

# KLASIFIKASI UNIT SIMPAN PINJAM KOPERASI DENGAN METODE *NAÏVE BAYES CLASSIFIER* PADA DINAS KOPERASI USAHA MIKRO KECIL DAN MENENGAH

Noor Rahmat Effendy<sup>1</sup>, Setyo Wahyu Saputro<sup>2</sup>

Program Studi Teknik Informatika, STMIK Banjarbaru

Jl. A. Yani Km. 33,3 Banjarbaru, Telp (0511) 4782881

<sup>1</sup>fendy.anjapu@gmail.com, <sup>2</sup>setyows@gmail.com

## Abstrak

Untuk terwujudnya pengelolaan USP Koperasi yang sehat dan sesuai ketentuan peraturan perundang-undangan. Maka dilakukan penilaian terhadap USP Koperasi. Aspek penilaian meliputi permodalan, kualitas aktiva produktif, manajemen, efisiensi, likuiditas, kemandirian dan pertumbuhan, dan jati diri koperasi. Hasil perhitungan penilaian terhadap tujuh aspek tersebut dijumlah, dan diperoleh skor secara keseluruhan. Skor 100 sampai dengan 80 dikategorikan sehat, skor di bawah 80 sampai dengan 60 dikategorikan cukup sehat, skor di bawah 60 sampai dengan 40 dikategorikan kurang sehat, skor di bawah 40 sampai dengan 20 dikategorikan tidak sehat, skor di bawah 20 sampai dengan 0 dikategorikan sangat tidak sehat. Ada USP Koperasi yang mendapatkan skor mendekati 60, walaupun skornya hanya sedikit di bawah 60, dengan ketentuan seperti yang sudah disebutkan, USP Koperasi tersebut dikategorikan kurang sehat, namun jika dihitung rasio skor masing-masing aspek terhadap skor maksimalnya, USP Koperasi tersebut memiliki lebih banyak nilai aspek yang cukup sehat dari pada nilai yang kurang sehat, seharusnya koperasi ini dikategorikan cukup sehat.

*Naïve Bayes Classifier* merupakan salah satu metode di dalam *data mining* untuk mengklasifikasikan data. Metode *Naïve Bayes Classifier* akan diaplikasikan dalam penelitian ini untuk menentukan predikat USP Koperasi. Hasil penelitian menunjukkan *Naïve Bayes Classifier* dapat memecahkan masalah dengan cukup baik. Hasil penelitian menunjukkan kinerja sistem sebesar 93%.

**Kata Kunci:** Simpan Pinjam, Klasifikasi, *Data mining*, *Naïve bayes classifier*

## Abstract

*For the realization of healthy management of Cooperatives and in accordance with the provisions of legislation. Then the assessment of Cooperatives. Assessment aspects include capital, productive asset quality, management, efficiency, liquidity, self-reliance and growth, and cooperative identity. The results of the assessment of the seven aspects are summed, and the overall score is obtained. Scores of 100 to 80 are categorized as healthy, scores below 80 to 60 are categorized as quite healthy, scores below 60 to 40 are categorized as less healthy, scores below 40 to 20 are categorized as unhealthy, scores under 20 to 0 are highly categorized not healthy. There is a Cooperative that scores close to 60, although the score is only slightly below 60, with the provision as already mentioned, the Cooperative is categorized as less healthy, but if the calculated ratio of each aspect ratio to the maximum score, the cooperative has more the value of healthy aspects of the value of less healthy, this cooperative should be categorized quite healthy.*

*Naïve Bayes Classifier is one method in data mining to classify data. Method Naïve Bayes Classifier will be applied in this research to determine predicate Cooperative. The results show Naïve Bayes Classifier can solve the problem quite well. The results show the system performance of 93%.*

**Keyword:** *Saving and Loans, Classification, Data mining, Naïve bayes classifier*

## 1. Pendahuluan

Unit Simpan Pinjam Koperasi yang selanjutnya disebut USP Koperasi adalah unit koperasi yang bergerak di bidang usaha simpan pinjam sebagai bagian dari kegiatan usaha Koperasi yang bersangkutan. [1]

Untuk terwujudnya pengelolaan USP Koperasi yang sehat dan sesuai ketentuan peraturan perundang-undangan. Maka dilakukan penilaian terhadap USP Koperasi. Aspek penilaian meliputi permodalan, kualitas aktiva produktif, manajemen, efisiensi, likuiditas, kemandirian dan pertumbuhan, dan jati diri koperasi. Hasil perhitungan penilaian terhadap tujuh aspek tersebut dijumlah, dan diperoleh skor secara keseluruhan. Skor 100 sampai dengan 80 dikategorikan sehat, skor di bawah 80 sampai dengan 60 dikategorikan cukup sehat, skor di bawah 60 sampai dengan 40 dikategorikan kurang sehat, skor di bawah 40 sampai dengan 20 dikategorikan tidak sehat, skor di bawah 20 sampai dengan 0 dikategorikan sangat tidak sehat. Ada USP Koperasi yang mendapatkan skor mendekati 60, walaupun skornya hanya sedikit di bawah 60, dengan ketentuan seperti yang sudah disebutkan, USP Koperasi tersebut dikategorikan kurang sehat, namun jika dihitung rasio skor masing-masing aspek terhadap skor maksimalnya, USP Koperasi tersebut memiliki lebih banyak nilai aspek yang cukup sehat dari pada nilai yang kurang sehat, seharusnya koperasi ini dikategorikan cukup sehat. Akibat dikategorikan kurang sehat, Koperasi yang bersangkutan tidak bisa melakukan pinjaman pada bank untuk mengembangkan USP Koperasinya.

*Naïve Bayes Classifier* merupakan salah satu metode di dalam *data mining* untuk mengklasifikasikan data. Cara kerja dari metode *Naïve Bayes Classifier* menggunakan perhitungan probabilitas. *Naïve Bayes Classifier* menghitung peluang dari satu kelas dari masing-masing kelompok atribut yang ada, dan menentukan kelas mana yang paling optimal. Proses pengelompokan atau klasifikasi dibagi menjadi dua fase yaitu *learning/training* dan *testing/classify*. Pada fase *learning*, sebagian data yang telah diketahui kelasnya, datanya diumpungkan untuk membentuk model perkiraan. Kemudian pada fase *testing*, model yang sudah terbentuk diuji dengan sebagian data.

Penggunaan *Naïve Bayes Classifier* telah diuji oleh Efendi (2012) dalam kasus klasifikasi Dokumen [2], Amalia (2015) dalam kasus Klasifikasi Ketepatan Waktu Lulus Mahasiswa [3], dan Karyadiputra (2016) dalam kasus Klasifikasi Status Kesejahteraan Rumah Tangga Keluarga Binaan Sosial [4].

Paper ini mengusulkan penggunaan model *Naïve Bayes Classifier* untuk klasifikasi Unit Simpan Pinjam Koperasi berdasarkan sehat tidaknya.

## 2. Metode Penelitian

### 2.1. *Naïve Bayes Classifier*

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. Klasifikasi *Naive Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya[5].

### 2.2. Kebutuhan Sistem

Penggunaan *Naïve Bayes Classifier* hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian.

Karena yang diasumsikan sebagai *variable independent*, maka hanya *varians* dari suatu *variable* dalam sebuah kelas yang dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan dari *matriks kovarians*.

Dalam klasifikasi sehat USP Koperasi kelas dibagi menjadi sehat, cukup sehat, dan kurang sehat. Dan atribut untuk setiap kelas ada 7, yaitu permodalan, kualitas aktivitas produktif, manajemen, efisiensi, likuiditas, kemandirian dan pertumbuhan, dan jati diri koperasi.

Untuk data pelatihan (*training data*) menggunakan data penilaian USP Koperasi tahun-tahun sebelumnya.

Persamaan dari teorema Bayes adalah:

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Keterangan:

X : Data dengan class yang belum diketahui

H : Hipotesis data X merupakan suatu kelas spesifik

- P(H|X) : Probabilitas hipotesis H berdasar kondisi X (*posteriori-probability*)
- P(H) : Probabilitas hipotesis H (*prior probability*)
- P(X|H) : Probabilitas X berdasarkan kondisi pada hipotesis H
- P(X) : Probabilitas X

Untuk menjelaskan teorema *Naive Bayes*, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema *Bayes* di atas disesuaikan sebagai berikut:

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)} \tag{2}$$

Dimana Variabel C merepresentasikan kelas, sementara variabel F1 ... Fn merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (*Posterior*) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik karakteristik sampel pada kelas C (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga *evidence*). Karena itu, rumus diatas dapat pula ditulis secara sederhana sebagai berikut:

$$Posterior = \frac{Prior \times likelihood}{evidence} \tag{3}$$

Nilai *Evidence* selalu tetap untuk setiap kelas pada satu sampel. Nilai dari *posterior* tersebut nantinya akan dibandingkan dengan nilai-nilai *posterior* kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut rumus *Bayes* tersebut dilakukan dengan menjabarkan  $(C|F_1, \dots, F_n)$  menggunakan aturan perkalian sebagai berikut:

$$\begin{aligned} P(C|F_1, \dots, F_n) &= P(C) P(F_1, \dots, F_n|C) \\ &= P(C) P(F_1|C) P(F_2, \dots, F_n|C, F_1) \\ &= P(C) P(F_1|C) P(F_2|C, F_1) P(F_3, \dots, F_n|C, F_1, F_2) \\ &= P(C) P(F_1|C) P(F_2|C, F_1) P(F_3|C, F_1, F_2) \dots P(F_n|C, F_1, F_2, F_3, \dots, F_{n-1}) \end{aligned} \tag{4}$$

Dapat dilihat bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor-faktor syarat yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk dianalisa satu persatu. Akibatnya, perhitungan tersebut menjadi sulit untuk dilakukan. Disinilah digunakan asumsi independensi yang sangat tinggi (*naif*), bahwa masing-masing petunjuk ( $F_1, F_2 \dots F_n$ ) saling bebas (independen) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan sebagai berikut:

$$P(P_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i) \tag{5}$$

Untuk  $i \neq j$ , sehingga

$$P(F_i|C, F_j) = P(F_i|C) \tag{6}$$

Persamaan di atas merupakan model dari teorema *Naive Bayes* yang selanjutnya akan digunakan dalam proses klasifikasi. Untuk klasifikasi dengan data kontinyu digunakan rumus *Densitas Gauss*:

$$P(X_i = x_i|Y = y_i) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} e^{-\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}} \tag{7}$$

- Dimana:
- P : peluang
  - Xi : atribut ke i
  - xi : nilai atribut ke i

- Y : kelas yang dicari  
 $y_i$  : sub kelas Y yang dicari  
 $\mu$  : *mean*, menyatakan rata-rata dari seluruh atribut  
 $\sigma$  : standar deviasi, menyatakan varian dari seluruh atribut

Adapun alur dari metode *Naive Bayes* adalah sebagai berikut [6]:

1. Baca data *training*
2. Hitung Jumlah dan probabilitas, namun apabila data numerik maka:
  - a. Cari nilai *mean* dan standar deviasi dari masing-masing parameter yang merupakan data numerik. Adapun persamaan yang digunakan untuk menghitung nilai rata-rata hitung (*mean*) dapat dilihat sebagai berikut:

$$\mu = \frac{\sum_{i=1}^n x_i}{n} \quad (8)$$

atau

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \quad (9)$$

dimana:

- $\mu$  : rata-rata hitung (*mean*)  
 $x_i$  : nilai sampel ke-i  
 n : jumlah sampel

dan persamaan untuk menghitung nilai simpangan baku (standar deviasi) dapat dilihat sebagai berikut:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}} \quad (10)$$

dimana:

- $\sigma$  : standar deviasi  
 $x_i$  : nilai x ke-i  
 $\mu$  : rata-rata hitung (*mean*)  
 n : jumlah sampel

- b. Cari nilai probabilitik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
3. Mendapatkan nilai dalam tabel mean, standar deviasi dan probabilitas.
4. Solusi kemudian dihasilkan.

### 3. Hasil Dan Pembahasan

#### 3.1. Hasil

Gambar 1 *Form Penilaian*

*Form* ini digunakan untuk melakukan penilaian usp koperasi. Hasil penilaian diklasifikasikan menjadi kurang sehat, cukup sehat, atau sehat.

Gambar 2 *Form* Penilaian Permodalan

*Form* ini digunakan untuk menghitung skor permodalan.

Gambar 3 *Form* Penilaian Aktiva

*Form* ini digunakan untuk menghitung skor kualitas aktiva produktif.

Gambar 4 *Form* Penilaian Manajemen

*Form* ini digunakan untuk menghitung skor manajemen.

Gambar 5 *Form Penilaian Efisiensi*

*Form* ini digunakan untuk menghitung skor efisiensi.

Gambar 6 *Form Penilaian Likuiditas*

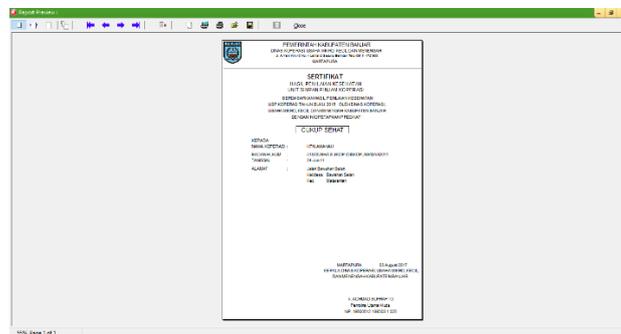
*Form* ini digunakan untuk menghitung skor likuiditas.

Gambar 7 *Form Penilaian Kemandirian dan Pertumbuhan*

*Form* ini digunakan untuk menghitung skor kemandirian dan pertumbuhan.

Gambar 8 Form Penilaian Jati Diri Koperasi

Form ini digunakan untuk menghitung skor jati diri.



Gambar 9 Sertifikat

Form ini digunakan untuk menampilkan sertifikat hasil penilaian.

NO	USP KOPERASI	NOMOR SAKSI (USP)	WILAYAH	KABUPATEN	PERINGKAT
1	KOPKASINDO KAMPULAN	00201001-001-001	Banjarnegara	Kab. Banjarnegara	Kategori II
2	KOPKASINDO KAMPULAN	00201001-001-002	Banjarnegara	Kab. Banjarnegara	Kategori II
3	KOPKASINDO KAMPULAN	00201001-001-003	Banjarnegara	Kab. Banjarnegara	Kategori II
4	KOPKASINDO KAMPULAN	00201001-001-004	Banjarnegara	Kab. Banjarnegara	Kategori II
5	KOPKASINDO KAMPULAN	00201001-001-005	Banjarnegara	Kab. Banjarnegara	Kategori II
6	KOPKASINDO KAMPULAN	00201001-001-006	Banjarnegara	Kab. Banjarnegara	Kategori II
7	KOPKASINDO KAMPULAN	00201001-001-007	Banjarnegara	Kab. Banjarnegara	Kategori II
8	KOPKASINDO KAMPULAN	00201001-001-008	Banjarnegara	Kab. Banjarnegara	Kategori II

Gambar 10 Rekapitulasi

Form ini digunakan untuk melihat rekapitulasi penilaian usp koperasi satu periode.

### 3.2. Pembahasan

Untuk menilai kinerja sistem, dilakukan pengujian terhadap beberapa data. Tabel 1 menunjukkan hasil pengujian untuk semua data.

Tabel 1 Hasil Pengujian

Data ke-	Hasil		Sesuai
	Data riil	Hasil Naïve Bayes	
1	Kurang Sehat	Kurang Sehat	Ya
2	Kurang Sehat	Kurang Sehat	Ya
3	Kurang Sehat	Kurang Sehat	Ya
4	Kurang Sehat	Kurang Sehat	Ya
5	Kurang Sehat	Kurang Sehat	Ya
6	Kurang Sehat	Kurang Sehat	Ya
7	Kurang Sehat	Kurang Sehat	Ya
8	Kurang Sehat	Cukup Sehat	Tidak
9	Kurang Sehat	Kurang Sehat	Ya
10	Kurang Sehat	Kurang Sehat	Ya
11	Kurang Sehat	Kurang Sehat	Ya
12	Kurang Sehat	Kurang Sehat	Ya
13	Kurang Sehat	Kurang Sehat	Ya
14	Kurang Sehat	Cukup Sehat	Tidak
15	Kurang Sehat	Cukup Sehat	Tidak
16	Cukup Sehat	Kurang Sehat	Tidak
17	Cukup Sehat	Cukup Sehat	Ya
18	Cukup Sehat	Kurang Sehat	Tidak
19	Cukup Sehat	Cukup Sehat	Ya
20	Cukup Sehat	Cukup Sehat	Ya
21	Cukup Sehat	Cukup Sehat	Ya
22	Cukup Sehat	Cukup Sehat	Ya
23	Cukup Sehat	Cukup Sehat	Ya
24	Cukup Sehat	Cukup Sehat	Ya
25	Cukup Sehat	Cukup Sehat	Ya
26	Cukup Sehat	Cukup Sehat	Ya
27	Cukup Sehat	Cukup Sehat	Ya
28	Cukup Sehat	Cukup Sehat	Ya
29	Cukup Sehat	Cukup Sehat	Ya
30	Cukup Sehat	Cukup Sehat	Ya
31	Cukup Sehat	Cukup Sehat	Ya
32	Cukup Sehat	Cukup Sehat	Ya
33	Cukup Sehat	Cukup Sehat	Ya
34	Cukup Sehat	Cukup Sehat	Ya
35	Cukup Sehat	Cukup Sehat	Ya
36	Cukup Sehat	Cukup Sehat	Ya
37	Cukup Sehat	Cukup Sehat	Ya
38	Cukup Sehat	Cukup Sehat	Ya
39	Cukup Sehat	Cukup Sehat	Ya
40	Cukup Sehat	Cukup Sehat	Ya
41	Cukup Sehat	Cukup Sehat	Ya
42	Cukup Sehat	Cukup Sehat	Ya
43	Cukup Sehat	Cukup Sehat	Ya
44	Cukup Sehat	Cukup Sehat	Ya
45	Cukup Sehat	Cukup Sehat	Ya
46	Cukup Sehat	Cukup Sehat	Ya
47	Sehat	Sehat	Ya
48	Sehat	Sehat	Ya
49	Sehat	Sehat	Ya
50	Sehat	Sehat	Ya

Berdasarkan hasil pengujian terhadap semua data, diperoleh hasil bahwa terdapat 45 yang sesuai dengan kelas yang diberikan dan 5 hasil yang tidak sesuai dengan hasil yang

diberikan. Pada tabel 2 terlihat ke-*valid*-an dari setiap kelas, proses penilaian kinerja dilakukan dengan konsep sensitivitas, kekhususan, nilai benar (FPV), dan nilai salah (FNV). Sensitivitas (TPV) adalah rasio USP Koperasi yang mendapat predikat X dan model klasifikasi juga memutuskan USP Koperasi tersebut mendapat predikat X dibagi dengan USP Koperasi yang memiliki predikat X. Kekhususan (TNV) adalah rasio USP Koperasi yang tidak mendapat predikat X dan model klasifikasi juga memutuskan USP Koperasi tersebut tidak mendapat predikat X dibagi dengan USP Koperasi bukan predikat X. FPV adalah rasio USP Koperasi yang memiliki predikat X, namun model klasifikasi memutuskan USP Koperasi tersebut tidak mendapat predikat X dibagi dengan jumlah keduanya. FNV adalah rasio USP Koperasi bukan predikat X, namun model klasifikasi memutuskan USP Koperasi tersebut mendapat predikat X dibagi dengan jumlah keduanya. Kinerja sistem dapat dihitung dengan dengan formula berikut:

$$\text{Kinerja} = (\text{TP}) + (\text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (11)$$

Tabel 2 Kinerja Setiap Kelas

Predikat	TP (TPV)	TN (TNV)	FP (FPV)	FN (FNV)	Kinerja
Kurang Sehat	12 (12/15)	33 (33/35)	3 (3/15)	2 (2/35)	45/50
Cukup Sehat	29 (29/31)	16 (16/19)	2 (2/31)	3 (3/19)	45/50
Sehat	4 (4/4)	46 (46/46)	0 (0/4)	0 (0/46)	50/50

Apabila diambil nilai total kinerja untuk semua kelas/predikat, akan diperoleh nilai total kinerja sebesar 0,93. Hasil tersebut menunjukkan bahwa nilai cukup bagus sehingga dapat disimpulkan bahwa model sistem telah memiliki kinerja yang baik.

#### 4. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa algoritma *Naïve Bayes Classifier* dapat digunakan sebagai salah satu metode untuk klasifikasi sehat USP Koperasi. Model sistem yang dibangun memiliki kinerja yang baik karena hasil pengujian menunjukkan total kinerja sebesar 0,93 atau 93%.

**Referensi**

- [1] Sembiring, M. (2016). Peraturan Menteri dan Peraturan Deputi Tentang Pengawasan Koperasi. Jakarta: Deputi Bidang Pengawasan.
- [2] Karyadiputra, E. (2016). Analisis Algoritma Naive Bayes Untuk Klasifikasi Status Kesejahteraan Rumah Tangga Keluarga Binaan Sosial. *Technologia: Jurnal Ilmiah*, 7(4). 199-208
- [3] Efendi, R., & MilaSari, J. (2012). Klasifikasi dokumen berbahasa indonesia Menggunakan naive bayes classifier. *Journal Research Computer Science & Application*, 1(1).7-13
- [4] Amalia, N., Shaufiah, S., & Sa'adah, S. (2015). Penerapan Teknik Data Mining untuk Klasifikasi Ketepatan Waktu Lulus Mahasiswa Teknik Informatika Universitas Telkom Menggunakan Algoritma Naive Bayes Classifier. *eProceedings of Engineering*, 2(3):1-11
- [5] Saleh, A. (2015). Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *Citec Journal*, 2, 207-2014.
- [6] Bustami. (2013). Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi. *TECHSI : Jurnal Penelitian Teknik Informatika*, 3, 128-132.