

# IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK MENGLASIFIKASI DATA NASABAH ASURANSI (Studi Kasus : Asuransi Prudential Cabang Jember)

Laras Dwi Mawarni

*Program Studi Teknik Informatika – Universitas Muhammadiyah Jember*

*Email: larasmawarni@gmail.com*

## Abstrak

Dalam bidang asuransi terkhusus asuransi Ayah, Ibu, Istri maupun Suami dan Anak, memprediksi suatu kejadian resiko tinggi pada keluarga sehingga kemunculan resiko secara dini bias ditanggulangi. Model pendekatan Bayesian berupa Klasifikasi Naïve Bayes dipakai memprediksi pembayaran premi yang akan dialami oleh nasabah dengan karakteristik Jenis Kelamin, Usia, Status Perkawinan, Tanggungan, Pekerjaan, Penghasilan dan Cara Pembayaran Premi. Semua data didiskritkan berdasar batasan yang dipakai Prudential dan hasil prediksi berupa probabilitas terjadinya resiko, bisa dipakai sebagai kelayakan melakukan asuransi.

Kata Kunci : *premi, asuransi, klasifikasi, algoritma Naive Bayes*

## 1. Latar Belakang

Premi merupakan pendapatan bagi perusahaan asuransi, yang jumlahnya ditentukan dalam suatu persentase atau tarif tertentu dari jumlah yang dipertanggungjawabkan. Bagi tertanggung premi merupakan beban karena membayar premi merupakan beban tertanggung. Pendapatan premi untuk perusahaan asuransi ditentukan oleh jumlah premi yang dibayar oleh nasabah.

Permasalahan yang sering timbul dalam perusahaan asuransi adalah banyaknya nasabah yang menunggak dalam membayar premi, sehingga diperlukan sebuah sistem yang dapat mengklasifikasikan nasabah mana yang masuk ke dalam kelompok lancar, kelompok kurang lancar dan nasabah mana yang masuk ke dalam kelompok tidak lancar dalam membayar iuran premi. Sehingga pihak asuransi bisa mengatasi sejak dini permasalahan tersebut.

Sebuah perusahaan asuransi pastilah mempunyai data yang begitu besar. Banyak yang belum menyadari bahwa dari pengolahan data-data tersebut dapat memberikan informasi berupa klasifikasi

data nasabah yang akan bergabung pada perusahaan itu sendiri. Penggunaan teknik data *mining* diharapkan mampu memberikan informasi yang berguna tentang teknik klasifikasi data nasabah yang akan bergabung dalam kelompok lancar, kelompok kurang lancar atau tidak lancar dalam membayar premi.

PT. Prudential Life Assurance Indonesia merupakan bagian dari Prudential plc, sebuah group perusahaan jasa keuangan terkemuka dari Inggris yang mengelola dana sebesar lebih dari US\$ 510 miliar dan melayani lebih dari 21 juta nasabah di seluruh dunia. Perusahaan ini juga yang pertama kali meluncurkan produk asuransi yang dikaitkan dengan investasi (*unit link*) di Indonesia, yaitu tahun 1999. Selain itu, Prudential Indonesia juga menyediakan berbagai produk yang di rancang untuk memenuhi dan melengkapi setiap kebutuhan para nasabahnya di Indonesia. Adapun produk dari Prudential Indonesia meliputi PRU*link fixed pay*, PRU*life*, PRU*life for juveniles*, PRU*major medical*, PRU*accident plus*, PRU*protector plan*, PRU*med*, PRU*link assurance account*

plus, PRUlink investor account, dan PRUlink syariah assurance account.

## 2. Tinjauan Pustaka

### 2.1. Naïve Bayes Classifier

#### 2.1.1. Teorema Bayes

*Bayes* merupakan teknik prediksi berbasis probabilistik sederhana yang berdasar pada penerapan teorema *Bayes* (atau aturan *Bayes*) dengan asumsi independensi (ketidaktergantungan) yang kuat (naïf). Dengan kata lain, *Naïve Bayes*, model yang digunakan adalah “model fitur independen”.

Dalam *Bayes* (terutama *Naïve Bayes*), maksud independensi yang kuat pada fitur adalah bahwa sebuah fitur pada sebuah data tidak berkaitan dengan ada atau tidaknya fitur lain dalam data yang sama.

Prediksi Bayes didasarkan pada teorema Bayes dengan formula umum sebagai berikut :

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)} \dots \dots \dots (3)$$

Penjelasan dari formula (3) tersebut adalah sebagai berikut :

Parameter	Keterangan
P(H E)	Probabilitas akhir bersyarat ( <i>conditional probability</i> ) suatu hipotesis H terjadi jika diberikan bukti ( <i>evidence</i> ) E terjadi.
P(E H)	Probabilitas sebuah bukti E terjadi akan memengaruhi hipotesis H.
P(H)	Probabilitas awal (priori) hipotesis H terjadi tanpa memandang bukti apapun.
P(E)	Probabilitas awal (priori) bukti E terjadi tanpa memandang hipotesis/bukti yang lain.

Ide dasar dari aturan *Bayes* adalah bahwa hasil dari hipotesis atau

peristiwa (H) dapat diperkirakan berdasarkan pada beberapa bukti (E) yang diamati. Ada beberapa hal penting dari aturan *Bayes* tersebut, yaitu :

1. Sebuah probabilitas awal/prior H atau P(H) adalah probabilitas dari suatu hipotesis sebelum bukti diamati.
2. Sebuah probabilitas akhir H atau P(H|E) adalah probabilitas dari suatu hipotesis setelah bukti diamati.

#### 2.1.2. *Naïve Bayes* Untuk Klasifikasi

Kaitan antara *Naïve Bayes* dengan klasifikasi, korelasi hipotesis dan bukti klasifikasi adalah bahwa hipotesis dalam teorema *Bayes* merupakan label kelas yang menjadi target pemetaan dalam klasifikasi, sedangkan bukti merupakan fitur-fitur yang menjadikan masukkan dalam model klasifikasi. Jika X adalah vektor masukkan yang berisi fitur dan Y adalah label kelas, *Naïve Bayes* dituliskan dengan P(X|Y). Notasi tersebut berarti probabilitas label kelas Y didapatkan setelah fitur-fitur X diamati. Notasi ini disebut juga probabilitas akhir (*posterior probability*) untuk Y, sedangkan P(Y) disebut probabilitas awal (*prior probability*) Y.

Selama proses pelatihan harus dilakukan pembelajaran probabilitas akhir P(Y|X) pada model untuk setiap kombinasi X dan Y berdasarkan informasi yang didapat dari data latih. Dengan membangun model tersebut, suatu data uji X' dapat diklasifikasikan dengan mencari nilai Y' dengan memaksimalkan nilai P(X'|Y') yang didapat.

Formulasi *Naïve Bayes* untuk klasifikasi adalah :

$$P(Y|X) = \frac{P(Y) \sum_{i=1}^q P(X_i|Y)}{P(X)}$$

$P(Y|X)$  adalah probabilitas data dengan vektor  $X$  pada kelas  $Y$ .  $P(Y)$  adalah probabilitas awal kelas  $Y$ .  $\sum_{i=1}^q P(X_i|Y)$  adalah probabilitas independen kelas  $Y$  dari semua fitur dalam vektor  $X$ . Nilai  $P(X)$  selalu tetap sehingga dalam perhitungan prediksi nantinya kita tinggal menghitung bagian  $P(Y) \sum_{i=1}^q P(X_i|Y)$  dengan memilih yang terbesar sebagai kelas yang dipilih sebagai hasil prediksi. Sementara probabilitas independen  $\sum_{i=1}^q P(X_i|Y)$  tersebut merupakan pengaruh semua fitur dari data terhadap setiap kelas  $Y$ , yang dinotasikan dengan :

$$P(X|Y = y) = \sum_{i=1}^q P(X_i|Y = y)$$

Setiap set fitur  $X = \{X_1, X_2, X_3, \dots, X_q\}$  terdiri atas  $q$  atribut ( $q$  dimensi).

Umumnya, *Bayes* mudah dihitung untuk fitur bertipe kategoris seperti pada kasus klasifikasi hewan dengan fitur “penutup kulit dengan nilai {bulu, rambut, cangkang} atau kasus fitur “jenis kelamin” dengan nilai {pria, wanita}. Namun untuk fitur dengan tipe *numerik (kontinu)* ada perlakuan khusus sebelum dimasukkan dalam *Naïve Bayes*. Caranya adalah :

1. Melakukan diskretisasi pada setiap fitur kontinu dan mengganti nilai fitur kontinu tersebut dengan nilai interval diskret. Pendekatan ini dilakukan dengan mentransformasikan fitur kontinu ke dalam fitur ordinal.
2. Mengasumsikan bentuk tertentu dari distribusi probabilitas untuk fitur kontinu dan memperkirakan parameter distribusi dengan data pelatihan. Distribusi Gaussian biasanya dipilih

untuk merepresentasikan probabilitas bersyarat dari fitur kontinu pada sebuah kelas  $P(X_i|Y)$ , sedangkan distribusi Gaussian dikarakteristikan dengan dua parameter : mean,  $\mu$  dan varian,  $\sigma^2$ . Untuk setiap kelas  $y_j$ , probabilitas bersyarat kelas  $y_j$  untuk fitur  $X_i$  adalah :

$$P(X_i = x_i | Y = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} \exp \left( -\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2} \right)$$

Parameter  $\mu_{ij}$  bisa didapat dari *mean* sampel  $X_i$  ( $\bar{x}$ ) dari semua data latih yang menjadi milik kelas  $y_j$ , sedangkan  $\sigma_{ij}^2$  dapat diperkirakan dari *varian* sampel ( $s^2$ ) dari data latih.

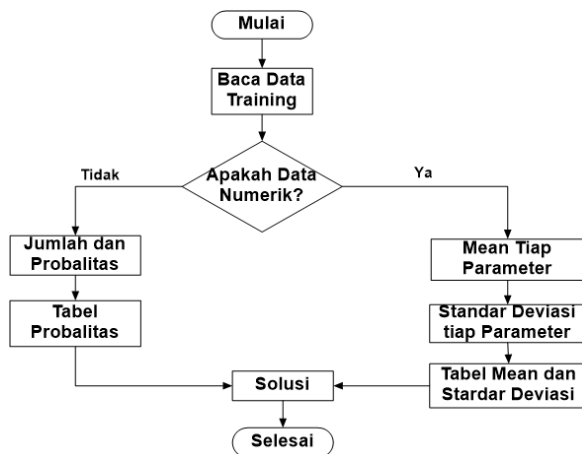
### 2.1.3. Karakteristik *Naïve Bayes*

Klasifikasi dengan *Naïve Bayes* bekerja berdasarkan teori probabilitas yang memandang semua fitur dari data sebagai bukti dalam probabilitas. Hal ini memberikan karakteristik *Naïve Bayes* sebagai berikut :

1. Metode *Naïve Bayes* bekerja teguh (*robust*) terhadap data-data yang terisolasi yang biasanya merupakan data dengan karakteristik berbeda (*outliner*). *Naïve Bayes* juga bisa menangani nilai atribut yang salah dengan mengabaikan data latih selama proses pembangunan model dan prediksi.
2. Tangguh menghadapi atribut yang tidak relevan.
3. Atribut yang mempunyai korelasi bisa mendegradasi kinerja klasifikasi *Naïve Bayes* karena asumsi independensi atribut tersebut sudah tidak ada.

### 3. Metodologi Penelitian

#### 3.1. Skema Naïve Bayes



**Gambar 3.1** Skema *Naïve Bayes*

*Teorema Bayes* adalah teorema yang digunakan dalam statistika untuk menghitung peluang untuk suatu hipotesis. *Bayes Optimal Classifier* menghitung peluang dari suatu kelas dari masing-masing kelompok atribut yang ada, dan menentukan kelas mana yang paling optimal. Pada penelitian ini penulis tidak menggunakan data numerik, data yang digunakan data nasabah Prudential Asuransi Jember.

### 4. Implementasi Sistem

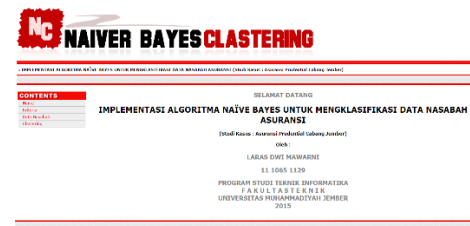
#### 4.1. Implementasi Sistem

Setelah melalui tahapan perancangan sistem, database selanjutnya adalah implementasi sistem. Implementasi sistem merupakan bagian akhir daripada perancangan sistem yang telah dibangun dimana tahapan ini juga merupakan testing program.

##### 4.1.1. Halaman Menu Utama

Halaman menu utama berfungsi untuk mengakses segala perintah yang terdapat dalam aplikasi. Form Utama tersebut adalah halaman yang pertama kali di jumpai ketika program dijalankan, Pada form ini terdapat beberapa menu yaitu Home yang berisi halaman utama, Kriteria, Data

Nasabah dan Klasifikasi. Adapun tampilannya dapat dilihat pada gambar berikut :



**Gambar 4.1** Halaman Utama

Pada halaman utama berisi informasi nama sistem dan penyusun sistem “Implementasi Algoritma Naïve Bayes Untuk Mengklasifikasi Data Nasabah Asuransi Pada Asuransi Prudential Cabang Jember”, nama fakultas dan program studi penyusun dan tahun perancangan sistem.

#### 4.1.2. Halaman Kriteria dan Sub Kriteria

Halaman kriteria dan sub kriteria disini digunakan untuk data kriteria dan sub kriteria yang nantinya akan menjadi score dari setiap nasabah kriteria penilai meliputi “Jenis Kelamin, Umur Nasabah, Status Perkawinan, Pekerjaan, Penghasilan, Cara Pembayaran Premi, Masa Asuransi dan Klasifikasi” dari kriteria tersebut terdapat sub kriteria yang nantinya akan menjadi penilaian dari nasabah, untuk sub kriteria bersifat fleksibel yang artinya dapat di rubah sesuai kebutuhan dalam mengklasifikasikan nadabah. Adapun tampilannya dapat dilihat pada gambar berikut :

### Kriteria Pembobotan

Kode Kriteria	Kriteria	Operasi
1 JK	Jenis Kelamin	~
1.1	Laki-Laki	~
1.2	Perempuan	~
2 Age	Umur Nasabah	[ Tambah Sub.Kriteria ]
2.1	20 - 29 Tahun	[ Edit ] [ Delete ]
2.2	30 - 40 Tahun	[ Edit ] [ Delete ]
2.3	>40 Tahun	[ Edit ] [ Delete ]
3 SK	Status Perkawinan	[ Tambah Sub.Kriteria ]
3.1	Belum Kawin	[ Edit ] [ Delete ]
3.2	Kawin	[ Edit ] [ Delete ]
3.3	Janda atau Duda	[ Edit ] [ Delete ]
4 PK	Pekerjaan	[ Tambah Sub.Kriteria ]
4.1	PNS	[ Edit ] [ Delete ]
4.2	Wiraswasta	[ Edit ] [ Delete ]
4.3	Pegawai Swasta	[ Edit ] [ Delete ]
5 PH	Penghasilan	[ Tambah Sub.Kriteria ]
5.1	< 25 Juta	[ Edit ] [ Delete ]
5.2	25 - 50 Juta	[ Edit ] [ Delete ]
5.3	> 50 Juta	[ Edit ] [ Delete ]

**Gambar 4.2** Daftar Kriteria dan Sub Kriteria

Pada gambar diatas terdapat menu tambah sub kriteria yang fungsi untuk menambahkan sub kriteria baru dan menu edit untuk melakukan perubahan pada sub kriteri sedangkan menu delete untuk menghapus sub kriteria. Untuk tampilan form tambah atau form edit sub kriteria seperti gambar 4.3 dibawah ini :

#### Formulir Sub.Kriteria Pembobotan

Kriteria	Umur Nasabah
Sub Kriteria	<input type="text"/>
<input type="button" value="Simpan"/> <input type="button" value="Kembali"/>	

**Gambar 4.3** Form Sub Kriteria

Pada halaman form tambah atau edit sub keriteri terdapat menu simpan dan kembali menu tersebut memiliki fungsi yang berbeda pada menu simpan disini untuk menyimpan penambah sub kriteria atau update perubahan pada sub kriteria dan sedangkan pada menu kembali disini digunakan untuk kembali daftar kriteria dan sub kriteria yang arti membatalkan penambahan atau update sub kriteria.

### 4.1.3. Halaman Data Nasabah

Pada halaman data nasabah adalah data nasabah training maksudnya adalah data yang menjadi acuan pengklasifikasian atau data maining data maining yang dijadikan data training berjumlah 20 data nasabah dari data set yang berjumlah 54 nasabah yang diambil secara acak. Pada halaman data nasabah terdapat dua menu yaitu tambah dan edit nasabah. Untuk tampilan halaman data nasabah dapat dilihat pada gambar dibawah ini:

No.	Nama Nasabah	Jenis Kelamin	Umur Nasabah	Status Perkawinan	Pekerjaan	Penghasilan	Cara Pembayaran Premi	Masa Asuransi	Klasifikasi
1	Dani Lulman	Laki-Laki	30 - 40 Tahun	Kawin	PNS	< 25 Juta	Tahunan	> 15 Tahun	Tidak Lancar
2	Zulakha	Perempuan	20 - 29 Tahun	Kawin	Wiraswasta	25 - 50 Juta	Triwulan	11-15 Tahun	Tidak Lancar
3	Culhem	Laki-Laki	20 - 29 Tahun	Kawin	PNS	< 25 Juta	Tahunan	11-15 Tahun	Kurang Lancar
4	Hidayatullah	Laki-Laki	30 - 40 Tahun	Belum Kawin	Wiraswasta	< 25 Juta	Tahunan	11-15 Tahun	Lancar
5	Ade Gunawan	Laki-Laki	30 - 40 Tahun	Kawin	PNS	< 25 Juta	Semesteran	11-15 Tahun	Tidak Lancar
6	Fuadhi	Perempuan	20 - 29 Tahun	Belum Kawin	PNS	< 25 Juta	Tahunan	11-15 Tahun	Lancar
7	Rozan Hadi	Laki-Laki	20 - 29 Tahun	Belum Kawin	PNS	< 25 Juta	Tahunan	11-15 Tahun	Lancar
8	Rahmat Saputra	Laki-Laki	30 - 40 Tahun	Belum Kawin	Wiraswasta	< 25 Juta	Semesteran	11-15 Tahun	Lancar
9	Rasyidi	Perempuan	20 - 29 Tahun	Kawin	Pegawai Swasta	< 25 Juta	Triwulan	5-10 Tahun	Tidak Lancar
10	Dina Sudha	Perempuan	30 - 40 Tahun	Belum Kawin	PNS	< 25 Juta	Triwulan	5-10 Tahun	Lancar
11	Wika Riki	Laki-Laki	30 - 40 Tahun	Kawin	Wiraswasta	< 25 Juta	Tahunan	5-10 Tahun	Kurang Lancar
12	Niam Sari	Perempuan	20 - 29 Tahun	Kawin	Wiraswasta	25 - 50 Juta	Tahunan	> 15 Tahun	Kurang Lancar
13	Nahar Indri	Laki-Laki	30 - 40 Tahun	Kawin	Wiraswasta	< 25 Juta	Triwulan	11-15 Tahun	Lancar
14	Vandier	Perempuan	>40 Tahun	Kawin	PNS	< 25 Juta	Semesteran	> 15 Tahun	Kurang Lancar
15	Ursento	Laki-Laki	20 - 40 Tahun	Belum Kawin	Wiraswasta	> 50 Juta	Semesteran	11-15 Tahun	Lancar
16	Tari Sulandari	Perempuan	20 - 40 Tahun	Kawin	Wiraswasta	< 25 Juta	Triwulan	11-15 Tahun	Lancar
17	Leni Suwastika	Perempuan	20 - 29 Tahun	Belum Kawin	Wiraswasta	25 - 50 Juta	Bulanan	5-10 Tahun	Tidak Lancar
18	Orlania	Perempuan	20 - 40 Tahun	Kawin	PNS	< 25 Juta	Semesteran	5-10 Tahun	Lancar
19	M. Saiful	Laki-Laki	>40 Tahun	Kawin	Pegawai Swasta	< 25 Juta	Tahunan	11-15 Tahun	Tidak Lancar
20	M. Irfan	Laki-Laki	30 - 40 Tahun	Kawin	Pegawai Swasta	25 - 50 Juta	Tahunan	11-15 Tahun	Tidak Lancar

**Gambar 4.4** Halaman Data Nasabah

#### Formulir Nasabah

Nomor KTP/ID	<input type="text"/>
Nama Nasabah	<input type="text"/>
Alamat	<input type="text"/>
Jenis Kelamin	<input type="text"/>
Umur Nasabah	<input type="text"/>
Status Perkawinan	<input type="text"/>
Pekerjaan	<input type="text"/>
Penghasilan	<input type="text"/>
Cara Pembayaran Premi	<input type="text"/>
Masa Asuransi	<input type="text"/>
Klasifikasi	<input type="text"/>
<input type="button" value="Simpan"/> <input type="button" value="Kembali"/>	

**Gambar 4.5** Halaman Formulir Nasabah

### 4.1.4. Halaman Klasifikasi Nasabah

Halaman klasifikasi adalah halaman untuk mengklasifikasi calon nasabah untuk memperoleh kelas nasabah, nantinya akan termasuk ke kelas lancar, kurang lancar atau tidak lancar. Untuk tampilan klasifikasi nasabah seperti berikut :

Klasifikasi Calon Nasabah

Nomor KTP/ID	<input type="text"/>
Nama Nasabah	<input type="text"/>
Alamat	<input type="text"/>
Jenis Kelamin	<input type="text"/>
Umur Nasabah	<input type="text"/>
Status Perkawinan	<input type="text"/>
Pekerjaan	<input type="text"/>
Penghasilan	<input type="text"/>
Cara Pembayaran Premi	<input type="text"/>
Masa Asuransi	<input type="text"/>
<input type="button" value="Hitung Clustering"/> <input type="button" value="Data Nasabah"/>	

**Gambar 4.6** Halaman Form  
Klasifikasi Nasabah

Pada gambar diatas user harus mengimput data nasabah dan nilai kriteria nasabah jika data sudah diisi maka user harus mengklik menu hitung clustering untuk memperoleh nasabah tersebut tergolong kekelas lancar, kurang lancar maupun tidak lancar. Untuk tampilan contoh perhitungan dan implementasi algoritma Naïve Bayes sebagai berikut :

Apabila diberikan input baru, maka klasifikasi data nasabah asuransi dapat ditentukan melalui langkah berikut:

1. Menghitung jumlah class / label

- $P(Y=Lancar)= \frac{10}{20}$   
(Jumlah data Lancar pada data penelitian dibagi dengan jumlah keseluruhan data)
- $P(Y=Kurang Lancar)= \frac{4}{20}$   
(Jumlah data Kurang Lancar pada data penelitian dibagi dengan jumlah keseluruhan data)
- $P(Y=Tidak Lancar)= \frac{7}{20}$   
(Jumlah data Tidak Lancar pada data penelitian dibagi dengan jumlah keseluruhan data)

2. Menghitung jumlah kasus yang sama dengan class yang sama

- $P(\text{Jenis Kelamin} = \text{Laki-Laki} \mid Y=Lancar) = 6/10$
- $P(\text{Jenis Kelamin} = \text{Laki-Laki} \mid Y=Kurang Lancar) = 2/4$

- $P(\text{Jenis Kelamin} = \text{Laki-Laki} \mid Y=Tidak Lancar) = 4/7$
- $P(\text{Umur Nasabah} = >40 \text{ Tahun} \mid Y=Lancar) = 0/10$
- $P(\text{Umur Nasabah} = >40 \text{ Tahun} \mid Y=Kurang Lancar) = 1/4$
- $P(\text{Umur Nasabah} = >40 \text{ Tahun} \mid Y=Tidak Lancar) = 1/7$
- $P(\text{Status Perkawinan} = \text{Kawin} \mid Y=Lancar) = 3/10$
- $P(\text{Status Perkawinan} = \text{Kawin} \mid Y=Kurang Lancar) = 4/4$
- $P(\text{Status Perkawinan} = \text{Kawin} \mid Y=Tidak Lancar) = 6/7$
- $P(\text{Pekerjaan} = \text{PNS} \mid Y=Lancar) = 5/10$
- $P(\text{Pekerjaan} = \text{PNS} \mid Y=Kurang Lancar) = 2/4$
- $P(\text{Pekerjaan} = \text{PNS} \mid Y=Tidak Lancar) = 2/7$
- $P(\text{Penghasilan} = < 25 \text{ Juta} \mid Y=Lancar) = 8/10$
- $P(\text{Penghasilan} = < 25 \text{ Juta} \mid Y=Kurang Lancar) = 2/4$
- $P(\text{Penghasilan} = < 25 \text{ Juta} \mid Y=Tidak Lancar) = 4/7$
- $P(\text{Cara Pembayaran Premi} = \text{Bulanan} \mid Y=Lancar) = 1/10$
- $P(\text{Cara Pembayaran Premi} = \text{Bulanan} \mid Y=Kurang Lancar) = 0/4$
- $P(\text{Cara Pembayaran Premi} = \text{Bulanan} \mid Y=Tidak Lancar) = 1/7$
- $P(\text{Masa Asuransi} = >15 \text{ Tahun} \mid Y=Lancar) = 0/10$
- $P(\text{Masa Asuransi} = >15 \text{ Tahun} \mid Y=Kurang Lancar) = 2/4$

- $P(\text{Masa Asuransi} = >15 \text{ Tahun} \mid Y = \text{Tidak Lancar}) = 1/7$
3. Kalikan semua hasil class Lancar, Kurang Lancar, Tidak Lancar

- $P(\text{Jenis Kelamin} = \text{Laki-Laki} \mid Y = \text{Lancar}) *$

$$P(\text{Umur Nasabah} = 30 - 40 \text{ Tahun} \mid Y = \text{Lancar}) *$$

$$P(\text{Status Perkawinan} = \text{Belum Kawin} \mid Y = \text{Lancar}) *$$

$$P(\text{Pekerjaan} = \text{Wiraswasta} \mid Y = \text{Lancar}) *$$

$$P(\text{Penghasilan} < 25 \text{ Juta} \mid Y = \text{Lancar}) *$$

$$P(\text{Cara Pembayaran Premi} = \text{Tahunan} \mid Y = \text{Lancar}) *$$

$$P(\text{Masa Asuransi} = 11-15 \text{ Tahun} \mid Y = \text{Lancar}) *$$

$$P(\text{Klasifikasi} = \text{Lancar} \mid Y = \text{Lancar}) =$$

$$6/10 \times 7/10 \times 7/10 \times 5/10 \times 8/10 \times 3/10 \times 7/10 \times 10/10 = 0,0247$$

- $P(\text{Jenis Kelamin} = \text{Laki-Laki} \mid Y = \text{Kurang Lancar}) *$

$$P(\text{Umur Nasabah} = 20 - 29 \text{ Tahun} \mid Y = \text{Kurang Lancar}) *$$

$$P(\text{Status Perkawinan} = \text{Kawin} \mid Y = \text{Kurang Lancar}) *$$

$$P(\text{Pekerjaan} = \text{PNS} \mid Y = \text{Kurang Lancar}) *$$

$$P(\text{Penghasilan} = 25 - 50 \text{ Juta} \mid Y = \text{Kurang Lancar}) *$$

$$P(\text{Cara Pembayaran Premi} = \text{Triwulan} \mid Y = \text{Kurang Lancar}) *$$

$$P(\text{Masa Asuransi} = 11-15 \text{ Tahun} \mid Y = \text{Kurang Lancar}) *$$

$$P(\text{Klasifikasi} = \text{Kurang Lancar} \mid Y = \text{Kurang Lancar}) =$$

$$2/4 \times 1/4 \times 4/4 \times 2/4 \times 2/4 \times 1/4 \times 1/4 \times 4/4 = 0,0020$$

- $P(\text{Jenis Kelamin} = \text{Laki-Laki} \mid Y = \text{Tidak Lancar}) *$

$$P(\text{Umur Nasabah} = 30 - 40 \text{ Tahun} \mid Y = \text{Tidak Lancar}) *$$

$$P(\text{Status Perkawinan} = \text{Kawin} \mid Y = \text{Tidak Lancar}) *$$

$$P(\text{Pekerjaan} = \text{PNS} \mid Y = \text{Tidak Lancar}) *$$

$$P(\text{Penghasilan} < 25 \text{ Juta} \mid Y = \text{Tidak Lancar}) *$$

$$P(\text{Cara Pembayaran Premi} = \text{Tahunan} \mid Y = \text{Tidak Lancar}) *$$

$$P(\text{Masa Asuransi} > 15 \text{ Tahun} \mid Y = \text{Tidak Lancar}) *$$

$$P(\text{Klasifikasi} = \text{Tidak Lancar} \mid Y = \text{Tidak Lancar}) =$$

$$4/7 \times 3/7 \times 6/7 \times 2/7 \times 4/7 \times 3/7 \times 1/7 \times 7/7 = 0,0021$$

Bandingkan hasil class Lancar, Kurang Lancar dan Tidak Lancar Dari hasil diatas, terlihat bahwa nilai probabilitas tertinggi ada pada kelas  $P(\text{Lancar})$  sehingga dapat disimpulkan bahwa status calon nasabah tersebut termasuk ke golongan class **“Lancar”** dengan nilai probabilitas : **0,0247**

## 5. Kesimpulan Dan Saran

### 5.1. Kesimpulan

Berdasarkan hasil pembahasan maka penulis dapat mengambil beberapa kesimpulan antara lain :

1. Sistem klasifikasi data nasabah ini digunakan untuk menampilkan informasi klasifikasi

lancar, kurang lancar atau tidak lancarnya calon nasabah dalam membayar premi asuransi dengan menggunakan algoritma Naive Bayes.

2. Dengan adanya sistem ini maka mempermudah pihak asuransi dalam memperkirakan nasabah yang bergabung, sehingga perusahaan bisa mengambil keputusan untuk menerima atau menolak calon nasabah tersebut.

3. Algoritma Naive Bayes didukung oleh ilmu Probabilistik dan ilmu statistika khususnya dalam penggunaan data petunjuk untuk mendukung keputusan pengklasifikasian. Pada algoritma Naive Bayes, semua atribut akan memberikan kontribusinya dalam pengambilan keputusan, dengan bobot atribut yang sama penting dan setiap atribut saling bebas satu sama lain.

4. Variabel penentu yang digunakan dalam penelitian ini adalah jenis kelamin, usia, status, pekerjaan, penghasilan per tahun, masa pembayaran asuransi, dan cara pembayaran asuransi.

## 5.2. Saran

Dikarenakan perkembangan teknologi sangat pesat saran dari penulis untuk pengembang berikut bisa

menggunakan perbandingan algoritma sehingga tidak hanya menggunakan algoritma *Naive Bayes* saja, pengembang berikut bias menggunakan algoritma Naive Bayes dibandingkan dengan algoritma K-Means, Decision Tree (ID3).

## DAFTAR PUSTAKA

- [1]. Rokhmah, Dewi Pyriana, 2011, Klasifikasi Data Menggunakan Metode K-Nearest Neighbour dan Teorema Bayes, Program Studi Teknik Informatika, Universitas Brawijaya Malang
- [2]. Susanto, Sani, Ph.D, Suryadi, Dedy, 2010, Pengantar Data Mining: Menggali Pengetahuan dari Bongkahan Data, Andi, Yogyakarta
- [3]. Hermawati, Fajar Astuti, 2013, Data Mining, Andi, Yogyakarta
- [4]. Kusumadewi, Sri, 2009, Klasifikasi Status Gizi Menggunakan Algoritma Naive Bayes Classification, Jurusan Teknik Informatika, universitas Islam Indonesia
- [5]. Bustami, 2010, Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi, Aceh Utara, Universitas Malikussaleh