ISSN Online : 2685-5615

SEMINAR NASIONAL TEKNOLOGI INFORMASI DAN KOMUNIKASI 2019

SENATIK

PROSIDING

SEMINAR NASIONAL TEKNOLOGI INFORMASI DAN KOMUNIKASI

"TEKNOLOGI HUMANIS DI ERA SOCIETY 5.0"

PROGRAM STUDI INFORMATIKA
UNIVERSITAS PGRI MADIUN
2019

SENATIK 2019
PROSIDING
SEMINAR NASIONAL TEKNOLOGI INFORMASI DAN KOMUNIKASI
"TEKNOLOGI HUMANIS DI ERA SOCIETY 5.0"

Klasifikasi Penentuan Pinjaman Nasabah di Bank XYZ Menggunan Naïve Bayes

Abdul Rozaq
Universitas PGRI Madiun
e-mail: zarozaq@gmail.com

Abstrak

Persaingan yang tinggi di bidang perbankan semakin ketat sehingga penting untuk mengklasifikasikannya berdasarkan jaminan pinjaman, jumlah pinjaman sampai perilaku nasabah sangat penting. *Machine Learning* menyediakan teknologi untuk menganalisis volume data yang besar atau mendeteksi data untuk mengubah data mentah menjadi informasi berharga. Salah satu metode klasisifikasi dalam *machine learning* adalah naive bayes. Dalam penelitian ini penulis menggunakan naive bayes untuk menentukan pinjaman di bank XYZ dengan atribut jenis pinjaman, agunan, pendapatan dan karaktek sedang kelasnya yaitu 1 artinya pinjaman di setujui dan 0 menandakan pinjaman tidak disetujui. Berdasarkan 16 data training dan data testing dengan atribut jenis pinjaman = KUR, agunan = sertifikat, pendapatan = cukup, karakter = kurang baik maka hasil klasifikasinya ke dalam kelas 1 yang artinya pinjaman di setujui. Hasil uji data dari 13 data training dan 3 data testing maka hasil nilai akurasi 93.75%.

Kata kunci : machine learning, klasifikasi, pinjaman bank

1. Pendahuluan

Persaingan yang tinggi di bidang perbankan semakin ketat sehingga penting untuk menganalisis data nasabah yang besar dan mengklasifikasikannya berdasarkan jaminan pinjaman, jumlah pinjaman sampai perilaku nasabah sangat penting. *Machine Learning* menyediakan teknologi untuk menganalisis volume data yang besar atau mendeteksi data untuk mengubah data mentah menjadi informasi berharga.

Machine learning telah menarik banyak perhatian di industri informasi dan di masyarakat secara keseluruhan dalam beberapa tahun terakhir, karena ketersediaan data yang sangat besar dan kebutuhan yang sangat cepat untuk mengubah data tersebut menjadi informasi dan pengetahuan yang berguna. Informasi dan pengetahuan yang diperoleh dapat digunakan untuk aplikasi mulai dari analisis pasar, deteksi penipuan, dan analisa data nasabah.

Customer service memasukkan data nasabah berupa besarnya pinjaman, agunan yang digunakan, jenis pinjaman, pekerjaan nasabah hingga histori usaha nasabah. Dari data tersebut machine learning akan menganalisa, apakah nasabah tersebut akan di setujui untuk meminjam uang di bank atau nasabah tersebut tidak di setujui untuk meminjam di bank yang bersangkutan.

Berdasarkan masalah ini, metode pengolahan data yang diperlukan, yaitu klasifikasi. Klasifikasi adalah salah satu metode untuk menemukan model atau fungsi yang menjelaskan dan membedakan konsep atau kelas data. Model ini diturunkan berdasarkan analisis dari serangkaian data pelatihan dan dapat digunakan untuk memprediksi label kelas dari suatu objek dengan label yang tidak diketahui.

2. Klasifikasi dan prediksi

Klasifikasi data adalah proses dua langkah. Langkah pertama adalah proses pembelajaran. Dalam hal ini data pelatihan dianalisis dengan algoritma klasifikasi, atribut label kelas adalah keputusan pinjaman. Proses kedua adalah klasifikasi untuk pengujian dimana data uji digunakan untuk memperkirakan keakuratan algoritma klasifikasi. Jika akurasi klasifikasi lebih baik dari itu algoritma dapat diterapkan pada klasifikasi data baru tuple.

2.1 Algoritma Naïve Bayes

Pengklasifikasi Bayesian adalah pengklasifikasi Statistik. Bayesian dapat memprediksi probabilitas keanggotaan kelas, seperti probabilitas bahwa tupel tertentu milik kelas tertentu. Klasifikasi Bayesian didasarkan pada teorema Bayes.

Pengklasifikasi Naïve Bayesian menganggap bahwa efek dari nilai atribut pada kelas tertentu tidak tergantung pada nilai atribut lainnya. Asumsi ini disebut kemandirian kelas bersyarat. Hal ini dibuat untuk menyederhanakan perhitungan yang terlibat dalam pengertian ini, dianggap " Naïve." Jaringan kepercayaan Bayesian adalah model grafis, yang tidak seperti pengklasifikasi naive Bayesian memungkinkan representasi dependensi di antara subset dari atribut. Jaringan kepercayaan Bayesian juga dapat digunakan untuk klasifikasi. Berikut persamaan yang digunakan dalam naive bayes.

$$P(H|X) = \frac{P(X|H) * P(H)}{P(X)}$$

X = Data dengan kelas yang belum diketahui
H = Hipotesis data merupakan suatu kelas spesifik
P (H | X) = Probabilitas hipotesis H berdasar kondisi X (posteori probabilitas)

P (H) = Probabilitas hipotesis H (prior probabilitas)
P (X | H) = Probabilitas X berdasarkan kondisi pada hipotesis H
P (X) =Probabilitas X
Proses klasifikasi untuk menentukan data yang termasuk ke dalam kelas yang paling cocok berdasarkan data testing. Pada metode naive bayes, semua atribut bersifat indepent di mana semua atribut sama penting. Apabila diberikan n atribut maka nilai probabilitas dapat dimasukkan kedalam persamaan berikut:

$$\text{likelihood}(c_i) = \arg\max_{c_i \epsilon C} P(c_i) \prod_{f=1}^{F} P(t_f|c_i)$$

3. Hasil dan Pembahasan

4.1 Data Training

Untuk menganalisa data dengan metode Naive Bayes maka langkah pertama yang dilakukan adalah membaca data latih. Data latih yang digunakan dalam penelitian ini terdapat pada tabel 1.

Tabel 1. Data Training

No	Jenis Pinjaman	Agunan	Pendapatan	Karakter	Label
1	Kur	Bpkb	Baik	Baik	1
2	Kur	Bpkb	Baik	Kurang Baik	1
3	Kur	Sertifikat	Baik	Tidak Baik	0
4	Kur	Bpkb	Cukup	Baik	1
5	Kur	Bpkb	Cukup	Kurang Baik	1
6	Kur	Sertifikat	Kurang	Baik	1
7	Swaguna	Mou	Baik	Baik	1
8	Swaguna	Mou	Kurang	Baik	0
9	Swaguna	Mou	Baik	Kurang Baik	1
10	Swaguna	Mou	Cukup	Kurang Baik	1
11	Mikro	Bpkb	Baik	Baik	1
12	Mikro	Bpkb	Baik	Kurang Baik	1
13	Mikro	Sertifikat	Baik	Tidak Baik	0
14	Mikro	Bpkb	Cukup	Baik	1
15	Mikro	Bpkb	Cukup	Kurang Baik	1
16	Mikro	Sertifikat	Kurang	Baik	1

Label 1 menunjukkan pinjaman nasabah di terima sedangkan label 0 menunjukkan pinjaman nasabah di tolak
Perhitungan data training naive bayes berdasarkan table

a. Menentukan kelas
P (Label = 0) = 3/16
= 0.1875
P (Label = 1) = 13/16
= 0.8125

b. Jenis Pinjaman
P (JP = KUR | Label = 0) = 1/3
= 0.333
P (JP = KUR | Label = 1) = 5/13
= 0.384
P (JP = SWAGUNA | Label = 0) = 1/3
= 0.333
P (JP = SWAGUNA | Label = 1) = 3/13
= 0.231
P (JP = MIKRO | Label = 0) = 1/3
= 0.333
P (JP = MIKRO | Label = 1) = 5/13
= 0.384

c. Agunan
P (Agunan = BPKB | Label =0) = 0/3
= 0
P (Agunan = BPKB | Label =1) = 8/13
= 0.615
P (Agunan = SERTIFIKAT | Label =0) = 2/3
= 0.666
P (Agunan = SERTIFIKAT | Label =1) = 2/13 = 0.1538
P (Agunan = MOU | Label =0) = 1/3
= 0.333
P (Agunan = MOU | Label =1) = 3/13
= 0.231

d. Pendapatan
P (Pendapatan = BAIK | Label =0) = 2/3
= 0.667
P (Pendapatan = BAIK | Label =1) = 6/13
= 0.4615
P (Pendapatan = CUKUP | Label =0) = 0/3
= 0
P (Pendapatan = CUKUP | Label =1) = 5/13
= 0.3846

P (Pendapatan = KURANG | Label =0) = 1/3
= 0.333
P (Pendapatan = KURANG | Label =1) = 2/13 = 0.1538

e. Karakter
P (Karakter = BAIK | Label =0) = 1/3
= 0.333
P (Karakter = BAIK | Label =1) = 9/13
= 0.6923
P (Karakter = KURANG BAIK | Label =0) = 0/3 = 0
P (Karakter = KURANG BAIK | Label =1) = 4/13 = 0.3076
P (Karakter = TIDAK BAIK | Label =0) = 2/3 = 0.666
P (Karakter = TIDAK BAIK | Label =1) = 0/13 = 0

3.2 Data Testing

Data testing merupakan data yang akan diprediksi ke dalam sebuah kelas 0 atau masuk ke kelas 1, data testing ditunjukkan tabel 2.

Tabel 2 Data Testing

Jenis Pinjaman	Agunan	Pendapatan	Karakter	Label
Kur	Sertifikat	Cukup	Kurang Baik	?

Perhitungan data training pada tabel 2
P | (0) = 0.333 * 0.666 * 0 * 0.333 * 0.1875 = 0
P | (1) = 0.384 * 0.1538 * 0.3848 * 0.3076 * 0.8125 = 0.00567
Perbandingan nilai P | (0) dengan P | (1) maka nilai yang tertinggi yang dipilih, jadi hasil label dari data training pada tabel adalah 1.

3.3 Implementasi

Berdasarkan pada tabel 1, maka hasil implementasi dari kemunculan data sebagai berikut:

	JENIS_PINJAMAN_KUR	JENIS_PINJAMAN_MIKRO	JENIS_PINJAMAN_SWAGUNA
0	1	0	0
1	1	0	0
2	1	0	0
3	1	0	0
4	1	0	0
5	1	0	0
6	0	0	1
7	0	0	1
8	0	0	1
9	0	0	1
10	0	1	0
11	0	1	0
12	0	1	0
13	0	1	0
14	0	1	0
15	0	1	0

AGUNAN_BPKB	AGUNAN_MOU	AGUNAN_SERTIFIKAT	PENDAPATAN_BAIK	\
1	0	0	1	
1	0	0	1	
0	0	1	1	
1	0	0	0	
1	0	0	0	
0	0	1	0	
0	1	0	1	
0	1	0	0	
0	1	0	1	
0	1	0	0	
1	0	0	1	
1	0	0	1	
0	0	1	1	
1	0	0	0	
1	0	0	0	
0	0	1	0	

PENDAPATAN_CUKUP	PENDAPATAN_KURANG	KARAKTER_BAIK	KARAKTER_KURANG BAIK
0	0	1	0
0	0	0	1
0	0	0	0
1	0	1	0
1	0	0	1
0	1	1	0
0	0	1	0
0	1	1	0
0	0	0	1
1	0	0	1
0	0	1	0
0	0	0	1
0	0	0	0
1	0	1	0
1	0	0	1
0	1	1	0

KARAKTER_TIDAK BAIK
0
0
1
0
0
0
0
0
0
0
0
1
0
0
0

Gambar 1. Frekuensi Data Training lanjutan

kemudian memasukkan data testing pada tabel 2 di masukan maka akan di dapat hasil seperti pada gambar 2.

```
Hasil Prediksi dengan Fitur KUR, SERTIFIKAT, CUKUP, KURANG BAIK = [1]
```

Gambar 2. Hasil dari data testing

3.4. Akurasi Data

Nilai akurasi merupakan cara untuk menentukan kedakatan hasil terhadap nilai sebenarnya.

$$Akurasi = \frac{Data\ yang\ benar}{Total\ Data} * 100\%$$

$$Akurasi = \frac{15}{16} = 0.9375$$

$$Akurasi = 93.75\%$$

4. Kesimpulan dan Saran

Berdasarkan hasil perhitungan dari data training dan data testing maka bisa di tarik kesimpulan:

1. Penerapan klasifikasi dengan metode naive bayes untuk menentukan pinjaman di bank XYZ dapat membantu bank yang bersangkutan.
2. Hasil perhitungan dari 16 data training pada tabel 1 dan data testing pada tabel 2 dapat diklasifikasikan ke dalam kelas 1 dengan nilai 0.00567
3. Uji data dengan 12 data training dan 3 data testing didapat nilai akurasi 93.75%

DAFTAR PUSTAKA

Annur, H. 2018. Klasifikasi Masyarakat Miskin Menggunakan Metode Naïve Bayes. ILKOM Jurnal Ilmiah Volume 10 Nomor 2 Agustus 2018.

Antaristi. M dan Kurniawan. I., Y. 2017. Aplikasi Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan Metode Naive Bayes di Bank BNI Syariah Surabaya. Jurnal Teknik Elektro Vol. 9 No. 2.

Ariani. F., Amir, Alam. N., Rizal. K. 2018. Klasifikasi Penetapan Status Karyawan

Dengan Menggunakan Metode Naïve Bayes. Paradigma Volume XX No. 2 September 2018.

Manalu, E., Ariwisanto, F., Manalu, R. M. 2017. Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan Dan Jumlah Pemesanan Pada CV. Papadan Mama Pastries. Jurnal Mantik Penusa. Volume 1 No 2 Desember 2017.

Saleh, A. 2015. Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. Citec Journal, Vol. 2, No. 3, Mei 2015 – Juli 2015.