

**KLASIFIKASI KELAYAKAN PESERTA
PENGAJUAN KREDIT RUMAH DENGAN ALGORITMA NAÏVE BAYES DI PERUMAHAN
AZZURA RESIDENCIA**

Karsito¹⁾, Santi Susanti²⁾

Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa
karsito_awb@pelitabangsa.ac.id

Disetujui, 25 Maret 2019

Abstraksi

Rumah adalah kebutuhan yang sangat mendasar bagi kelangsungan hidup manusia belakangan ini banyak sekali pengembang yang bergerak di bidang property berupa perumahan. Azzura Residencia hadir sebagai perumahan syariah tanpa riba, di kembangkan oleh perusahaan Azzura Griya Utama yang menjalankan usaha sejak tahun 2013. Banyaknya calon peserta pengajuan kredit rumah yang berasal dari latar belakang ekonomi yang berbeda-beda dan belum diketahui layak atau tidak untuk menjadi peserta kredit. dengan memanfaatkan data tersebut peneliti ingin menerapkan salah satu teknik data mining dengan perhitungan statistika dalam melakukan klasifikasi kelayakan kredit rumah. Metode yang digunakan adalah Naive Bayes dengan menggunakan sebanyak 250 data peserta kredit rumah. Prediksi kelayakan kredit menggunakan metode Naive Bayes merupakan aplikasi juga bertujuan membantu tim marketing untuk lebih cepat dalam mengambil keputusan layak atau tidak layak. Hasil analisis menunjukkan bahwa pekerjaan, penghasilan, jumlah tanggunga, harga rumah, DP, lama kredit bisa menjadi indikator untuk pengambilan keputusan dalam menentukan layak atau tidak layaknya calon peserta kredit. Hasil analisis juga menunjukkan bahwa ketepatan klasifikasi peserta kredit menggunakan metode naive bayes pada penelitian ini adalah sebesar 97.33%.

Kata kunci : Rumah, *Data Mining*, *Naive Bayes*

Abstract

The house is a very basic need for human survival lately a lot of developers engaged in property in the form of housing. Azzura Residencia is present as sharia housing without usury, developed by the Azzura Griya Utama company which has been in business since 2013. Many prospective applicants apply for home loans from different economic backgrounds and are not yet known to be eligible or not to become credit participants. by utilizing the data the author wants to apply one of the data mining techniques with static calculations in doing the classification of home credit worthiness. The method used is Naive Bayes by using 250 data on home loan participants. The credit worthiness prediction using the Naive Bayes method is an application that also aims to help the marketing team more quickly in making feasible or improper decisions. The results of the analysis show that employment, income, amount of interest, house price, DP, credit duration can be indicators for decision making in determining whether or not a candidate is worthy. The analysis also shows that the accuracy of the classification of credit participants using the naive bayes method in this study amounted to 97.33%.

Keywords: *Houses, Data Mining, Naive Bayes*

1. Pendahuluan

Azzura Residencia hadir sebagai perumahan syariah tanpa riba, di kembangkan oleh perusahaan Azzura Griya. Permasalahan yang sering timbul dalam perusahaan pengembang perumahan yang memberikan kredit adalah banyaknya calon peserta pengajuan kredit rumah yang berasal dari latar belakang ekonomi yang berbeda-beda dan belum diketahui layak atau tidak untuk menjadi peserta kredit rumah, kemudian dalam proses pengecekan data pun membutuhkan waktu yang cukup lama karna masih dilakukan dengan cara manual.

Penggunaan teknik data mining diharapkan mampu memberikan informasi yang berguna tentang teknik klasifikasi data peserta yang akan bergabung dalam kelompok kriteria yang kreditnya layak dan kriteria yang kreditnya tidak layak. Dengan menggunakan teknik data mining ini juga diharapkan dapat memberi gambaran mengenai target pemasaran kredit perumahannya di masa yang akan datang.

2. Tinjauan Studi

2.1 Kredit Rumah

Pengertian kredit dalam buku Seri Manajemen Bank No. 5 (1997: 31) adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara Bank dengan pihak lain yang mewajibkan peminjam untuk melunasi utangnya setelah jangka waktu tertentu dengan jumlah bunga, imbalan, atau pembagian hasil keuntungan. Selain itu, kredit juga bisa berarti kemampuan untuk melaksanakan suatu pembelian atau mengadakan suatu pinjaman dengan suatu janji pembayarannya akan dilakukan atau ditanggguhkan pada suatu jangka waktu yang disepakati.

2.2

ata Mining

Data mining adalah kegiatan yang meliputi pengumpulan, pemakaian data historis yang menemukan keteraturan. Maksud dari pengertian ini yaitu proses pencarian informasi yang tidak diketahui sebelumnya dari sekumpulan data yang besar.

D

2.3 Naïve Bayes

Naïve Bayes merupakan sebuah pengklasifikasian probalistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan *Teorema Bayes* dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variable kelas. *Naïve Bayes* juga didefinisikan sebagai pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan inggris Thomas Bayes yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya (Saleh, 2015).

2.4 Split Validation

Split Validation adalah model pengklasifikasian Naïve Bayes yang bertugas membagi dua data menjadi dua subset data yaitu data training dan testing. Data *training* merupakan data yang digunakan untuk pelatihan, sedangkan data testing akan digunakan untuk pengujian. Adapaun untuk melihat secara lebih jelas dari model *split validation* dapat dilihat pada gambar berikut.

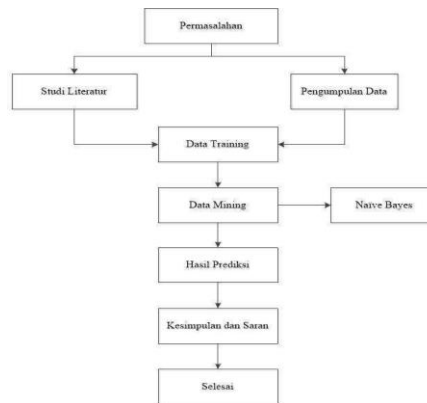
2.5 Confusion Matrix

Confusion matrix merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi. Pada dasarnya confusion matrix mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya. Pada pengukuran kinerja menggunakan *confusion matrix*, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah True Positive (TP), True Negative (TN), False Positive (FP) dan False Negative (FN). Nilai True Negative (TN) merupakan jumlah data negatif yang terdeteksi dengan benar, sedangkan False Positive (FP) merupakan data negatif namun terdeteksi sebagai data positif.

3. Kerangka Konsep

Kerangka pemikiran dalam penelitian ini berisi landasan teori yang menjadi dasar dalam menjawab tujuan penelitian. Teori yang diuraikan meliputi konsep dasar dari metode naïve bayes beserta Teknik yang digunakan untuk mengetahui kelayakan pengajuan calon peserta kredit rumah. Dan menggunakan rapid miner untuk mengetahui akurasi prediksi kelayakan pengajuan calon peserta kredit rumah.

Berdasarkan kerangka teori yang telah dijelaskan pada gambaran umum objek, maka dikembangkan kerangka pemikiran penelitian prediksi kelayakan pengajuan calon peserta kredit rumah yang dipengaruhi oleh jenis pekerjaan dan besar pendapatan perbulan dari calon peserta pengajuan kredit rumah.



Gambar 1. Kerangka Pemikiran

Sementara itu, True Positive (TP) merupakan data positif yang terdeteksi benar. False Negative (FN) merupakan kebalikan dari True Positive, sehingga data positif, namun terdeteksi sebagai data negatif (Untari, 2010). Ada 6 tahap, yaitu :

1. Tahap pertama yang dilakukan adalah menggali permasalahan yang ditemukan pada objek yang diteliti guna mencari alternatif solusi yang terkait dengan permasalahan tersebut.
2. Tahap kedua yang dilakukan dalam melakukan penelitian yaitu melakukan studi literatur dengan cara mempelajari teori dan pengetahuan dasar mengenai semua yang berkaitan dengan penelitian ini agar dapat memahami dasar – dasar teori dan konsep – konsep yang mendukung penelitian.
3. Tahap ketiga yaitu pengumpulan data, setelah tahap studi literatur maka tahap selanjutnya yang harus dilakukan adalah mengumpulkan data. Data yang dibutuhkan adalah data peserta kredit pada Perumahan Azzura Recidensia.
4. Tahap keempat yaitu membuat data training / data latih dari data – data yang sudah dikumpulkan, karena tidak semua data dibutuhkan untuk pengujian model.
5. Tahap kelima adalah membuat data mining yang digunakan untuk memilih teknik dan algoritma yang sesuai untuk menemukan pola dari data peserta kredit tersebut dengan menggunakan algoritma naive bayes.
6. Tahap keenam adalah hasil prediksi dari penerapan data mining dengan algoritma naive bayes. Tahap ketujuh merupakan kesimpulan yang didapat dari penelitian tersebut serta saran.

4. Desain Penelitian/Methodologi

Peneliti melakukan penelitian di kantor marketing perumahan Azzura Recidensia yang berlokasi di Jl. Cisaat setu, Telajung Cikarang Barat, Bekasi, Jawa Barat. Azzura Residencia hadir sebagai perumahan syariah tanpa riba, di kembangkan oleh perusahaan Azzura Griya Utama yang menjalankan usaha sejak tahun 2013, mengikuti perkembangan pasar dan terus menerus mengalami peningkatan yang sangat signifikan di 3 tahun terakhir. Pada tahun 2016 Azzura Griya Utama melahirkan produk syariah tanpa bank perdananya yang terjual dalam waktu singkat, dan menjadi market leader dan trend setter dari dari pemain property syariah khususnya di wilayah Bekasi.

Peneliti mendapatkan data dari kantor marketing Azzura Recidensia berupa data peserta kredit. Data tersebut ke mudian dianalisis guna mendapatkan data yang spesifik dan menuangkan data yang di dapatkan dalam bentuk excel, guna mempermudah pengolahan data. Total data yang diambil sebanyak 250 kasus, peserta kredit yang pengajuan kreditnya dinyatakan layak sebanyak 214 orang dan yang pengajuannya tidak layak sebanyak 36 orang.

Tahapan praproses adalah sebuah proses awal dalam melakukan pengolahan data yaitu dengan menyiapkan data yang diperoleh dari hasil wawancara dan pembacaan dokumen peserta kredit. Data yang telah diperoleh tersebut akan diseleksi menjadi beberapa informasi penting yang akan digunakan oleh tim marketing dalam melakukan prediksi kelayakan kredit rumah yaitu dari kriteria kelayakan kredit dan faktor yang mempengaruhi kelayakan kredit antara lain latar belakang pekerjaan yang beragam, pendapatan perbulan, pengeluaran perbulan, status pernikahan, jumlah tanggungan keluarga, harga rumah yang akan diambil,, uang muka yang akan dibayarkan dan lama kredit. Kriteria kelayakan kredit tersebut selanjutnya digunakan sebagai data input prediksi dengan rapidminer.

Model yang diusulkan untuk klasifikasi menggunakan algoritma Naive bayes adalah menggunakan model *split validation*. *Split validation* membagi data menjadi dua subset data yaitu data *training* dan data

testing. Data *training* merupakan data yang digunakan untuk pelatihan, sedangkan data testing akan digunakan untuk pengujian.

Tahapan validasi yang dilakukan dalam penelitian ini adalah untuk memberikan penilaian dari hasil penggunaan algoritma *naive bayes* saja dan *naive bayes* yang disertai dengan *confusion matrix* untuk mengklasifikasi prediksi kelayakan peserta kredit rumah. Bagian yang akan divalidasi adalah presentase data, jumlah data training, jumlah data testing dan nilai akurasi yang di hasilkan.

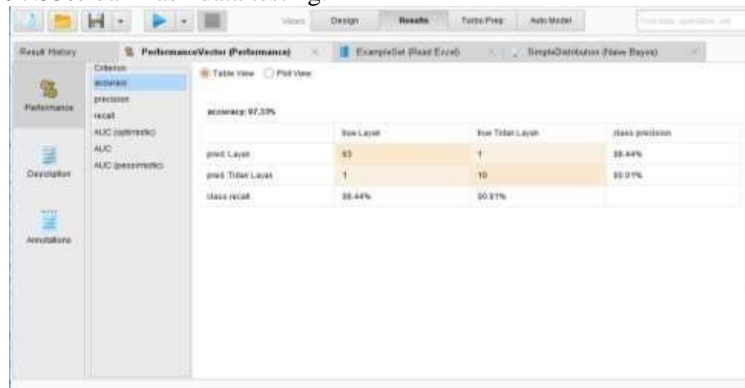
5. Hasil Penelitian dan Pembahasan

5.1 Hasil Penelitian

Proses klasifikasi dengan *rapidminer* dengan metode *naive bayes* yang digunakan mengklasifikasi data peserta kredit pada penelitian ini sehingga diperoleh nilai *Accuracy*, *Precision* dan *Recall*.

a) *Accuracy* / akurasi

Dengan mengetahui jumlah data yang di klasifikasikan secara benar maka dapat diketahui akurasi hasil prediksi yaitu 97.33% dari hasil data testing.



Gambar 2. Accuracy

Perhitungan secara manual

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100\%$$

$$= \frac{(63 + 10)}{63 + 10 + 1 + 1} \times 100\% = \frac{73}{75} \times 100\% = 0.973 \times 100\%$$

$$= 97.33\%$$

b) *Precision*

Precision adalah jumlah data yang *true positive* (jumlah data positif yang dikenali secara benar sebagai positif) dibagi dengan jumlah data dikenali sebagai positif. Dari hasil pengujian nilai *precision* yaitu 98.44% untuk *class* Positif dan nilai *precision* 90.91% untuk *class* Negatif.



Gambar 3. Precision

Perhitungan secara manual

$$\begin{aligned} \text{Precision (Positive)} &= \frac{TP}{(TP+FP)} \times 100\% \\ &= \frac{63}{(63+1)} \times 100\% = \frac{63}{64} \times 100\% = 0.9844 \times 100\% = 98.44\% \end{aligned}$$

$$\begin{aligned} \text{Precision (Negative)} &= \frac{TN}{(TN+FN)} \times 100\% \\ &= \frac{10}{(10+1)} \times 100\% = \frac{10}{11} \times 100\% = 0.9091 \times 100\% = 90.91\% \end{aligned}$$

c) **Recall**

Recall merupakan jumlah data yang *true* idibagi dengan jumlah data yang sebenarnya *positive* (*true positive* + *true negative*). Untuk nilai *recall* yaitu 100% pada *class* positif dan nilai *recall* 90.91% pada *class* negatif.

	Real Layak	Real Tidak Layak	Class precision
pred Layak	63	1	98.44%
pred Tidak Layak	1	10	90.91%
Class recall	98.44%	90.91%	

Gambar 4.8 Recall
Perhitungan secara manual

$$\begin{aligned} \text{Recall (Positive)} &= \frac{TP}{(TP+FN)} \times 100\% \\ &= \frac{63}{(63+1)} \times 100\% = \frac{63}{64} \times 100\% = 0.9844 \times 100\% = 98.44\% \end{aligned}$$

$$\begin{aligned} \text{Recall (Negative)} &= \frac{TN}{(TN+FP)} \times 100\% \\ &= \frac{10}{(10+1)} \times 100\% = \frac{10}{11} \times 100\% = 0.9091 \times 100\% = 90.91\% \end{aligned}$$

6. Kesimpulan

Kesimpulan yang dapat diambil dari penelitian ini adalah :

1. Dengan menggunakan data mining metode klasifikasi dan algoritma naïve bayes dengan tools rapidminer dapat mengetahui klasifikasi kelayakancalon peserta kredit.
2. Dengan menggunakan data mining metode klasifikasi dan algoritma naïve bayes dengan tools rapidminer dapat mempercepat pengambilan keputusankelayakan calon peserta kredit.
3. Dari hasil penelitian didapatkan Accuracy sebanyak 97.33%, Precision Positif / Layak 98.44%, Precision Negatif / Tidak Layak sebanyak 90.91%, Recall Positif / Layak sebanyak 90.91%, Recall Negatif / Tidak Layak sebanyak 98.44%.

Daftar Pustaka

Agung Kurniadi Hermawan, Selvia Lorena Br G. 2017. Aplikasi Data Mining Menggunakan Naïve Bayes Classifier untuk Persetujuan Pengajuan Kredit. Bandung.

Bustami “Penerapan Algoritma Naive Bayes untuk mengklasifikasi Data Nasabah Asuransi”. Jurnal Penelitian Teknik Informatika (TECHSI) Vol. 2 : 2, 127-146, 2013 (<http://journal.uad.ac.id/index.php/JIFO/article/view/2086>).

- C, Dennis Aprilla, dkk. 2013. Belajar Data Mining dengan Rapidminer. Diakses 10 Oktober 2018
- Cahya, Intan G. 2014. Prediksi Persediaan Obat dengan Metode Naive Bayes. Fakultas Komunikasi dan Informatika Universitas Muhammadiyah Surakarta : Surakarta.
- Connoly, T., & Begg. C. 2005. Database Systems ; A Practical Approach to Design, Implementation and Management, Harlow : Addison Wesley.
- Dedy Ahmad Kurniawan, Danny Kriestanto. Penerapan Naive Bayes untuk Prediksi Kelayakan Kredit. Jurnal Informatika dan Komputer (JIKO) Vol.1 No. 1. 2016 (<https://ejournal.akakom.ac.id/index.php/jiko/article/view/10/3>).
- Fadli, A. 2011. Konsep Data Mining. Sumber : IlmuKomputer.com (diakses pada 07 September 2018)
- Hermawati, F.A. 2013. Data Mining. Andi : Yogyakarta.
- Jananto Arief. Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa. Jurnal Teknologi Informasi DINAMIK Vol. 18, No. 1 2013 (<https://www.unisbank.ac.id/ojs/index.php/fti1/article/view/1669>).
- Kusrini dan Luthfi, Emha Taufiq. 2010. Algoritma Data Mining. Andi : Yogyakarta.
- Larose, Daniel T. 2005. Discovering Knowledge in Data : An Introduction to Data Mining.
- Mahadewi. M. Ch. Agung Ayu Bulan. 2015. Klasifikasi Kain Tradisional Nusantara dengan menggunakan Algoritma Naive Bayes Classifier. Universitas Sanata Dharma : Yogyakarta.
- Moertini, V. 2002. Data Mining sebagai Solusi Bisnis. Penelitian Staff Pengajar Jurusan Ilmu Komputer Universitas Katolik Parahyangan : Bandung.
- Muslehatin Wiwik, dkk. 2017. Penerapan Naive Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Sistem Informasi.
- UIN Suska Riau. Jurnal SNTKI 9. 2017 : Pekanbaru, Riau (<http://ejournal.uin-suska.ac.id/index.php/SNTIKI/article/view/3276>).
- Nuraeni, Nia. Penentuan Kelayakan Kredit Dengan Algoritma Naive Bayes Classifier. Jurnal Teknik Komputer AMIK BSI vol. III, No. 1. 2017 (<https://ejournal.bsi.ac.id/ejurnal/index.php/jtk/article/view/1337.pdf>).
- Ridwan M, Suyono H dan Sarosa M. Penerapan Data Mining untuk evaluasi kinerja Akademik Mahasiswa menggunakan Algoritma Naive Bayes Classifier. Jurnal EECCIS Vol. 7 No. 1. 2013 (<https://media.neliti.com/media/publications/61002-ID-penerapan-data-mininguntuk-evaluasi-kin.pdf>).
- Sitanggang, Sartika. 2015. Pengkategorian Buku untuk Fakultas Teknik Menggunakan Metode Naive Bayes Classifier di Perpustakaan Universitas Widyatama : Bandung.