

ANALISIS DATA MINING KELAYAKAN CALON DEBITUR DI PT. BANK BRI DENGAN MENGGUNAKAN METODE KLASIFIKASI *NAIVE BAYES*

Ahmad Turmudzy¹, Donny Maulana², MaulanaSofyan

Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa
¹turmudi@pelitabangsa.ac.id, ²donny.maulana@pelitabangsa.ac.id

Disetujui, 05 Maret 2018

Abstraksi

PT. Bank Rakyat Indonesia (BRI) adalah salah satu bank milik pemerintah yang terbesar di Indonesia. PT. Bank Rakyat Indonesia merupakan salah satu badan usaha perbankan yang melakukan usaha menghimpun dan menyalurkan dan masyarakat, terutama memberikan kredit dan jasa di lalu lintas pembayaran dan peredaran uang, serta memiliki beberapa unit produk, diantaranya pelayanan penyimpanan atau tabungan dan pelayanan perkreditan. Produk pembiayaan pada PT. Bank Rakyat Indonesia terdiri dari berbagai jenis produk, yaitu Kredit Umum Pedesaan (Kupedes) yang terdiri dari kredit Skala Micro (KSM), Kredit Komersial (KOM), Kredit BRIGuna (KBG), dan kredit yang bekerja sama dengan pemerintah yaitu Kredit Usaha Rakyat (KUR). Dengan begitu banyaknya calon debitur yang mengajukan kredit pada PT. Bank Rakyat Indonesia. Kemudian akan dilakukan ujikelayakan pada calon debitur PT. Bank Rakyat Indonesia menggunakan salah satu metode klasifikasi yaitu metode *naive bayesian*, untuk mengetahui calon debitur yang layak mendapatkan kredit pada PT. Bank Rakyat Indonesia dengan tingkat akurasi yang tinggi, agar tidak terjadi masalah ketika sedang berlangsungnya proses perkreditan pada PT. Bank Rakyat Indonesia.

Kata kunci : Analisis, *data mining*, kelayakan, calon debitur, PT. Bank BRI , klasifikasi, *naive bayes*

Abstract

PT. Bank Rakyat Indonesia (BRI) is one of the largest government-owned banks in Indonesia. PT. Bank Rakyat Indonesia is one of the banking business entities that make efforts to collect and channel public funds, primarily providing credit and services in payment traffic and money circulation, and has several product units, including storage services or savings and credit services. Financing products at PT. Bank Rakyat Indonesia consists of various types of products, namely Rural Public Credit (Kupedes) which consists of Micro Scale Credit (KSM), Commercial Credit (KOM), BRIGuna Credit (KBG), and credit in collaboration with the government, namely People's Business Credit (KUR). With so many debtor candidates applying for credit at PT. Bank Rakyat Indonesia. Then a feasibility test will be conducted on prospective borrowers of PT. Bank Rakyat Indonesia uses one of the classification methods, namely the naive bayesian method, to find out which prospective borrowers deserve credit at PT. Bank Rakyat Indonesia with a high degree of accuracy, so that no problems occur while the credit process is taking place at PT. Bank Rakyat Indonesia.

Keyword: *Analysis, data mining, feasibility, prospective borrowers, PT. Bank BRI, classification, naive bayes*

1. Pendahuluan

Bank Rakyat Indonesia merupakan salah satu badan usaha perbankan yang melakukan usaha menghimpun dan menyalurkan dana masyarakat, terutama memberikan kredit dan jasa di lalu lintas pembayaran dan peredaran uang, serta memiliki beberapa unit produk, diantaranya pelayanan penyimpanan atau tabungan dan pelayanan perkreditan. Produk pembiayaan pada Bank Rakyat Indonesia terdiri dari berbagai jenis produk, yaitu Kredit Umum Pedesaan (Kupedes) yang terdiri dari kredit Skala Micro (KSM), Kredit Komersial (KOM), Kredit BRIGuna (GBT), dan kredit yang bekerja sama dengan pemerintah yaitu Kredit Usaha Rakyat (KUR). Kredit adalah kemampuan untuk melaksanakan suatu pembelian atau pinjaman menggunakan suatu perjanjian, pembayaran yang dilaksanakan pada jangka waktu yang telah ditentukan dan disepakati bersama. Dalam praktek sehari-hari pinjaman kredit dinyatakan dalam

bentuk perjanjian tertulis baik akta dibawah tangan maupun secara akta otentik. Sebagai jaminan pengaman, pihak peminjam akan memenuhi kewajiban dan menyerahkan suatu jaminan yaitu baik bersifat kebendaan ataupun bukan kebendaan.

Dalam hal ini penyalur kredit Bank Rakyat Indonesia terdapat permasalahan pada calon debitur seperti BI *Checking*, kekurangannya persyaratan, angsuran yang tidak disanggupi oleh calon debitur, dan usaha yang tidak sesuai ketentuan yang diberikan oleh Bank Rakyat Indonesia.

Hal ini menyebabkan sulitnya debitur dalam proses peminjaman uang kepada Bank Rakyat Indonesia. Sering terjadinya *Double Credit*, artinya nasabah yang sedang dalam proses kredit dapat mengajukan kembali dengan nama pemohon kredit yang lain. Nasabah yang sudah pernah mengajukan dengan status tidak disetujui juga dapat mengajukan kembali dikemudian hari. Dengan melihat hal-hal tersebut sehingga perlu dibutuhkan analisa dalam menentukan calon debitur yang sesuai persyaratan dan ketentuan yang telah berlakudi Bank Rakyat Indonesia.

2. Tinjauan Pustaka

2.1. Data Warehouse

Data warehouse adalah sebuah database yang berfungsi sebagai wadah dalam mengumpulkan informasi dari berbagai sumber berbeda dan menyimpannya di lokasi sentral untuk akses dan analisis yang mudah. Data yang disimpan didalam gudang data adalah sebuah rekaman *statis*, atau *snapshot*, dari apa yang setiap *item* tampak seperti pada titik waktu tertentu (Warehouse, Warehousing, Ware-, & Tantawi PhD, 2016).

Data warehouse adalah suatu pengumpulan data dari berbagai sumber yang digunakan untuk analisis data dan penelitian. Karena keputusan organisasi seringkali dibuat berdasarkan data yang disimpan di gudang data, semua komponennya harus diuji secara ketat. Para peneliti telah mengusulkan sejumlah pendekatan dan alat untuk menguji dan mengevaluasi berbagai komponen sistem *data warehouse* (Homayouni, Ghosh, & Ray, 2018).

2.2. Business Intelligence

Business intelligence adalah ikhtisar lengkap tentang apa yang terdiri dari kecerdasan bisnis. Hal ini dimaksudkan untuk memberikan pengenalan kepada konsep-konsep untuk tidak rumitnya proses pembelajaran ketika

mengimplementasikan program intelijen bisnis (Loshin, 2013).

Business Intelligence adalah sekumpulan teknik dan alat untuk mentransformasi data mentah menjadi informasi yang berguna dan bermakna untuk tujuan analisis bisnis. Teknologi *Business Intelligence* dapat menangani data yang tak terstruktur dalam jumlah yang sangat besar untuk membantu mengidentifikasi, mengembangkan, dan selain itu membuat kesempatan strategi bisnis yang baru.

2.3. Data Mining

Data Mining adalah sebuah rangkaian proses untuk menggali nilai tambah berupa informasi dari kumpulan suatu basis data yang selama ini tidak diketahui secara manual dengan melakukan penggalian pola-pola dari basis data tersebut dengan tujuan untuk memanipulasi data menjadi informasi yang lebih berharga yang diperoleh dengan cara mengekstraksi dan mengenali pola yang penting dan menarik dari data yang terdapat dalam basis data. *Data mining* biasa juga dikenal dengan nama lain seperti : *Knowledge discovery (mining) in databases* (KDD), ekstraksi pengetahuan (*knowledge extraction*) Analisa data atau pola dan kecerdasan bisnis (*business intelligence*) dan merupakan alat yang penting untuk memanipulasi data untuk penyajian informasi sesuai kebutuhan user dengan tujuan untuk membantu dalam analisis koleksi pengamatan perilaku, secara umum definisi *data mining* dapat diartikan sebagai berikut : Proses penemuan pola yang menarik dari data yang tersimpan dalam jumlah besar. Ekstraksi dari suatu informasi yang berguna atau menarik (*non-trivial, implisit*, sebelumnya belum diketahui potensial kegunaannya) pola atau pengetahuan dari data yang disimpan dalam jumlah besar. Ekplorasi dari analisa secara otomatis atau semi otomatis terhadap data-data dalam jumlah besar untuk mencari pola dan aturan yang berarti (Bogor Institution, 2013).

Data Mining merupakan suatu cara untuk menemukan informasi dengan mencari pola atau aturan tertentu dari data dalam jumlah besar yang diharapkan dapat mengatasi kondisi tersebut (Meilani & Susanti, 2015).

Untuk memudahkan pengguna dalam penerapan *data mining*, diperlukan sebuah metodologi standar yang dikenal dengan nama dengan CRISP-DM (*Cross Industry Standard Process for Data Mining*). CRISP-DM (*Cross Industry Standard Process for Data Mining*) adalah suatu *data mining methodology* yang pada awalnya dikembangkan oleh tiga perusahaan, yaitu SPSS (ISL by then), NCR dan Dainler Chrysler pada tahun 1996 dan baru pada bulan

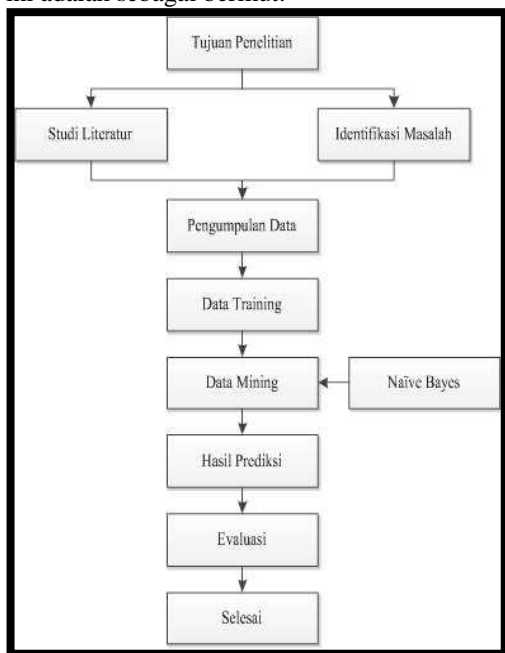
Agustus 2000, versi 1.0 CRISP-DM dipublikasikan. Kemudian tahun 2009, CRISP-DM dan dikenal dengan SEMMA (*sample, explore, modify, model and access*) yang dikembangkan oleh SAS Institute dan CRISP-DM merupakan metodologi *data mining* yang paling banyak yang digunakan (Annisa Paramitha Fadillah, 2015).

2.4. Rapid Miner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (*open source*). *RapidMiner* adalah sebuah solusi untuk melakukan analisis terhadap *data mining, text mining* dan analisis prediksi. *RapidMiner* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. *RapidMiner* memiliki kurang lebih 500 operator *data mining*, termasuk untuk *input, output, data preprocessing* dan visualisasi. *RapidMiner* merupakan *software* yang berdiri sendiri untuk analisis data dan sebagai mesin *data mining* yang dapat diintegrasikan pada produknya sendiri. *RapidMiner* ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi (Dennis, 2013).

3. Kerangka Konsep

Berikut ini merupakan metode rancangan yang akan dilakukan pada penelitian ini adalah sebagai berikut:



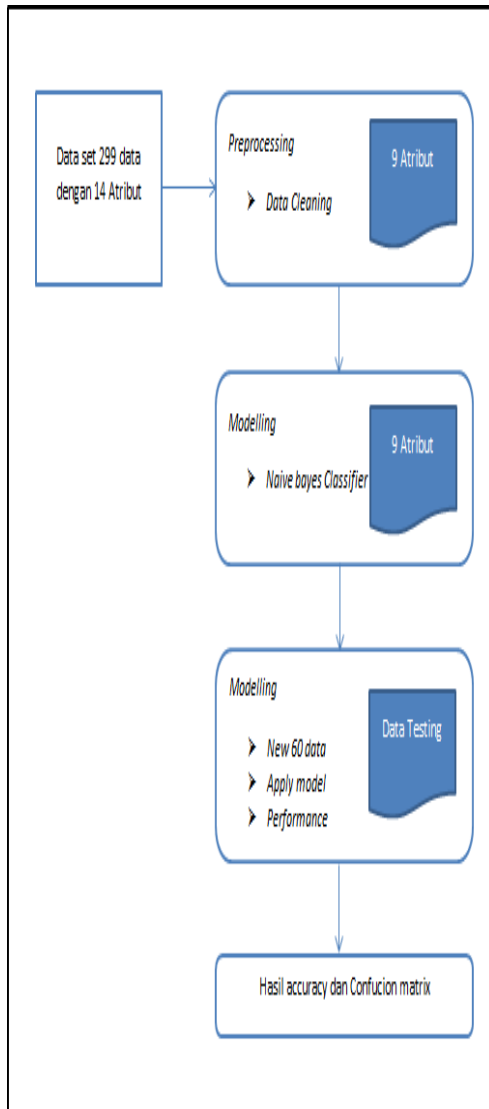
Gambar 1. Rancangan Penelitian

Sumber : modifikasi dari Saefudin Anwar (2013)

- a. Tahap awal yang akan dilakukan pada penelitian ini adalah melakukan studi literatur dengan menggunakan cara mempelajari suatu teori dan pengetahuan dasar mengenai semua yang berkaitan dengan penelitian ini agar dapat memahami dasar-dasar teori dan konsep – konsep yang mendukung penelitian.
- b. Tahap kedua yaitu adalah mengidentifikasi masalah, pada tahap ini proses yang dilakukan adalah dengan cara menggali permasalahan yang ditemukan pada objek yang diteliti serta mengidentifikasi kebutuhan yang dibutuhkan oleh pengguna untuk mencari alternatif solusi yang terkait dengan permasalahan tersebut.
- c. Tahap ketiga yaitu adalah pengumpulan data, setelah tahap identifikasi masalah dilakukan dan semua kebutuhan sudah didapatkan, maka tahap selanjutnya yang harus dilakukan adalah mengumpulkan data. Data yang digunakan adalah data diri calon debitur PT. Bank Rakyat Indonesia. Sumber data didapatkan dari PT. Bank Rakyat Indonesia.
- d. Tahap keempat yaitu adalah membuat data *training* dari data – data yang sudah dikumpulkan, karena tidak semua data dibutuhkan untuk pengujian model.
- e. Tahap kelima yaitu adalah membuat *data mining* yang digunakan untuk memilih teknik dan algoritma yang sesuai untuk menentukan pola yang terbentuk dan tersembunyi dari data calon debitur PT. Bank Rakyat Indonesia dengan algoritma *naive bayes*.
- f. Tahap keenam adalah hasil prediksi dari penerapan *data mining* dengan menggunakan algoritma *naive bayes*.
- g. Tahap ketujuh adalah evaluasi, pada tahap ini dilakukannya evaluasi untuk mengetahui apakah teknik dan algoritma yang digunakan sudah sesuai digunakan dalam penelitian sehingga bisa menemukan pola hubungan calon debitur yang pantas mendapatkan kredit dari PT. Bank Rakyat Indonesia.

4. Desain Penelitian/ Metodologi

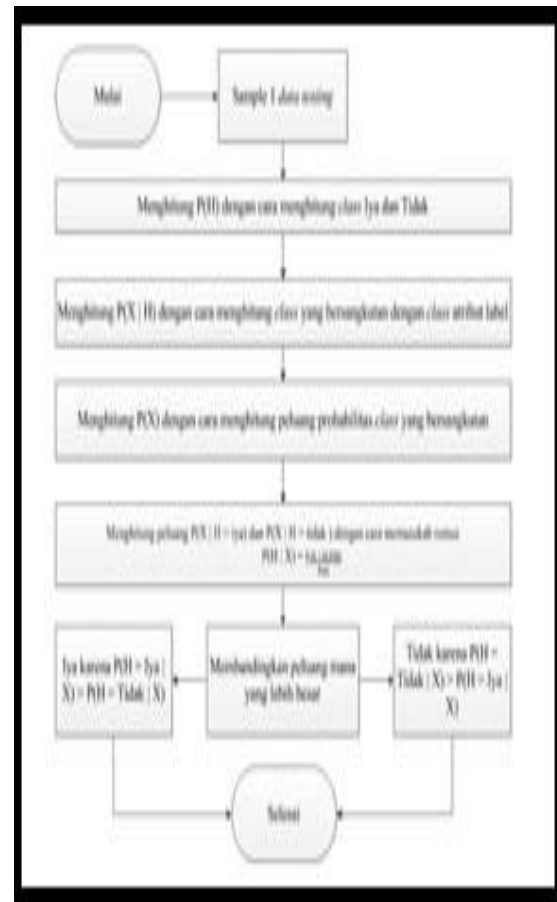
Metode yang akan digunakan pada penelitian ini adalah algoritma *naive bayes*. Dalam permodelan algoritma *naive bayes* akan dicari performanya yaitu *performance vector (accuracy)* dan *confusion matrix*. Data yang digunakan sudah melalui proses *learning data*, untuk melakukan pengukuran dalam penelitian ini menggunakan *tool rapidminer*.



Gambar 2. Model dan Tahap menemukan *accuracy* dan *confusion matrix*
 Sumber : modikasi dari Oman Somantri dan Slamet Wiyono (2017)

Dalam proses pengujian *data set* pertama dilakukan melakukan proses *praprosesing* yaitu data mentah (data yang benar – benar asli dari fakta) akan dilakukan pembersihan dengan cara *data cleaning*, dari 14 *field* menjadi 9 *field*. Setelah dilakukan *data cleaning*, maka data yang sudah diproses tersebut siap diujidengan menggunakan *data testing*. Pada pengujian kali ini, proses inimenggunakan sample 60 *data testing*, selanjutnya memasukkan operator pada *rapidminer* yaitu *apply model* dan *performance*. *Apply model* itu berfungsi untuk pengujian *data testing* yang nantinya akan diuji. Pada *Apply model* juga terdapat *data training* sebagai patokan untuk perhiungan pada *data testing*. *Performance* merupakan proses untuk mengevaluasi keakuratan prediksi dari model.

Permodelan Algoritma Naive Bayes



Gambar 3. Model tahap perhitungan naive bayes
 Sumber : modikasi dari Oman Somantri dan Slamet Wiyono (2017)

Dalam permodelan algoritma *naive bayes* akan dicari prediksi dari *field* “keputusan” yaitu dengan cara menentukan prior dari *class* yang ada dan *likelihood* yang ada. Dengan kemungkinan dari *evidence*, maka *posterior* yang tadi akan diketahui dari peluang – peluang yang telah dihasilkan.

5. Hasil Penelitian Dan Pengujian

a. Penggunaan Data pada Aplikasi RapidMiner

Data yang digunakan merupakan dataset calon debitur PT. Bank Rakyat Indonesia pada 2017 – 2018 yang terdiri dari 299 data. Data yang digunakan dalam aplikasi adalah data training dan data testing. Dari 299 data tersebut akan dijadikan 150 data training, 149 data testing yang akan diprediksi dan untuk menguji keakuratan.

b. Pembahasan Tentang Hasil Prediksi

Data prediksi dikarenakan untuk mengetahui pola dari suatu keputusan. Semakin

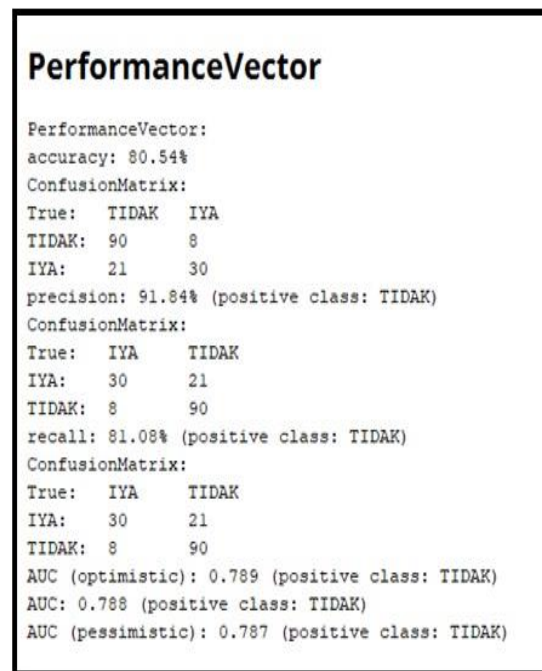
banyak data, maka semakin akurat hasil yang diprediksi. Ketika 149 *data testing* yang ingin diprediksi, maka kita harus memprediksi data tersebut dengan langkah yaitu sebagai berikut

Row No.	KEPUTUSAN	prediction	precision	recall	JENIS KEL.	STATUS PE.	JENIS
1	TIDAK	IYA	0.999	0.001	LAKI-LAKI	BELUM MENIKAH	TOKOH
2	TIDAK	IYA	1.000	0.000	LAKI-LAKI	BELUM MENIKAH	TOKOH
3	TIDAK	IYA	1.000	0.000	LAKI-LAKI	BELUM MENIKAH	TOKOH
4	TIDAK	IYA	0.624	0.376	LAKI-LAKI	MENIKAH	TOKOH
5	TIDAK	IYA	1.000	0.000	LAKI-LAKI	BELUM MENIKAH	TOKOH
6	TIDAK	IYA	0.525	0.475	LAKI-LAKI	BELUM MENIKAH	KAWAN
7	TIDAK	TIDAK	0.000	0.000	LAKI-LAKI	BELUM MENIKAH	TOKOH
8	TIDAK	TIDAK	0.107	0.893	PEREMPUAN	BELUM MENIKAH	KAWAN
9	TIDAK	TIDAK	0.004	0.916	LAKI-LAKI	MENIKAH	TOKOH
10	TIDAK	TIDAK	0.114	0.886	LAKI-LAKI	BELUM MENIKAH	TOKOH
11	TIDAK	TIDAK	0.147	0.853	LAKI-LAKI	BELUM MENIKAH	TOKOH
12	IYA	IYA	1.000	0.000	PEREMPUAN	BELUM MENIKAH	TOKOH
13	TIDAK	TIDAK	0.154	0.846	PEREMPUAN	BELUM MENIKAH	KAWAN
14	TIDAK	TIDAK	0.121	0.879	PEREMPUAN	BELUM MENIKAH	TOKOH
15	TIDAK	TIDAK	0.122	0.878	LAKI-LAKI	BELUM MENIKAH	KAWAN
16	IYA	IYA	1.000	0.000	LAKI-LAKI	BELUM MENIKAH	TOKOH

Gambar 5. Perbandingan antara data aktual dan data prediksi

Setelah perbandingan aktual dan prediksi dihitung sesuai golongannya, maka langkah selanjutnya adalah proses *confusion matrix*. Pada proses *confusion matrix* pada penelitian ini menampilkan nilai *accuracy*, *recall* dan *precision*. Maka hasil tabel *confusion matrix* pada penelitian ini adalah

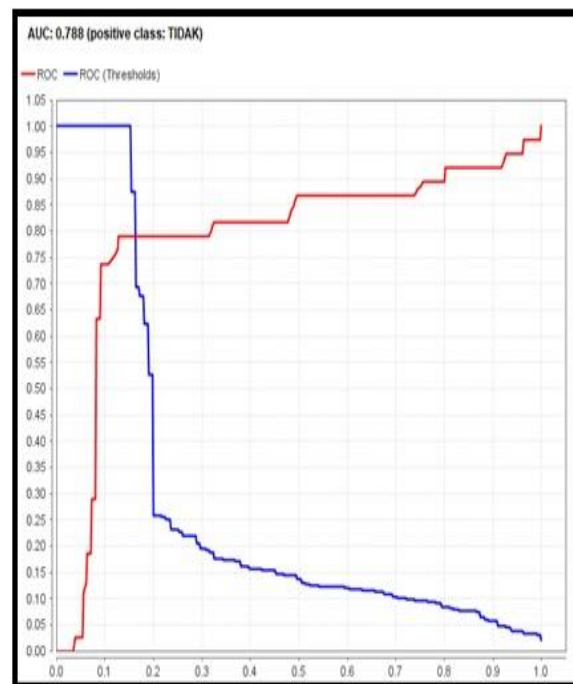
Dari perhitungan yang sebenarnya terlihat sama dengan perhitungan pada *rapidminer*. Hasil perhitungan *precision* yang diperoleh nilainya mencapai 91,84 % dan perhitungan *recall* sebesar 81,08 % pada *rapidminer*. Dari tabel tersebut sebanyak 149 data calon debitur, dengan hasil keputusan kredit disetujui ada 38 calon debitur dan tidak disetujui ada 117 calon debitur. Setelah dilakukan pengujian, diprediksi terdapat 51 calon debitur yang diprediksi keputusan disetujui dan 98 debitur yang diprediksi tidak disetujui.



Gambar 6. Deskripsi hasil *precision* dan *recall* pada *RapidMiner*

c. AUC (Area Under Curve)

Setelah dilakukan seluruh tahapan evaluasi untuk *confusion matrix*, maka selanjutnya dilakukan analisa selanjutnya yaitu perhitungan *Receiver Operating Characteristic* (ROC). Maka inilah yang disajikan kurva *Receiver Operating Characteristic* (ROC) yaitu sebagai berikut



Gambar 4. 8 Perhitungan yang disajikan ke dalam AUC (Area Under Curve)

Kurva ROC digunakan untuk mengekspresikan data *confusion matrix*. Dari gambar 4.9 dapat diketahui bahwa nilai AUC (*Area Under Curve*) dengan model algoritma *naive bayes* adalah 0,788. Hal ini menunjukkan bahwa model algoritma *naive bayes* mencapai klasifikasi yang baik. Maka akurasi tingkat diagnosa tersebut digolongkan ke dalam *fair classification* karena nilainya berkisar antara 0,70 sampai 0,80 atau 70% sampai dengan 80% akurasi yang didapatkan.

6. Kesimpulan

Kesimpulan yang dapat diambil dari penelitian ini adalah sebagai berikut

1. Analisa data mining dengan metode klasifikasi dan algoritma *naive bayes* dapat mengetahui pola calon debitur yang keputusan diditerima atau tidaknya pengajuan kredit pada PT. Bank Rakyat Indonesia.
2. Algoritma *naive bayes* memiliki tingkat akurasi yang bagus yaitu 78,88 % dengan begitu algoritma ini bisa diterapkan dalam memprediksi calon debitur pada PT. Bank Rakyat Indonesia.
3. Pola calon debitur pada PT. Bank Rakyat Indonesia dapat digunakan sebagai salah satu alternatif landasan keputusan kredit oleh pihak manajemen PT. Bank Rakyat Indonesia untuk meningkatkan pertimbangan keakuratan data calon debitur

Saran yang dapat diberikan dalam penelitian adalah sebagai berikut:

1. Perlu dilakukan percobaan untuk memakai algoritma selain *naive bayes* seperti *decision tree*, *support machine* dan lain sebagainya.
2. Perlu dilakukan pengembangan metode selain klasifikasi yaitu seperti klastering dan asosiasi. Setiap metode diuji tingkat akurasinya yaitu baik dari akurasi metode pada data mining itu sendiri maupun uji akurasi antara aktual dan prediksi.
3. Pada penelitian selanjutnya diperlukan penambahan aplikasi untuk mempermudah dalam menganalisa calon debitur pada PT. Bank Rakyat Indonesia.

Daftar Pustaka

Adi, Sumarni, & Winarko, Edi. 2015. Klasifikasi Data NAP (Nota Analisis Pembiayaan) Untuk Prediksi Tingkat Keamanan Pemberian Kredit Studi Kasus : Bank Syariah Mandiri Cabang Luwuk Sulawesi Tengah. Tidak Diterbitkan. Universitas Gajah Mada : Yogyakarta. Termuat Di

<https://jurnal.ugm.ac.id/ijccs/article/view/6635> (Diakses Tanggal 15 September 2018)

Agus Sugianto, Castaka 2015. Analisis Komparasi Algoritma Klasifikasi Untuk Menangani Data Tidak Seimbang Pada Data Kebakaran Hutan. Tidak Diterbitkan. Politeknik TEDC Bandung : Bandung. Termuat Di <https://publikasi.dinus.ac.id/index.php/technoc/article/view/992> (Diakses Tanggal 15 September 2018)

Agustianto, Khafidurrohman Dkk 2015. TEACHER MODELING UNTUK Mendukung Adaptive Learning Dalam Proses Pembelajaran Face To Face Learning Environments (KASUS: KELAS X BIDANG STUDI KEAHLIAN TIK SMK NEGERI DI KABUPATEN PONOROGO).

Tidak Diterbitkan. Universitas Gajah Mada : Daerah Istimewa Yogyakarta. Termuat Di <https://ojs.amikom.ac.id/index.php/semansteknimedia/article/view/1012> (Diakses Tanggal 15 September 2018)

Aprilla, Dennis Dkk 2013. BELAJAR DATA MINING DENGAN RAPIDMINER. Tidak Diterbitkan. Jakarta. Termuat Di <https://id.scribd.com/doc/294428531/Belajar-Data-Mining-Dengan-Rapidminer-Pdf> (Diakses Tanggal 15 September 2018)

Asriningtias, Yuli, & Mardhiyah, Rodhyah 2014. APLIKASI DATA MINING UNTUK MENAMPILKAN INFORMASI TINGKAT KELULUSAN. Tidak

Diterbitkan, Universitas Teknologi Yogyakarta : Yogyakarta. Termuat Di <http://journal.uad.ac.id/index.php/jifo/article/view/2082> (Diakses Tanggal 15 September 2018)

Bogor Institution. (2013). Pengertian Dan Konsep Data Mining – GSB-IPB. Retrieved From <http://gsbipb.com/?P=821>. Termuat Di <https://www.mendeley.com/newsfeed/> (Diakses Tanggal 15 September 2018)

Brookshier, D. (2014). DWH- Data Warehouse. Tutorialpoint, 99. Termuat Di <https://www.mendeley.com/newsfeed/> (Diakses Tanggal 15 September 2018)

- Defiyanti, Sofi, & Jajuli, Mohamad. 2015. Integrasi Metode Klasifikasi Dan Clustering Dalam Data Mining. Tidak Diterbitkan. Universitas Singaperbangsa Karawang : Karawang. Termuat Di https://www.researchgate.net/publication/314266899_Integrasi_Metode_Klasifikasi_Dan_Clustering_Dalam_Data_Mining (Diakses Tanggal 15 September 2018)
- Fadillah, Annisa Paramitha 2015. Penerapan Metode CRISP-DM Untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ). Tidak Diterbitkan, UNIKOM Bandung : Bandung. Termuat Di <https://media.neliti.com/media/publications/133327-ID-Penerapan-Metode-Crisp-Dm-Untuk-Prediksi.Pdf> (Diakses Tanggal 15 September 2018)
- Homayouni, H., Ghosh, S., & Ray, I. (2018). Data Warehouse Testing. *Advances In Computers*. <https://doi.org/10.1016/Bs.adcom.2017.12.005>. <https://www.mendeley.com/newsfeed/> (Diakses Tanggal 15 September 2018)
- Iskandar, Derick, & K. Suprapro, Yoyon. 2013. Perbandingan akurasi klasifikasi tingkat kemiskinan antara algoritma C4.5 dan Naïve Bayes Classifier. Tidak diterbitkan. Institut Teknologi Sepuluh Nopember : Surabaya. Termuat di <http://javajournal.its.ac.id/index.php/java/article/view/10> (Diakses tanggal 15 September 2018)
- Larose, 2014. *Discovering Knowledge in Data : An Introduction to Data Mining*. Termuat di https://doc.lagout.org/Others/Data%20Mining/Discovering%20Knowledge%20in%20Data_%20An%20Introduction%20to%20Data%20Mining%20%282nd%20ed.%29%20%5BLarose%20%26%20Larose%202014-06-30%5D.pdf (Diakses tanggal 15 September 2018)
- Loshin, D. (2013). *Business Intelligence. Business Intelligence*. <https://doi.org/10.1016/C2010-0-67240-3>. Termuat di <https://www.mendeley.com/newsfeed/> (Diakses tanggal 15 September 2018)
- Meilani, B. D., & Susanti, N. (2015). Aplikasi Data Mining Untuk Menghasilkan Pola Kelulusan Siswa Dengan Metode Naïve Bayes. *Jurnal Ilmiah NERO*, 1(3), 182–189. Termuat di <https://www.mendeley.com/newsfeed/> (Diakses tanggal 15 September 2018)
- Muktamar, Burhan Alifroni dkk 2015. PEMBOBOTAN KORELASI PADA NAÏVE BAYES CLASSIFIER. Tidak diterbitkan. Universitas Gadjah Mada : Daerah Istimewa Yogyakarta. Termuat di <https://ojs.amikom.ac.id/index.php/semnasteknomedia/article/download/814/780> (Diakses tanggal 15 September 2018)
- Nia, Nuraeni. 2017. Pentuan Kelayakan Kredit Dengan Algoritma Naive Bayes Classifier. Studi Kasus Bank Mayapada Mitra Usaha Cabang PGC. Tidak diterbitkan. AMIK BSI. Termuat di <https://ejournal.bsi.ac.id/ejurnal/index.php/jtk/article/view/1337> (Diakses tanggal 15 September 2018)
- Novita Putri, Astrid. 2017. Penerapan Naive Bayesian untuk Perangkingan Kegiatan di Fakultas TIK Universitas Semarang. Tidak diterbitkan. Universitas Semarang : Semarang. Termuat di <http://jurnal.umk.ac.id/index.php/simet/article/view/1545> (Diakses tanggal 15 September 2018)
- Ridwan, M., Suyono, H., & Sarosa, M. (2013). Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier, 7(1), 59–64. Termuat di <https://www.mendeley.com/newsfeed/> (Diakses tanggal 15 September 2018)
- Suyanto.2017. Data mining untuk klasifikasi dan klasterisasi data. Bandung. Informatika.
- Warehouse, D., Warehousing, D., Ware-, D., & Tantawi PhD, R. (2016). *Data warehouse. Salem Press Encyclopedia*, (6). Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&site=eds-live&db=ers&AN=90558282> OP - Salem Press Encyclopedia, January, 2016.%5Cn<http://widgets.ebscohost.com/prod/customerspecific/ns000290/authentication/index.php?url=http%253A%252F%252Fsearch.ebscohost.com> Termuat di <https://www.mendeley.com/newsfeed/> (Diakses tanggal 15 September 2018)

Wikipedia, F. (2016). Data mining. <https://doi.org/10.1007/978-3-642-21533-9> (Diakses tanggal 15 September 2018)

Wikipedia. (2018). Sejarah dan Company Profile PT. Bank Rakyat Indonesia. Termuat di https://id.wikipedia.org/wiki/Bank_Rakyat_Indonesia (Diakses tanggal 15 September 2018)