiana, julce sidette, eko sediyono, dan oky dewi nurhayati. 2014. "pendekatan metode pohon keputusan menggunakan algoritma ID3 untuk sistem informasi pengukuran kinerja PNS." *sistem informasi bisnis* 02 (1): 75-86.
- fauziah, dedy hartama, dan irfan damanik sudahri. 2020. "analisa kepuasan pelanggan menggunakan klasifikasi data mining." *brahmana* 2 (1): 41-48.
- Febtadianrano, Hakam Putro, retno tri vulandari, dan wawan laksito. 2020. "penerapan metode naive bayes untuk klasifikasi pelanggan." *Tikomsin* 8 (2): 19-24.
- Inda, Putri sari, dan muhammad nasir. 2019. "implementasi model decision tree menggunakan algoritma ID3 untuk menentukan strategi pemasaran universitas bina darma Palembang." *bina darma conference on computer science* 1 (1): 739-754.
- izzaty, annisak jamhur. 2016. "penerapan data mining untuk menganalisa jumlah pelanggan aktif dengan menggunakan algoritma c4.5." *Majalah Ilmiah* 23 (2): 12-20.
- Pudjo, Prabowo widodo, Tria Rahmadya Handayanto, dan Helawati. 2013. *Penerapan data mining dengan matlab*. Bandung: Rekayasa Sains.
- santoso, T.B. 2014. "analisa dan penerapan metode C4.5 untuk prediksi loyalitas pelanggan." *universitas satya negara indonesia*.
- subhan, muhamad efendi, dan helmie arif wibawa. 2018. "prediksi penyakit diabetes menggunakan algoritma ID3 dengan pemilihan atribut terbaik." *Juita* 6 (1): 29-36.
- Tama, B.A. 2015. "Data mining for predicting customer satisfaction." *J Theor appl Inf Technol* 75 (1): 3-7.
- Ulin, Khilyanti fitri, dan Azizah Fatmawati. 2019. "Sistem Informasi pelanggan pada bengkel marno jaya motor." *teknik elektro* 19 (1): 29-35.