

PENERAPAN ADABOOST UNTUK METODE NAIVE BAYES DALAM MENGLASIFIKASI KUALITAS PEMBIAYAAN NASABAH BANK

Edi Noersasongko^a, Purwanto^b, M. Arief Soeleman^c, Akhmad Kuncoro^d
edi.noersasongko@dsn.dinus.ac.id^a, purwanto@dsn.dinus.ac.id^b, m.arief.soeleman@dsn.dinus.ac.id^c,
akhmad.22005@mhs.unesa.ac.id^d

Abstract

The problem that often occurs in providing bank customer financing is not on target, so that many financing customers are not current due to a lack of more accurate initial analysis to select prospective customers. Because the dataset of prospective bank customers has a lot of missing attribute information. The Naive Bayes method is known to provide good accuracy compared to other methods in dealing with missing attributes. However, if the results obtained are still not optimal, optimization with adaboost is used to improve accuracy. This study discusses the application of the Naive Bayes method and Adaboost optimization to classify the quality of bank customer financing. The results obtained using the Naive Bayes method achieved an accuracy of 99.485% while the Naive Bayes method and optimization using Adaboost succeeded in increasing the accuracy of 99.691%. Adaboost managed to increase accuracy by 0.206%. It was concluded that the optimization of AdaBoost in the Naive Bayes method can increase the accuracy for the classification of the quality of bank customer financing.

Keywords: *Naive bayes, AdaBoost, Financing Quality, Classification, Bank Customers.*

Abstrak

Masalah yang sering terjadi dalam pemberian pembiayaan nasabah bank kurang tepat sasaran, sehingga banyak nasabah pembiayaan yang tidak lancar akibat kurangnya menganalisa diawal yang lebih akurat untuk memilih calon nasabah. Disebabkan dataset calon nasabah bank banyak informasi atribut yang hilang. Metode *naive bayes* dikenal memberikan akurasi yang baik dibandingkan dengan metode lain dalam menangani atribut yang hilang. Namun, apabila hasil yang diperoleh masih belum maksimal maka optimasi dengan *adaboost* digunakan untuk meningkatkan akurasinya. Penelitian ini membahas penerapan metode *naive bayes* dan optimasi *adaboost* untuk mengklasifikasi kualitas pembiayaan nasabah bank. Hasil yang diperoleh dengan menggunakan metode *naive bayes* mencapai akurasi sebesar 99.485% sedangkan metode *naive bayes* dan optimasi menggunakan *adaboost* berhasil meningkatkan akurasi sebesar 99.691%. *Adaboost* berhasil meningkatkan akurasi sebesar 0.206%. Disimpulkan optimasi *AdaBoost* dalam metode *Naive Bayes* dapat meningkatkan akurasi untuk klasifikasi kualitas pembiayaan nasabah bank.

Kata Kunci : *Naive bayes, AdaBoost, Kualitas Pembiayaan, Klasifikasi, Nasabah Bank.*

Pendahuluan

Kebutuhan adalah segala sesuatu yang harus dipenuhi oleh seseorang untuk keberlangsungan hidupnya, dari sesuatu yang paling penting untuk keberlangsungan hidupnya sampai sesuatu untuk kepuasan dirinya sendiri. Kebutuhan terdiri dari tiga macam, yaitu kebutuhan primer, kebutuhan sekunder, dan kebutuhan tersier. Kebutuhan primer adalah kebutuhan pokok atau kebutuhan yang wajib dipenuhi. Kebutuhan primer terdiri dari pangan (makanan), sandang (pakaian) dan papan (rumah) [1].

Kebutuhan primer papan yang berupa rumah sudah menjadi kebutuhan dengan proporsi utama. Namun, di Indonesia masih menghadapi persoalan *backlog* perumahan. *Backlog* perumahan adalah suatu kondisi di mana jumlah rumah yang ada belum mampu menampung seluruh rumah tangga. Berdasarkan konsep penghunian, pada tahun 2015 masih terjadi *backlog* perumahan sebesar 6,9 juta unit [2]. Adapun berdasarkan konsep kepemilikan, masih terdapat *backlog* sebesar 11,4 juta unit. Defisit perumahan telah menjadi perhatian pemerintah. Berbagai kebijakan seperti Fasilitas Likuiditas Pembiayaan Perumahan (FLPP) dan Program Sejuta Rumah disiapkan oleh pemerintah. Namun, persoalan *backlog* dan kesenjangan perumahan bukan persoalan mudah [2].

Empat puluh persen masyarakat termiskin (desil 1 sampai 4) tidak memiliki dana yang cukup untuk berinvestasi atau membeli rumah. Padahal, pada kelompok 40% masyarakat termiskin inilah terjadi masalah rumah tidak layak huni. Pada desil 1 (10 persen masyarakat termiskin) misalnya, rata-rata penghasilan rumah tangga per bulan hanya sebesar Rp 1.200.000,-. Penghasilan ini dipakai seluruhnya untuk membiayai pengeluaran selama sebulan penuh, sehingga kelompok masyarakat ini tidak memiliki dana yang cukup untuk berinvestasi dan membeli rumah [3].

Begitu juga pada desil 4 (40 persen masyarakat termiskin) yang memiliki rata-rata penghasilan sebesar Rp 2.600.000,- dan rata-rata pengeluaran sebesar Rp 1.800.000,-. Hal tersebut menyebabkan banyaknya permintaan pembiayaan pembelian melalui lembaga keuangan yang disebabkan oleh keterbatasan dana yang dimiliki oleh nasabah untuk memiliki sebuah rumah. Sebenarnya pemerintah telah melakukan kebijakan untuk masyarakat yang tidak memiliki rumah yaitu dengan adanya program rumah subsidi [4].

Pada umumnya lembaga keuangan konvensional menggunakan sistem bunga pada kredit kepemilikan rumah, yang sebenarnya dalam hukum Islam tidak diperbolehkan (dilarang). Dengan adanya sistem bunga yang diberlakukan dalam lembaga keuangan konvensional menyebabkan masyarakat sulit untuk menggunakan produk pembiayaan tersebut.

Sistem bunga yang diterapkan oleh lembaga keuangan konvensional khususnya dalam skema kredit kepemilikan rumah, menyebabkan masyarakat muslim ragu untuk menggunakan skema pembiayaan tersebut [5].

Sistem bunga pada kredit kepemilikan rumah konvensional, menyebabkan angsuran yang dibayarkan oleh nasabahnya selalu berubah-ubah, sehingga menyulitkan nasabah saat pembayaran angsuran terlebih lagi untuk nasabah yang memiliki pendapatan rendah (upah minimum) [6]. Sedangkan, lembaga keuangan syariah melakukan kegiatannya tidak sama dengan lembaga keuangan konvensional, yang tidak akan menyebabkan angsuran yang berubah-ubah dan tidak akan memberatkan nasabahnya dalam hal angsuran yang dibayarkan untuk setiap bulannya, karena besaran angsuran yang dibayarkan setiap bulannya ditetapkan secara bersama-sama dan sesuai dengan kemampuan nasabah dan sesuai dengan syarat dan ketentuan yang berlaku pada lembaga keuangan syariah tersebut [7].

Menurut Peraturan Menteri Tenaga Kerja dan Transmigrasi No.7 Tahun 2013 terdapat beberapa definisi terkait upah minimum, salah satunya mengatakan bahwa upah minimum adalah upah bulanan terendah yang terdiri atas upah pokok termasuk tunjangan tetap yang ditetapkan oleh gubernur sebagai jaring pengaman [8]. Menurut Undang-Undang No.78 Tahun 2015 tentang Pengupahan, sebagaimana pada Kebijakan Pengupahan pasal 3 ayat (1) menyatakan bahwa kebijakan pengupahan diarahkan untuk pencapaian penghasilan yang memenuhi penghidupan yang layak bagi pekerja/buruh [8]. Sesuai dengan peraturan di atas maka setiap kota

memiliki upah minimum yang berbeda sesuai dengan ketetapan dari gubernurnya, untuk di Bandar Lampung UMK/UMP (upah minimum kota/provinsi) pada tahun 2017 ialah sebesar Rp 2.054.365,- dan pada tahun 2018 ialah sebesar Rp 2.074.673,- [9].

Untuk masyarakat yang ingin memiliki rumah (hunian) namun tidak memiliki cukup uang (dana) untuk membeli secara *cash*, maka dapat menggunakan fasilitas pembiayaan Kredit Kepemilikan Rumah (KPR), namun biasanya hanya digunakan oleh masyarakat berpenghasilan menengah keatas atau masyarakat yang memiliki penghasilan tinggi, sedangkan untuk masyarakat berpenghasilan rendah menganggap KPR sulit dijangkau untuk mereka karena faktor ekonomi (pendapatan) yang menjadi alasannya [10].

Masyarakat di era digital seperti sekarang ini, semakin jeli dan semakin pintar dalam mengambil keputusan dengan banyaknya media, dan informasi yang mereka dapat mereka banyak melakukan pertimbangan, khususnya dalam keputusan Kredit Pemilikan Rumah (KPR) melalui lembaga keuangan. Masyarakat dapat dengan mudah untuk mengetahui mengenai sistem bunga KPR pada lembaga keuangan konvensional dan mengenai sistem bunga yang tidak diberlakukan pada lembaga keuangan syariah [4].

Di tengah situasi ekonomi yang mengalami inflasi, KPR syariah menjadi salah satu solusi untuk dapat memiliki rumah namun tidak dengan perubahan angsuran setiap bulannya karena tidak menggunakan sistem bunga. Terlebih lagi, pada KPR syariah angsurannya ditetapkan secara bersama-sama antara pihak bank dan nasabah sesuai dengan kemampuan nasabah [11]. Peluang inilah yang digunakan oleh lembaga keuangan syariah dalam mengeluarkan produk pembiayaan kepemilikan rumah secara syariah. Dengan adanya peluang tersebut, telah banyak bank syariah yang menyediakan produk pembiayaan kepemilikan rumah melalui proses syariah. Kredit Kepemilikan Rumah (KPR) syariah ini mendapatkan respon yang positif dari masyarakat [12].

Kredit Kepemilikan Rumah (KPR) syariah yang ada pada perbankan syariah berbeda dengan KPR pada perbankan konvensional. Perbedaan ini terdapat pada perbedaan prinsip antara perbankan syariah dengan perbankan konvensional. Jika dalam perbankan syariah biasa dikenal konsep berbasis bagi hasil dan juga perdagangan, sedangkan pada perbankan konvensional dikenal dengan kredit. KPR Syariah dapat menggunakan beberapa akad seperti akad *murabahah*, akad *Ijarah Muntahiyah Bit Tamlik* (IMBT).

Untuk penggunaan akad pada pembiayaan Kredit Kepemilikan Rumah (KPR) syariah sesuai dengan kesepakatan antara nasabah dengan pihak bank. Pembiayaan pemilikan rumah dengan akad *murabahah* atau jual beli sesuai dengan Fatwa Dewan Syariah No.04/DSN-MUI/IV/2000 tentang *murabahah*. Pembiayaan pemilikan rumah dengan akad *Ijarah Muntahiyah Bit Tamlik* (IMBT) sesuai dengan Fatwa Dewan Syariah No.27/DSN/MUI/III/2002 tentang *Ijarah Muntahiyah Bit Tamlik* [6].

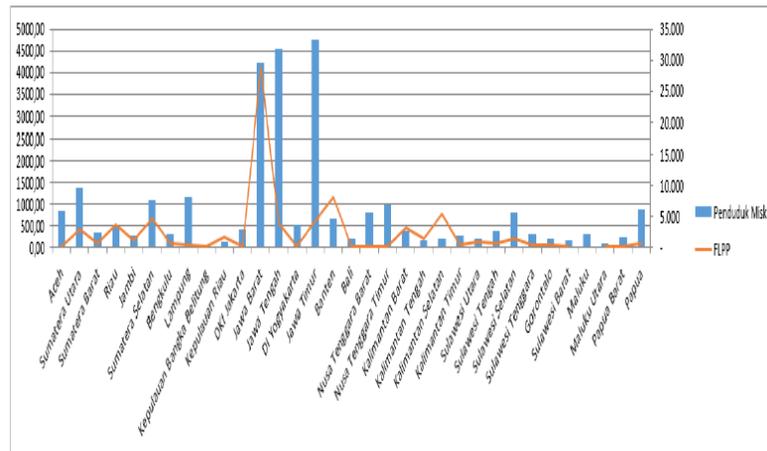
Menurut Peraturan Menteri Negara Perumahan Rakyat Nomor 13 Tahun 2010 tentang Tata Cara Pelaksanaan KPR Bersubsidi dan KPR Syariah Bersubsidi Serta KPR Sarusuna Bersubsidi dan KPR Sarusuna Syariah Bersubsidi, kelompok sasaran yang diperbolehkan menerima subsidi perumahan ini adalah yang memiliki penghasilannya per bulan paling banyak Rp 2.500.000,- (dua juta lima ratus ribu rupiah) [2].

Pada tahun 2015 sesuai Peraturan Menteri Pekerjaan Umum dan Perumahan Rakyat Republik Indonesia Nomor 48/PRT/M/2015 tentang Skema Selisih Angsuran Kredit/Pembiayaan Pemilikan Rumah Bagi Masyarakat Berpenghasilan Rendah Dengan Menggunakan Pendapatan Badan Layanan Umum Pusat Pengelolaan Dana Pembiayaan, kelompok sasaran yang

diperbolehkan menerima subsidi perumahan ini adalah yang memiliki gaji pokok per bulan paling banyak Rp 4.000.000,- (empat juta rupiah). Dari peraturan tersebut dapat diketahui bahwa tidak ada batas minimum pendapatan untuk mendapatkan pembiayaan KPR Syariah khususnya KPR subsidi [2].

Pola Ketersediaan pasokan rumah program FLLP di masing-masing provinsi belum selaras dengan potensi MBR yang masing-masing provinsi. Hal ini dibuktikan dengan grafik dibawah ini [23].

**Grafik Perbandingan antara
 Ketersediaan pasokan rumah program FLPP per provinsi (unit)
 dengan jumlah penduduk miskin per provinsi (orang)**



Sumber: Badan Pusat Statistik dan Kementerian Perumahan Rakyat 2014, diolah.

Gambar 1.1 Grafik Perbandingan Program FLPP dengan MBR [23]

Proses Pertumbuhan Pembiayaan sangat perlu, tetapi jangan lupa mengendalikan kualitas pembiayaan, karena asset pembiayaan adalah sumber utama pendapatan Bank, maka kualitas pembiayaan harus dinilai transparan, sehingga saat ini.

Penetapan kolektibilitas pembiayaan dapat hanya didasarkan pada ketepatan pembayaran. Lancar (kolektibilitas 1) yaitu apabila tidak terdapat tunggakan pembayaran pinjaman baik pokok ataupun margin/bagi hasil/fee, Dalam Perhatian Khusus (kolektibilitas 2) yaitu apabila terdapat tunggakan pinjaman pembayaran pokok dan atau margin/bagi hasil/fee sampai dengan 90 hari, Kurang Lancar (kolektibilitas 3) yaitu apabila terdapat tunggakan pembayaran pokok dan atau margin/bagi hasil/fee sampai dengan 120 hari, Diragukan (kolektibilitas 4) yaitu apabila terdapat tunggakan pembayaran pinjaman baik pokok dan atau margin/bagi hasil/fee sampai dengan 180 hari dan Macet (kolektibilitas 5) yaitu apabila terdapat tunggakan pembayaran pokok dan atau margin/bagi hasil/fee diatas 180 hari. Tujuan penetapan kolektibilitas pembiayaan adalah untuk mengetahui kualitas pembiayaan sehingga bank dapat mengantisipasi risiko secara dini karena risiko pembiayaan dapat mempengaruhi kelangsungan usaha bank. Disamping itu penetapan kolektibilitas pembiayaan digunakan untuk menetapkan tingkat cadangan potensi kerugian akibat pembiayaan bermasalah [24].

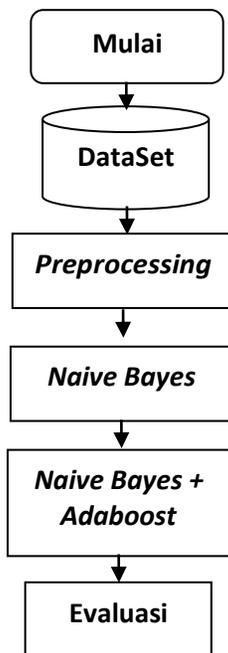
Algoritma *Naive Bayes* ditemukan oleh Thomas Bayes. *Simple Naive Bayesian (Naive Bayes = NB) Classifier* merupakan salah satu metode pengklasifikasi berpeluang sederhana yang berdasarkan pada penerapan teorema *Bayes* dengan asumsi antar variabel penjelas saling bebas (independen). *Naive Bayes* bekerja sangat efektif ketika diuji pada aktual datasets [13].

Algoritma *AdaBoost* pertamakali diperkenalkan pada tahun 1995 oleh Freund dan Schapire, telah banyak memecahkan berbagai masalah praktis dari algoritma *boosting* sebelumnya [14].

Dari hasil penelitian sebelumnya dapat disimpulkan bahwa metode *Naïve Bayes Classifier* cukup sukses dalam mengklasifikasikan data Penyakit ginjal kronik / (*Chronic Kidney Disease* = CKD). Meskipun dataset yang diteliti memiliki banyak *missing value*, *Naïve Bayes* dapat memperoleh nilai akurasi sebesar 0,95 dan *F1-score* sebesar 0,958. Sedangkan apabila dataset tersebut digantikan atau diisi *missing value*-nya dengan mean dan modus, hasil akurasi metode *Naïve Bayes* ini justru mengalami penurunan sebesar 0,005 yaitu menjadi 0,945. Untuk nilai *F1-score* hanya mengalami penurunan yang sangat rendah menjadi 0,954. Hal ini juga membuktikan bahwa metode *Naïve Bayes* cukup handal menghadapi data set yang memiliki banyak *missing value* sesuai dengan penelitian yang dilakukan oleh Liu, Lei, dan Wu. Penggunaan metode *boosting* (*AdaBoost*) juga terbukti mampu meningkatkan performa dari metode *Naïve Bayes*. Dari 10 kali perulangan yang diujicobakan, nilai akurasi untuk algoritma *AdaBoost* pada dataset asli sebesar 0,98 dan *F1-score* sebesar 0,984. Sedangkan untuk pengujian dengan dataset yang sudah melalui proses penggantian *missing value* nilai akurasi yang diperoleh meningkat menjadi 0,9825 dan *F1-score* sebesar 0,986 [25].

Metode Penelitian

Bab ini membahas tahapan metode penelitian. Tahapan yang digunakan untuk menjawab pertanyaan pada permasalahan penelitian dan untuk mencapai tujuan penelitian. Alur tahapan metode yang diusulkan pada penelitian ini diilustrasikan sebagaimana pada gambar 3.1 berikut:



Gambar 3.1 Metode Yang Diusulkan

3.1 Pengolahan Data

Dalam melakukan laporan penelitian ini dibutuhkan data dan informasi yang *real* dan

sesuai dengan masalah yang akan diselesaikan. Data yang akan digunakan berupa data dari nasabah Bank ABC dari April 2008 sampai dengan Desember 2018 dan data yang diperoleh ada 9.709 data *record*, berikut adalah 17 atribut data diantaranya adalah : Nama Produk, CIF, No Pemby, Nama, No Rek, Oust. Pokok, DD, Angsuran, RPC, Lokasi Agunan, Motivasi KPR, Kelengkapan Sarana & Prasarana, Status Perkawinan, Status Penghasilan, Alamat Agunan, Alamat KTP dan No Telp.

3.2 Pengolahan Data Awal

Setelah melakukan pengumpulan data maka diperoleh data sejumlah 9.709 record. Berikut atribut data yang diperoleh :

Tabel 3.1 Data Atribut

NO.	NAMA ATRIBUT	KETERANGAN
1.	Nama Produk	Nama Produk Pembiayaan
2.	CIF	No. CIF Rekening Nasabah
3.	No Pemby	No. Loan/Pembiayaan Nasabah
4.	Nama	Nama Nasabah
5.	No Rek	No. Rekening Nasabah
6.	RPC	Kemampuan Membayar Angsuran
7.	Lokasi Agunan	Lokasi agunan Marketable atau tidak
8.	Motivasi KPR	Rumah Ditempati/Untuk Investasi
9.	Kelengkapan Sarana & Prasarana	Kelengkapan Listrik, Air dan sarana prasarana perumahan
10.	Status Perkawinan	Kawin atau Belum Kawin
11.	Status Penghasilan	Penghasilan tetap atau penghasilan tidak tetap
12.	Oust. Pokok	Sisa Pinjaman Pembiayaan
13.	DD	Tanggal Akad Pembiayaan/Tanggal Jatuh Tempo Angsuran perbulan
14.	Angsuran	Nominal angsuran perbulan
15.	Alamat Agunan	Alamat Agunan/Alamat Rumah KPR
16.	Alamat KTP	Alamat KTP Nasabah
17.	No. Telp	Nomor Telp Nasabah

Untuk mendapatkan model yang baik maka terlebih dahulu, data dilakukan pengolahan tahap awal. Berikut tahap awal pengolahan data :

Data *Cleaning* adalah Data yang diperoleh dalam suatu penelitian yang dilakukan di Bank ABC memiliki isian yang belum sempurna, seperti data kurang lengkap, data salah tulis, yang disebabkan waktu penginputan data. Data seperti ini lebih baik dibuang agar tidak mengurangi mutu dan akurasi hasil data mining.

Setelah dilakukan konsultasi dengan bagian ahli Pembinaan Nasabah Pembiayaan di Bank ABC, atribut yang sangat mempengaruhi untuk penentuan nasabah kolektibilitas Lancar atau Dalam Perhatian Khusus Sebagai berikut :

Tabel 3.2 Data Atribut Setelah *Preprocessing*

No.	Nama Atribut	Keterangan
1.	RPC	Kemampuan Membayar Angsuran
2.	Lokasi Agunan	Lokasi Agunan Marketable Atau Tidak
3.	Motivasi KPR	Rumah Ditempati/Untuk Investasi
4.	Kelengkapan Sarana & Prasarana	Kelengkapan Listrik, Air Dan Sarana Prasarana Perumahan
5.	Status Perkawinan	Belum Kawin/Kawin
6.	Status Penghasilan	Penghasilan Tetap/Tidak Tetap

3.3 Proses Training *Naive Bayes*

Proses training dataset dalam penelitian ini untuk memprediksikan kriteria Kolektibilitas Pembiayaan Dalam Perhatian Khusus dan Kolektibilitas Lancar menggunakan metode *Naive Bayes*, dan tool weka.

3.4 Proses Training *Naive Bayes* dan *AdaBoost*

Proses training dataset dalam penelitian ini untuk memprediksikan kriteria Kolektibilitas Pembiayaan Dalam Perhatian Khusus dan Kolektibilitas Lancar menggunakan Metode *Naive Bayes* dan Optimasi menggunakan *AdaBoost*. Pada dasarnya, metode boosting ini bertujuan untuk meningkatkan akurasi klasifikasi setelah dilakukan training Dataset Menggunakan *Naive Bayes* dan tool weka.

3.5 Evaluasi Pengukuran Keakuratan Hasil

Dalam tahap ini dilakukan Evaluasi pengukuran keakuratan hasil yang dicapai model yang menggunakan *Confusion Matrix* untuk pengukuran tingkat akurasi model dan *split validasi*.

Salah satu tahapan akhir yang penting dari sebuah proses penelitian adalah evaluasi dari model yang dihasilkan. Evaluasi dari model yang dihasilkan bertujuan untuk mengukur performa model tersebut. Untuk melakukan sebuah tes, seorang peneliti harus memiliki sekumpulan dataset yang terpisah dan tidak berhubungan dengan dataset yang dipakai untuk membentuk model tersebut.

Salah satu cara untuk melakukan evaluasi sebuah model adalah dengan menggunakan metode *confusion matrix*. *Confusion matrix* yang sering juga disebut *classification matrix* merupakan sebuah matrik yang memberikan gambaran penuh mengenai tingkat kesalahan (*error rate*) serta kualitas prediksi sebuah model. Tabel 3.3 dibawah ini merupakan contoh sebuah *confusion matrix* dengan memakai dua buah kelas ("*true*", "*false*").

Hasil Penelitian

4.1. Data

Dalam melakukan laporan penelitian ini dibutuhkan data dan informasi yang *real* dan sesuai dengan masalah yang akan diselesaikan. Dataset yang akan digunakan berupa data private dari nasabah Bank ABC dari April 2008 sampai dengan Desember 2018, Data yang diperoleh ada 9.709 data *record*, setelah dilakukan *preprocessing* oleh ahlinya variabel yang sangat menentukan Kolektibilitas lancar atau Kolektibilitas DPK terdapat 6 variabel,

sebagai berikut :

Tabel 4.1 Data Preprocessing

NO	RPC	LOKASI AGUNAN	MOTIVASI KPR	KELENGKAPAN SARANA DAN PRASARANA	STATUS PERKAWINAN	STATUS PENGHASILAN	KOL
1	STABIL	MAKETABLE	DITEMPATI	LENGKAP	KAWIN	PENGHASILAN TETAP	LANCAR
2	MENURUN	MAKETABLE	INVESTASI	TIDAK LENGKAP	BELUM KAWIN	PENGHASILAN TIDAK TETAP	DPK
3	STABIL	MAKETABLE	DITEMPATI	LENGKAP	BELUM KAWIN	PENGHASILAN TETAP	LANCAR
4	MENURUN	MAKETABLE	INVESTASI	TIDAK LENGKAP	KAWIN	PENGHASILAN TIDAK TETAP	DPK
5	STABIL	MAKETABLE	DITEMPATI	LENGKAP	KAWIN	PENGHASILAN TETAP	LANCAR
6	STABIL	MAKETABLE	DITEMPATI	LENGKAP	BELUM KAWIN	PENGHASILAN TETAP	LANCAR
7	STABIL	MAKETABLE	DITEMPATI	LENGKAP	KAWIN	PENGHASILAN TETAP	LANCAR
8	MENURUN	MAKETABLE	DITEMPATI	LENGKAP	KAWIN	PENGHASILAN TETAP	LANCAR
9	MENURUN	MAKETABLE	DITEMPATI	LENGKAP	BELUM KAWIN	PENGHASILAN TETAP	LANCAR
10	MENURUN	MAKETABLE	DITEMPATI	LENGKAP	KAWIN	PENGHASILAN TETAP	LANCAR
11	MENURUN	MAKETABLE	INVESTASI	TIDAK LENGKAP	BELUM KAWIN	PENGHASILAN TIDAK TETAP	DPK
...							
...							
...							
9709	MENURUN	MAKETABLE	DITEMPATI	LENGKAP	BELUM KAWIN	PENGHASILAN TIDAK TETAP	LANCAR

4.2 *Naive Bayes*

4.2.1 Menghitung probabilitas tiap label data training

Perhitungan *Naive Bayes* untuk kualitas pembiayaan nasabah bank, contoh 10 data training pada tabel 4.2 data training sebagai berikut:

Tabel 4.2 Data Training

NO	RPC	LOKASI AGUNAN	MOTIVASI KPR	KELENGKAPAN SARANA DAN PRASARANA	STATUS PERKAWINAN	STATUS PENGHASILAN	KOL
1	STABIL	MAKETABLE	DITEMPATI	LENGKAP	KAWIN	PENGHASILAN TETAP	LANCAR
2	MENURUN	MAKETABLE	INVESTASI	TIDAK LENGKAP	BELUM KAWIN	PENGHASILAN TIDAK TETAP	DPK
3	STABIL	MAKETABLE	DITEMPATI	LENGKAP	BELUM KAWIN	PENGHASILAN TETAP	LANCAR
4	MENURUN	MAKETABLE	INVESTASI	TIDAK	KAWIN	PENGHASILAN	DPK

				LENGKAP		TIDAK TETAP	
5	STABIL	MAKETABLE	DITEMPATI	LENGKAP	KAWIN	PENGHASILAN TETAP	LANCAR
6	STABIL	MAKETABLE	DITEMPATI	LENGKAP	BELUM KAWIN	PENGHASILAN TETAP	LANCAR
7	STABIL	MAKETABLE	DITEMPATI	LENGKAP	KAWIN	PENGHASILAN TETAP	LANCAR
8	MENURUN	MAKETABLE	DITEMPATI	LENGKAP	KAWIN	PENGHASILAN TETAP	LANCAR
9	MENURUN	MAKETABLE	DITEMPATI	LENGKAP	BELUM KAWIN	PENGHASILAN TETAP	LANCAR
10	MENURUN	MAKETABLE	DITEMPATI	LENGKAP	KAWIN	PENGHASILAN TETAP	LANCAR

Untuk langkah pertama yang disampaikan pada tabel 4.2 Data Training diatas, dapat dilihat bahwa jumlah data Label Kol = Lancar sebanyak 8 data dan jumlah data training sebanyak 10 data, sehingga peluang Kol = Lancar adalah jumlah data Kol Lancar dibagi dengan keseluruhan jumlah data sehingga diperoleh 8/10.

Selanjutnya dari data training diatas, dapat dilihat Label Kol = DPK sebanyak 2 data sehingga diperoleh peluang dari Kol = DPK adalah 2/10 yang diperoleh dari jumlah Label Kol = DPK dibagi dengan jumlah data atau dapat ditampilkan probabilitas tiap label data training sebagai berikut :

$$P(\text{Kol} = \text{Lancar}) = 8 / 10$$

$$P(\text{Kol} = \text{DPK}) = 2 / 10$$

- **Langkah selanjutnya menghitung probabilitas kasus perlabel data training**

Menghitung probabilitas kasus perlabel pada Variabel RPC terlihat bahwa jumlah Variabel RPC = Stabil sebanyak 5 data, sehingga peluang jika Label Kol = Lancar dengan Variabel RPC = Stabil adalah 5/8 yang diperoleh dari jumlah Variabel RPC = Stabil dibagi dengan jumlah Label Kol = Lancar. Selanjutnya menghitung peluang Variabel RPC = Menurun ketika syaratnya Label Kol = Lancar dengan cara jumlah Label Lancar dengan cara jumlah Variabel RPC = Menurun dibagi dengan jumlah Label Kol = Lancar sehingga diperoleh 3/8, karena telah diketahui bahwa jumlah data Label Kol = DPK adalah sebanyak 0, sehingga peluang Variabel RPC = Stabil ketika syarat Label Kol = DPK adalah 0. Yang diperoleh dari jumlah Variabel RPC = Stabil dibagi jumlah Label Kol = DPK. Selanjutnya menghitung Variabel RPC = Menurun ketika syaratnya Label Kol = DPK adalah dengan cara jumlah Variabel RPC = Menurun dibagi dengan jumlah Label Kol = DPK sehingga diperoleh 2/2. Dapat dilihat pada tabel 4.3 Probabilitas Variabel RPC sebagai berikut:

Tabel 4.3 Probabilitas Variabel RPC

RPC	KOL = LANCAR	KOL = DPK
Stabil	5 / 8	0
Menurun	3 / 8	2 / 2

Menghitung probabilitas kasus perlabel pada Variabel Lokasi Agunan terlihat bahwa jumlah Variabel Lokasi Agunan = Marketable sebanyak 8 data, sehingga peluang jika Label Kol = Lancar dengan Variabel Lokasi Agunan = Marketable adalah 8/8 yang diperoleh dari jumlah Variabel Lokasi Agunan = Marketable dibagi dengan

jumlah Label Kol = Lancar. Selanjutnya menghitung peluang Variabel Lokasi Agunan = Tidak Marketable ketika syaratnya Label Kol = Lancar dengan cara jumlah Label Lancar dengan cara jumlah Variabel Lokasi Agunan = Tidak Marketable dibagi dengan jumlah Label Kol = Lancar sehingga diperoleh 0, karena telah diketahui bahwa jumlah data Label Kol = DPK adalah sebanyak 2, sehingga peluang Variabel Lokasi Agunan = Marketable ketika syarat Label Kol = DPK adalah 2. Yang diperoleh dari jumlah Variabel Lokasi Agunan = Marketable dibagi jumlah Label Kol = DPK sehingga diperoleh 2/2. Selanjutnya menghitung Variabel Lokasi Agunan = Tidak Marketable ketika syaratnya Label Kol = DPK adalah dengan cara jumlah Variabel Lokasi Agunan = Tidak Marketable dibagi dengan jumlah Label Kol = DPK sehingga diperoleh 0. Dapat dilihat pada tabel 4.3 Probabilitas Variabel Lokasi Agunan sebagai berikut:

Tabel 4.4 Probabilitas Variabel Lokasi Agunan

Lokasi Agunan	KOL = LANCAR	KOL = DPK
Marketable	8 / 8	2 / 2
Tidak Marketable	0	0

Menghitung probabilitas kasus perlabel pada Variabel Motivasi KPR terlihat bahwa jumlah Variabel Motivasi KPR = Ditempati sebanyak 8 data, sehingga peluang jika Label Kol = Lancar dengan Variabel Motivasi KPR = Ditempati adalah 8/8 yang diperoleh dari jumlah Variabel Motivasi KPR = Ditempati dibagi dengan jumlah Label Kol = Lancar. Selanjutnya menghitung peluang Variabel Motivasi KPR = Investasi ketika syaratnya Label Kol = Lancar dengan cara jumlah Label Lancar dengan cara jumlah Variabel Motivasi KPR = Investasi dibagi dengan jumlah Label Kol = Lancar sehingga diperoleh 0, karena telah diketahui bahwa jumlah data Label Kol = DPK adalah sebanyak 0, sehingga peluang Variabel Motivasi KPR = Ditempati ketika syarat Label Kol = DPK adalah 0. Yang diperoleh dari jumlah Variabel Motivasi KPR = Ditempati dibagi jumlah Label Kol = DPK. Selanjutnya menghitung Variabel Motivasi KPR = Investasi ketika syaratnya Label Kol = DPK adalah dengan cara jumlah Variabel Motivasi KPR = Investasi dibagi dengan jumlah Label Kol = DPK sehingga diperoleh 2/2. Dapat dilihat pada tabel 4.5 Probabilitas Variabel Motivasi KPR sebagai berikut:

Tabel 4.5 Probabilitas Variabel Motivasi KPR

Motivasi KPR	KOL = LANCAR	KOL = DPK
Ditempati	8 / 8	0
Investasi	0	2 / 2

Menghitung probabilitas kasus perlabel pada Variabel Kelengkapan Sarana & Prasarana terlihat bahwa jumlah Variabel Kelengkapan Sarana & Prasarana = Lengkap sebanyak 8 data, sehingga peluang jika Label Kol = Lancar dengan Variabel Kelengkapan Sarana & Prasarana = Lengkap adalah 8/8 yang diperoleh dari jumlah Variabel Kelengkapan Sarana & Prasarana = Lengkap dibagi dengan jumlah Label Kol = Lancar.

Selanjutnya menghitung peluang Variabel Kelengkapan Sarana & Prasarana = Tidak Lengkap ketika syaratnya Label Kol = Lancar dengan cara jumlah Label Lancar dengan cara jumlah Variabel Kelengkapan Sarana & Prasarana = Tidak Lengkap dibagi dengan

jumlah Label Kol = Lancar sehingga diperoleh 0, karena telah diketahui bahwa jumlah data Label Kol = DPK adalah sebanyak 0, sehingga peluang Variabel Kelengkapan Sarana & Prasarana = Lengkap ketika syarat Label Kol = DPK adalah 0. Yang diperoleh dari jumlah Variabel Kelengkapan Sarana & Prasarana = Lengkap dibagi jumlah Label Kol = DPK.

Selanjutnya menghitung Variabel Kelengkapan Sarana & Prasarana = Investasi ketika syaratnya Label Kol = DPK adalah dengan cara jumlah Variabel Kelengkapan Sarana & Prasarana = Investasi dibagi dengan jumlah Label Kol = DPK sehingga diperoleh 2/2. Dapat dilihat pada tabel 4.6 Data Training Variabel RPC sebagai berikut:

Tabel 4.6 Probabilitas Variabel Kelengkapan Sarana & Prasarana

Kelengkapan Sarana & Prasarana	KOL = LANCAR	KOL = DPK
Lengkap	8 / 8	0
Tidak Lengkap	0	2 / 2

Menghitung probabilitas kasus perlabel pada Variabel Status Perkawinan terlihat bahwa jumlah Variabel Status Perkawinan = Kawin sebanyak 5 data, sehingga peluang jika Label Kol = Lancar dengan Variabel Status Perkawinan = Kawin adalah 5/8 yang diperoleh dari jumlah Variabel Status Perkawinan = Kawin dibagi dengan jumlah Label Kol = Lancar.

Selanjutnya menghitung peluang Status Perkawinan = Belum Kawin ketika syaratnya Label Kol = Lancar dengan cara jumlah Label Lancar dengan cara jumlah Variabel Status Perkawinan = Belum Kawin dibagi dengan jumlah Label Kol = Lancar sehingga diperoleh 3/8,

karena telah diketahui bahwa jumlah data Label Kol = DPK adalah sebanyak 1, sehingga peluang Variabel Status Perkawinan = Kawin ketika syarat Label Kol = DPK adalah 1/2. Yang diperoleh dari jumlah Variabel Status Perkawinan = Kawin dibagi jumlah Label Kol = DPK.

Selanjutnya menghitung Variabel Status Perkawinan = Belum Kawin ketika syaratnya Label Kol = DPK adalah dengan cara jumlah Variabel Status Perkawinan = Belum Kawin dibagi dengan jumlah Label Kol = DPK sehingga diperoleh 1/2. Dapat dilihat pada tabel 4.7 Probabilitas Variabel Status Perkawinan sebagai berikut:

Tabel 4.7 Probabilitas Variabel Status Perkawinan

Status Perkawinan	KOL = LANCAR	KOL = DPK
Kawin	5 / 8	1 / 2
Belum Kawin	3 / 8	1 / 2

Menghitung probabilitas kasus perlabel pada Variabel Status Penghasilan terlihat bahwa jumlah Variabel Status Penghasilan = Penghasilan Tetap sebanyak 7 data, sehingga peluang jika Label Kol = Lancar dengan Variabel Status Penghasilan = Penghasilan Tetap adalah 7/8 yang diperoleh dari jumlah Variabel RPC = Stabil dibagi dengan jumlah Label Kol = Lancar. Selanjutnya menghitung peluang Variabel Status Penghasilan = Penghasilan Tidak Tetap ketika syaratnya Label Kol = Lancar dengan cara jumlah Label Lancar dengan cara jumlah Variabel Status Penghasilan = Penghasilan Tidak Tetap dibagi dengan jumlah Label Kol = Lancar sehingga diperoleh 1/8, karena telah diketahui bahwa jumlah data Label Kol = DPK adalah sebanyak 0, sehingga

peluang Variabel Status Penghasilan = Penghasilan Tetap ketika syarat Label Kol = DPK adalah 0. Yang diperoleh dari jumlah Variabel Status Penghasilan = Penghasilan Tetap dibagi jumlah Label Kol = DPK. Selanjutnya menghitung Variabel Status Penghasilan = Penghasilan Tidak Tetap ketika syaratnya Label Kol = DPK adalah dengan cara jumlah Variabel Status Penghasilan = Penghasilan Tidak Tetap dibagi dengan jumlah Label Kol = DPK sehingga diperoleh 2/2. Dapat dilihat pada tabel 4.8 DataTraining Variabel Status Penghasilan sebagai berikut:

Tabel 4.8 Probabilitas Variabel Status Penghasilan

Status Penghasilan	KOL = LANCAR	KOL = DPK
Penghasilan Tetap	7 / 8	0
Penghasilan Tidak Tetap	1 / 8	2 / 2

• **Menentukan probabilitas label persoalan data testing**

Misal $X^1 = (RPC = \text{Menurun}, \text{Lokasi Agunan} = \text{Marketable}, \text{Motivasi KPR} = \text{Investasi}, \text{Kelengkapan Sarana dan Prasarana} = \text{Tidak Lengkap}, \text{Status Perkawinan} = \text{Belum Kawin}, \text{Status Penghasilan} = \text{Penghasilan Tidak Tetap})$

$$P(DPK | X^1) = 2/2 \cdot 2/2 \cdot 2/2 \cdot 2/2 \cdot 1/2 \cdot 2/2 = 0,5$$

Probabilitas terjadinya kualitas pembiayaan Kol DPK

Misal $X^1 = (RPC = \text{Menurun}, \text{Lokasi Agunan} = \text{Marketable}, \text{Motivasi KPR} = \text{Ditempati}, \text{Kelengkapan Sarana dan Prasarana} = \text{Lengkap}, \text{Status Perkawinan} = \text{Kawin}, \text{Status Penghasilan} = \text{Penghasilan Tetap})$

$$P(Lancar | X^1) = 3/8 \cdot 8/8 \cdot 8/8 \cdot 8/8 \cdot 5/8 \cdot 7/8 = 0,205078125$$

Probabilitas terjadinya kualitas pembiayaan Kol lancar

• **Langkah Terakhir**

Karena $P(DPK | X^1) > P(Lancar | X^1)$ maka kelas dari X^1 adalah Kol = DPK

Kesimpulan, ketika nasabah RPC menurun, lokasi agunan marketable, motivasi KPR untuk investasi, kelengkapan sarana dan prasarana tidak lengkap, status perkawinan belum kawin, status penghasilan tidak tetap, maka kualitas pembiayaan yang akan terjadi Kol DPK.

Jika $P(Lancar | X^1) > P(DPK | X^1)$ maka kelas dari X^1 adalah Kol = Lancar

Kesimpulan, RPC Menurun, Lokasi Agunan Marketable, Motivasi KPR Ditempati, Kelengkapan Sarana dan Prasarana Lengkap, Status Perkawinan Kawin, Status Penghasilan Penghasilan Tetap, maka kualitas pembiayaan yang akan terjadi Kol Lancar.

4.2.2 Menghitung keseluruhan dataset kualitas pembiayaan nasabah bank

Data kualitas pembiayaan nasabah bank, jumlah data dari setiap label kualitas pembiayaan nasabah bank sebagai berikut :

Tabel 4.9 data awal kita hitung banyak data dari setiap label

Lancar	8.675
DPK	1.034
Jumlah	9.709

Total keseluruhan label ada 9.709 data terdiri dari label Kol Lancar 8.675 data dan Kol DPK sebesar 1.034 data.

Langkah selanjutnya hitung probabilitas setiap label :

Tabel 4.10 hitung probabilitas setiap label

Lancar	0.893501
DPK	0.106499

Jumlah Label Kol lancar sebesar 8.675 dibagi total keseluruhan label 9.709
 hasil probabilitas label lancar 0.893501

Jumlah Label Kol DPK sebesar 1.034 dibagi total keseluruhan label 9.709
 hasil probabilitas label DPK 0.106499

Hitung banyak data dalam setiap sub set data

Tabel 4.11 Hitung banyak data dalam setiap sub set data

	Lancar	DPK
Stabil	8640	99
Menurun	35	935
Marketable	8633	94
Tidak Marketable	42	940
Ditempati	8418	74
Investasi	257	960
Lengkap	8582	133
Tidak Lengkap	93	901
Kawin	4947	724
Belum Kawin	3728	310
Penghasilan Tetap	8637	135
Penghasilan Tidak Tetap	38	899

Hitung Probabilitas data yang ada diatas sebagai berikut :

Tabel 4.12 Hitung Probabilitas data yang ada

	Lancar	DPK
Stabil	0.889999	0.0103
Menurun	0.003708	0.096405
Marketable	0.889278	0.009785
Tidak Marketable	0.004429	0.09692
Ditempati	0.867134	0.007725
Investasi	0.026573	0.09898
Lengkap	0.884025	0.013802
Tidak Lengkap	0.009682	0.092903
Kawin	0.50963	0.074673
Belum Kawin	0.384077	0.032032
Tetap	0.88969	0.014008
Tidak Tetap	0.004017	0.092697

Hitung Probabilitas RPC = Stabil, Lokasi agunan = Marketable , Motivasi KPR = Ditempati, Kelengkapan Prasarana dan sarana = Lengkap, Status = Belum Kawin dan Status Penghasilan = Tetap setiap label yang ada (Lancar dan DPK)

Tabel 4.13 Hitung Probabilitas RPC

Lancar	0.185237888
DPK	5.13505E-13

Hasil implementasi data processing menggunakan Metode *Naive Bayes* menggunakan tools weka

```

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      9659           99.485 %
Incorrectly Classified Instances     50             0.515 %
Kappa statistic                     0.9731
Mean absolute error                  0.0055
Root mean squared error              0.0644
Relative absolute error              2.8849 %
Root relative squared error          20.8792 %
Total Number of Instances           9709

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
                0.996   0.017   0.998     0.996   0.997     0.973  1.000    1.000    LANCAR
                0.983   0.004   0.969     0.983   0.976     0.973  1.000    0.996    DPK
Weighted Avg.   0.995   0.016   0.995     0.995   0.995     0.973  1.000    1.000

=== Confusion Matrix ===

  a  b  <-- classified as
8643 32 | a = LANCAR
 18 1016 | b = DPK
    
```

Gambar 4.1 Hasil implementasi data processing menggunakan Metode *Naive Bayes*

4.3 Proses Training *Naive Bayes* dan *AdaBoost*

AdaBoost ini bertujuan untuk menghasilkan nilai mean dan standar deviasi yang bersumber dari nilai probabilitas yang ada. Dari hasil perhitungan dari probabilitas yang baru akan dihitung tingkat akurasi yang kemudian jika nilai eror dianggap

kurang memuaskan maka nilai mean dan standar deviasi akan terus diperbaiki sampai nilainya dianggap memenuhi kriteria yang diinginkan. Dari hasil diatas hitung mean dan standar deviasinya

Tabel 4.14 hitung mean dan standar deviasi

	Lancar	DPK
Mean	4852.0822	4879.4797
Standar Deviasi	2795.7059	2861.0196
weight sum	8675	1034

Hasil implementasi data processing menggunakan Metode *Naive Bayes* + *Adaboost* menggunakan tools weka

```

time taken to build model: 1.16 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      9679      99.691 %
Incorrectly Classified Instances     30       0.309 %
Kappa statistic                    0.9838
Mean absolute error                 0.005
Root mean squared error             0.0505
Relative absolute error             2.6166 %
Root relative squared error        16.3596 %
Total Number of Instances          9709

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
          0.998   0.010   0.999     0.998   0.998     0.984  1.000    1.000    LANCAR
          0.990   0.002   0.981     0.990   0.986     0.984  1.000    0.998    DPK
Weighted Avg.   0.997   0.009   0.997     0.997   0.997     0.984  1.000    1.000

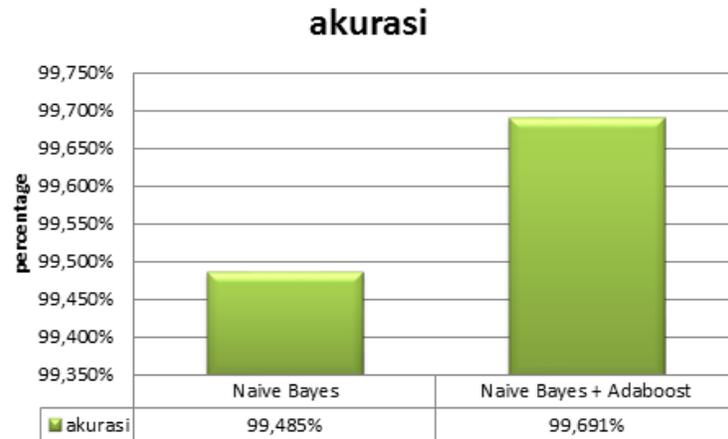
=== Confusion Matrix ===

  a  b  <-- classified as
8655 20 | a = LANCAR
 10 1024 | b = DPK
    
```

Gambar 4.2 Hasil implementasi data processing menggunakan Metode *Naive Bayes* dan *Adaboost*

4.4 Perbandingan

Dari data diatas yang sudah dilakukan dengan metode boosting yaitu *Adaboost* terbukti dapat meningkatkan akurasi dari *Naive bayes* dan Data kualitas pembiayaan bank dapat diklasifikasikan dengan baik.



Gambar 4.3 Hasil Perbandingan menggunakan Metode Naive Bayes dan Naive Bayes + Adaboost

Dari gambar 4.3 diatas, hasil akurasi menggunakan Metode *Naive Bayes* mencapai 99,485% dan menggunakan *Naive Bayes + Adaboost* berhasil meningkatkan akurasi sebesar 0,206 %.

Perhitungan Akurasi :

Tabel 4.15 Tabel Perhitungan Akurasi Metode Naive Bayes

	TP	TN
FP	8643	32
FN	18	1016

$$\text{Akurasi} = \frac{TP+TN}{FP+FN+TP+TN} = \frac{8643+1016}{8643+32+18+1016} = 99.485\%$$

Tabel 4.16 Tabel Perhitungan Akurasi Naive Bayes dan AdaBoost

	TP	TN
FP	8655	20
FN	10	1024

$$\text{Akurasi} = \frac{TP+TN}{FP+FN+TP+TN} = \frac{8655+1024}{8655+20+10+1024} = 99.691\%$$

Tabel 3.3 Contoh Confusion Matrix

	Nilai Sebenarnya		
		<i>Positive (P)</i>	<i>Negative (N)</i>
Hasil Prediksi	<i>Positive (P)</i>	<i>True Positive (TP) d</i>	<i>False Negative (FN) b</i>
	<i>Negative (N)</i>	<i>False Positive (FP) c</i>	<i>True Negative (TN) a</i>

Recall atau *True Positive* (TP), adalah proporsi dari sample bernilai “*true*” yang diprediksi secara benar. TP dihitung dengan menggunakan persamaan: $TP/(FP+TP)$.

False Positive (FP), yaitu proporsi antara sampel bernilai “*false*” yang salah diprediksi sebagai sample bernilai “*true*”. Persamaan yang digunakan adalah: $FN/(TN+FN)$.

True Negative (TN), didefinisikan sebagai perbandingan antara sampel bernilai “*false*” yang diprediksi secara benar. Persamaan yang digunakan adalah: $TN/(TN+FN)$.

False Negative (FN), didefinisikan sebagai proporsi sampel bernilai “*true*” yang salah diprediksi sebagai sampel bernilai “*true*”. Persamaan yang digunakan adalah : $FP/(FP+TF)$.

Akurasi (AC), didefinisikan sebagai proporsi jumlah sampel yang diprediksi secara tepat, terhadap jumlah seluruh sampel. Persamaan yang digunakan adalah: $(TP+TN)/(TN+FN+FP+TP)$.

Presisi (PR), didefinisikan sebagai proporsi jumlah sampel bernilai “*true*” yang berhasil diprediksi secara tepat. Persamaan yang digunakan adalah: $TP/(FN+TP)$.

Error Rate (ER) atau tingkat kesalahan dihitung dengan persamaan $ER = 1 - AC$ [29].

Kesimpulan dan Rekomendasi

1.1. Kesimpulan

Untuk meningkatkan akurasi Metode *Naive Bayes* dalam mengklasifikasi kualitas pembiayaan nasabah pada Bank ABC salah satu caranya dengan menggunakan metode *Naive Bayes* yang sudah dioptimasi dengan *AdaBoost*. Hal ini dikarenakan *AdaBoost* dapat meningkatkan akurasi yang ada pada *Naive Bayes*. Akurasi *Naive Bayes* menghasilkan akurasi sebesar 99.485% sedangkan *Naive Bayes* dengan menambahkan *AdaBoost* menghasilkan akurasi sebesar 99.691% maka mengalami peningkatan sebanyak 0.206% sehingga dapat kita tarik kesimpulan bahwa penambahan *AdaBoost* dalam metode *Naive Bayes* dapat meningkatkan akurasi Klasifikasi Kualitas Pembiayaan Nasabah Bank Khususnya Bank ABC.

1.2. Saran

1. Pengolahan data awal sangat berperan penting dalam akurasi yang dihasilkan sehingga dalam penelitian selanjutnya perlu memperhatikan pengolahan data awal.
2. Dapat menggunakan metode un supervised learning sehingga dapat mengantisipasi unseen data.
3. Menggunakan cara lain dalam menentukan decision tree seperti chi square, Gini Index , dsb.

Daftar Pustaka

- [1] S. Maulana, "Konsep Kebutuhan Dalam Perspektif Islam," *Academia*, 2016.
- [2] D. J. A. K. Keuangan, "Peranan Apbd Dalam Mengatasi Blacklog Perumahan Bagi Masyarakat Berpenghasilan Rendah," 2015.
- [3] B. Winarno, "Evaluasi Program Penyediaan Rumah Layak Huni Bagi Masyarakat Miskin

- Di Kabupaten Belitung," *Jurnal Pengembangan Kota*, 2018.
- [4] A. Savitri, "Analisis Pembiayaan KPR Syariah Terhadap Masyarakat Berpenghasilan Rendah," 2018.
- [5] R. N. Siregar, "Perbandingan Perjanjian Kredit Dalam Prespektif Perjanjian Syariah Islam Dan Kitab Undang-Undang Hukum Perdata," 2017.
- [6] P. P. Sari, "Perbandingan Kredit Pemilikan Rumah Pada Bank Konvensional Dengan Pembiayaan Kepemilikan Rumah Pada Bank Syariah," 2013.
- [7] Millaturrofi'ah, "Analisis Pelaksanaan Pengalihan Hutang (Take Over) Di Bank Jateng Cabang Syariah Semarang," 2017.
- [8] D. Katriani, "Proses Penetapan Upah Minimum Kabupaten Di Kabupaten Purbalingga".
- [9] S. Andriyani, "Dampak Upah Minimum Provinsi Terhadap Kesejahteraan Pekerja Buruh Di Lampung Selatan," 2017.
- [10] Wahyono, "Tinjauan Yuridis Prosedur Pelaksanaan Perjanjian Kredit Pemilikan Rumah (Studi Kasus PT. Bank Tabungan Negara. Tbk Cabang Surakarta)".
- [11] ILMI, "Implementasi Akad Murabahah Pada Produksi Pembiayaan KPR Muamalat Di Bank Muamalat Indonesia Kantor Cabang Malang," 2014.
- [12] S. M. Youdhi Prayogo, "Murabahah Produk Unggulan Bank Syariah Konsep, Prosedur, Penetapan Margin Dan Penerapan Pada Perbankan Syariah".
- [13] Emha Dan Lutfi, "Konsep Dan Teknik Data Mining," 2013.
- [14] Freund, Y.,Schapire, R.E., "A *Short Introduction to Boosting. Journal of Japanese Society For Artificial Intelligence*," 2013.
- [15] Arlan Adrianda," Analisis Faktor–Faktor yang mempengaruhi Kolektibilitas Kredit Pada Cabang Area II Jakarta-PT Bank XYZ," 2011.
- [16] Royan Aziz," Analisis Penyebab Terjadinya Kredit Bermasalah pada PD BPR Bank Gotong Royong Kabupaten Tegal,"2013.
- [17] Gulnur Dereliog lu dan Fikret Gurgens," *Knowledge discovery using neural approach for SME's credit risk analysis problem in Turkey.*"2011.
- [18] Yuliya Bidulya dan Elena Brunova Sentiment, "Analysis for Bank Service Quality:a Rule-based Classifier."
- [19] T. K. Das, *A Customer Classification Prediction Model Based on Machine Learning Techniques*,"2015.
- [20] S. Ummugulthum Natchiar dan Dr.S.Baulkani," *Customer Relationship Management Classification Using Data Mining Techniques*,"2014.
- [21] Youdhi Prayogo, "Murabahah Produk Unggulan Bank Syariah Konsep, Prosedur, Penetapan Margin Dan Penerapan Pada Perbankan Syariah,"2011.
- [22] Wu, X., et al, "Top 10 algorithms in data mining. *Knowledge and Information Systems*," 2007.
- [23] Primadhyta, Safyra, "Pembangunan Perumahan. Biro Analisa Anggaran dan Pelaksanaan APBN-Setjen DPR-RI, "2014.
- [24] Veithzal Rivai dan Arviyan Arifin, "Islamic Banking Sebuah Teori, Konsep, dan

- Aplikasi,” 2010.
- [25] Adhi Indra Irawan, Ristu Saptono dan Afrizal Doewes, “*Implementation Of Naive Bayes Classifier Method And Adaboost Algorithm For Prediction Of Chronic Kidney Disease,*” 2010.
- [26] Asma Feki, Anis Ben Ishak dan Saber Feki, “*Feature selection using Bayesian and multiclass Support Vector Machines approaches: Application to bank risk prediction,*”2011.
- [27] Mudrajad Kuncoro dan Suhardjono, “Manajemen Perbankan Teori dan Aplikasi,” 2012.
- [28] Han, J., Kamber, M., “*Data Mining Concepts and Techniques, 3rd ed.* Watham: Morgan Kauffman,”2012.
- [29] T. Fawcett, “*An introduction to ROC analysis,*” *Pattern Recognit. Lett.*,”2006.