ISSN Cetak 2745-925X
ISSN Online 2722-273X

JESSI Volume 01 Nomor 2 November 2020

PREDIKSI KELAYAKAN KREDIT KEPEMILIKAN RUMAH MENGGUNAKAN METODE K-NN BERBASIS FORWARD SELECTION

Rachmat
Rachmat27udinus@gmail.com
Politeknik Informatika Nasional (POLINAS) Makasar

Received : 16 June 20
Accepted : 19 Nov 20
Published : 22 Nov 20

Abstract

Everyone wants to have a house or private residence as a necessity to meet the needs of the board in accordance with the government's program for the welfare of its people with the Ministry of Housing, in line with business growth, credit is an interesting problem to study. Some computer research to reduce credit risk has been done in order to avoid the destruction of a finance company. This paper discusses the Forward Selectin-based k-Nearest Neighbor (kNN) algorithm that is applied to customer data on Home Ownership Loans (KPR) The results of testing for the performance of this algorithm with X-Validation and obtained a Mean Squared Error (MSE) of 24822,417 +/- 13389,369 and Squared Error (SE) of 795427569,033 +/- 674548543,837 to increase the MSE and the SE.

Keywords: house, people, credit, k-NN, Forward Selection

Abstrak

Setiap orang ingin memilki rumah atau tinggal pribadi sebagai kebutuhan dalam memenuhi kebutuhan papan sesuai juga dengan program pemerintah untuk mensejahterakan rakyatnya ini dengan adanya Kementerian Perumahaan Rakyat, Sejalan dengan pertumbuhan bisnis, kredit merupakah masalah yang menarik untuk diteliti. Beberapa riset bidang komputer untuk mengurangi resiko kredit telah banyak dilakukan dalam rangka menghindarai kehancuran suatu perusahaan pembiayaan. Paper ini membahas algoritma k-Nearest Neighbor (kNN) berbasis Forward Selectin yang diterapkan pada data nasabah terhadap Kredit Pemilikan Rumah (KPR) Hasil testing untuk performa algoritma ini dengan X-Validation dan didapat Mean Squared Error (MSE) sebesar 24822.417 +/- 13389.369 dan Squared Error (SE) sebesar 795427569.033 +/- 674548543.837 untuk menigkatkan MSEnya dan SEnya.

Kata kunci: : rumah, rakyat, kredit, k-NN, Forward Selection

This is an open access article under the CC BY-SA license

ISSN Cetak 2745-925X
ISSN Online 2722-273X

JESSI Volume 01 Nomor 2 November 2020

1. Pendahuluan

Dari penelitian-penelitian yang pernah dilakukan, evaluasi resiko kredit merupakan masalah yang menarik dalam analisa keuangan. Penelitian mengenai analisis kelayakan pemberian kredit untuk konsumen khususnya kredit kepemilikan barang dengan metode klasifikasi data mining telah banyak dilakukan. Dalam penulisan ini akan dibahas mengenai penerapan algoritma k-nearest neighbor untuk penentuan resiko kredit pemilikan rumah [1]. Tujuan adanya kredit pemilikan rumah beberapa sumber, maksud dan tujuan diberikannya layanan kredit pemilikan rumah sudah jelas artinya membantu para nasabah yang ingin memiliki rumah tetapi tidak mempunyai uang secara cash/tunai dalam jumlah banyak. Tujuan tersebut agar lebih ditekankan pada kebutuhan primer karena rumah merupakan tempat untuk tinggal dan unutk melakukan kegiatan lain. KPR merupakan sarana fasilitator untuk mendapatkan suatu kredit khususnya rumah. Dalam upaya meningkatkan daya guna dan hasil guna tanah yang jumlahnya terbatas tersebut, terutama bagi pembangunan perumahan dan permukiman, serta mengefektifkan penggunaan tanah terutama didaerah-daerah yang berpenduduk padat, maka perlu adanya pengaturan, penataan, dan penggunaan atas tanah, sehingga bermanfaat bagi masyarakat banyak. Apalagi jika dihubungkan dengan hak asasi, maka tempat tinggal (perumahan dan permukiman) merupakan hak bagi setiap Warga Negara, sebagaimana diatur dalam Pasal 28 H ayat (1) Undang-Undang Dasar 1945. Kebutuhan dasar tersebut wajib dihormati, dilindungi, ditegakkan, dan dimajukan oleh Pemerintah

Pemerintah melalui kalangan perbankan menyediakan fasilitas kredit pemilikan rumah yang biasa disebut Kredit Pemilikan Rumah (KPR) [1]. Kredit Pemilikan Rumah (KPR) adalah suatu fasilitas kredit yang diberikan oleh perbankan kepada para nasabah perorangan yang akan membeli atau memperbaiki rumah. Pada saat ini di Indonesia dikenal ada 2 (dua) jenis KPR. Pembangunan Rumah Susun dimaksudkan untuk penyediaan hunian yang layak bagi orang dan badan hukum. Oleh karena itu perumahan itu harus memenuhi standar sebagai hunian yang memenuhi syarat baik dari segi kesehatan, kenyamanan, dan keasrian dari rumah tersebut. Pembangunan Rumah Susun merupakan pemenuhan atas kebutuhan papan (tempat tinggal) khususnya bagi masyarakat yang berpenghasilan rendah, sebagaimana yang dinyatakan dalam Pasal 5 Undang-Undang No.16 Tahun 1985 Undang-Undang Nomor 16 Tahun 1985 tentang Rumah Susun yang menyatakan bahwa Rumah Susun dibangun sesuai dengan tingkat keperluan dan kemampuan masyarakat terutama masyarakat yang berpenghasilan rendah. Pembangunannya dapat dilaksanakan atau diselenggarakan oleh Badan Usaha Milik Negara atau Daerah, Koperasi, atau Badan Usaha Milik Swasta yang bergerak dibidang itu.

KPR Subsidi yaitu suatu kredit yang diperuntukan kepada masyarakat berpenghasilan menengah ke bawah dalam rangka memenuhi kebutuhan perumahan atau perbaikan rumah yang telah dimiliki. Bentuk subsidi yang diberikan berupa : Subsidi meringankan kredit dan subsidi menambah dana pembangunan atau perbaikan rumah. Kredit subsidi ini diatur tersendiri oleh Pemerintah, sehingga tidak setiap masyarakat yang mengajukan kredit dapat diberikan fasilitas ini. Secara umum batasan yang ditetapkan oleh Pemerintah dalam memberikan subsidi adalah penghasilan pemohon dan maksimum kredit yang diberikan.

KPR Non Subsidi yaitu suatu KPR yang diperuntukan bagi seluruh masyarakat. Ketentuan KPR ditetapkan oleh bank, sehingga penentuan besarnya kredit maupun suku bunga dilakukan sesuai kebijakan bank yang bersangkutan. Jenis Kredit Pemilikan Rumah (KPR) yang diberikan oleh Bank Tabungan Negara.

Data Mining didefinisikan sebagai proses penemuan pola dalam data. Berdasarkan tugasnya, data mining dikelompokkan menjadi deskripsi, estimasi, prediksi, klasifikasi, clustering dan asosiasi. Proses dalam tahap data mining (Gambar 1.) terdiri dari tiga langkah Utama, yaitu :

a. Data Preparation

Pada langkah ini, data dipilih, dibersihkan, dan dilakukan preprocessed mengikuti pedoman dan knowledge dari ahli domain yang menangkap dan mengintegrasikan data internal dan eksternal ke dalam tinjauan organisasi secara menyeluruh.

b. Algoritma data mining

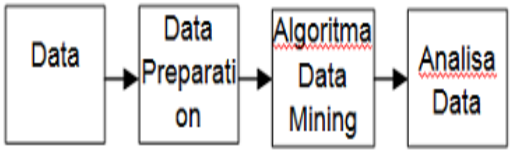
Penggunaan algoritma data mining dilakukan pada langkah ini untuk menggali data yang terintegrasi untuk memudahkan identifikasi informasi bernilai [2].

c. Fase analisa data

ISSN Cetak 2745-925X
ISSN Online 2722-273X

JESSI Volume 01 Nomor 2 November 2020

Keluaran dari data mining dievaluasi untuk melihat apakah knowledge domain ditemukan dalam bentuk rule yang telah diekstrak dari jaringan.

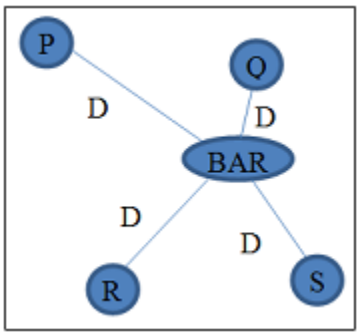


Gambar 1 Langkah-langkah dalam Proses Data Mining

K-nearest neighbor (kNN) termasuk kelompok instance-based learning. Algoritma ini juga merupakan salah satu teknik lazy learning. kNN dilakukan dengan mencari kelompok k objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing [3].

Contoh kasus, misal diinginkan untuk mencari solusi terhadap masalah seorang pasien baru dengan menggunakan solusi dari pasien lama. Untuk mencari solusi dari pasien baru tersebut digunakan kedekatan dengan kasus pasien lama, solusi dari kasus lama yang memiliki kedekatan dengan kasus baru digunakan sebagai solusinya.

Terdapat pasien baru dan 4 pasien lama, yaitu P, Q, R, dan S (Gambar 2). Ketika ada pasien baru maka yang diambil solusi adalah solusi dari kasus pasien lama yang memikili kedekatan terbesar.



Gambar2. ilustrasi kasus algoritma kNN

Misal D1 adalah jarak antara pasien baru dengan pasien P, D2 adalah jarak antara pasien baru dengan pasien Q, D3 adalah jarak antara pasien baru dengan pasien R, D4 adalah jarak antara pasien baru dengan pasien S. Dari ilustrasi gambar terlihat bahwa D2 yang paling terdekat dengan kasus baru. Dengan demikian maka solusi dari kasus pasien Q yang akan digunakan sebagai solusi dari pasien baru tersebut.

Ada banyak cara untuk mengukur jarak kedekatan antara data baru dengan data lama (data training), diantaranya euclidean distance dan manhattan distance (city block distance), yang paling sering digunakan adalah euclidean distance [4], yaitu:

$$\sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \ldots + (a_n - b_n)^2}$$

Dimana a = a_1,a_2, …, an, dan b = b_1, b_2, …,b_n mewakili n nilai atribut dari dua record. Untuk atribut dengan nilai kategori, pengukuran dengan euclidean distance tidak cocok. Sebagai penggantinya, digunakan fungsi sebagai berikut :

different (a_i,b_i) = { 0 jika a_i = b_i 1 }

Selainnya Dimana a_i dan b_i adalah nilai kategori. Jika nilai atribut antara dua record yang dibandingkan sama maka nilai jaraknya 0, artinya mirip, sebaliknya, jika berbeda maka nilai kedekatannya 1, artinya tidak mirip sama sekali. Misalkan atribut warna dengan nilai merah dan merah, maka nilai kedekatannya 0, jika merah dan biru maka nilai kedekatannya 1.

Untuk mengukur jarak dari atribut yang mempunyai nilai besar, seperti atribut pendapatan, maka dilakukan normalisasi. Normalisasi bisa dilakukan dengan min-max normalization atau Z-score standardization. Jika data training terdiri dari atribut campuran antara numerik dan kategori, lebih baik gunakan min-max normalization.

Untuk menghitung kemiripan kasus, digunakan rumus:

$$\text{Similarity (p, q)} = \frac{\sum_{i=1}^{n} f(pi, qi)\ X\ wi}{wi}$$

Keterangan :
P = Kasus baru
Q= Kasus yang ada dalam penyimpanan
n = Jumlah atribut dalam tiap kasus
i = Atribut individu antara 1 sampai dengan n
f = Fungsi similarity atribut i antara kasus p dan kasus q
w = Bobot yang diberikan pada atribut ke-i

ISSN Cetak 2745-925X
ISSN Online 2722-273X

JESSI Volume 01 Nomor 2 November 2020

Metode Evaluasi dan Validasi Algoritma Klasifikasi Data mining

Untuk mengukur akurasi algoritma klasifikasi, metode yang dapat digunakan antara lain cross validation, confusion matrix, dan kurva ROC (Receiver Operating Characteristic). Untuk mengembangkan aplikasi (development) berdasarkan model yang dibuat, digunakan Rapid Miner.

2. Metode

Metode yang digunakan dalam penelitian ini k-nearest neighbor (k-NN) ditambah feature selection (forward selection) untuk Mean Squared Error (MSE) dan Squared Error (SE)

2.1 k-Nearest Neighbor (k-NN)

Preprocessing yang dilakukan berupa penyeragaman pada data agar parameter-parameter data menjadi sama. Pada penelitian ini dilakukan proses normalisasi berupa normalisasi data kepemilikan rumah dapat dilakukan oleh persamaan (1).

$$S(i) = S(i) - \frac{1}{N}\sum_{i=1}^{N} S(i)$$

Sedangkan normalisasi Data dilakukan dengan persamaan berikut :

$$S(i) = \frac{S(i)}{S\ \max}$$

2.2. Forward Selection

Forward Selection merupakan salah satu metode pemodelan (pembangunan model linier) untuk menemukan kombinasi peubah yang “terbaik” dari suatu gugus peubah. Dalam Prosedur Forward selection, sekalinya variable masuk kedalam persamaan maka tidak bisa dihilangkan.

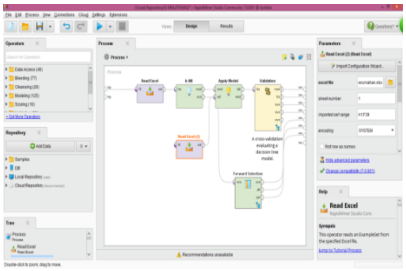
Selain itu, Forward selection dapat berarti memasukkan variabel bebas yang memiliki korelasi yang paling erat dengan variabel tak bebasnya (variabel yang paling potensial untuk memiliki hubungan linier dengan Y). kemudian secara bertahap memasukkan variabel bebas yang potensial berikutnya dan nanti akan terhenti sampai tidak ada lagi variabel bebas yang potensial.

3. Hasil dan Pembahasan

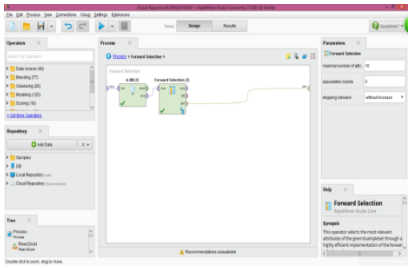
3.1 Rapidminer

RapidMiner merupakan perangakat 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 operator 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 [5].

Produknya sendiri. RapidMiner ditulis dengan munggunakan bahasa java sehingga dapat bekerja di semua sistem operasi. RapidMiner menyediakan GUI (Graphic User Interface) untuk merancang sebuah pipeline analitis. GUI ini akan menghasilkan file XML)Extensible Markup Language) yang mendefenisikan proses analitis keingginan pengguna untuk diterpkan ke data. File ini kemudian dibaca oleh RapidMiner untuk menjalankan analis secara otomatis.



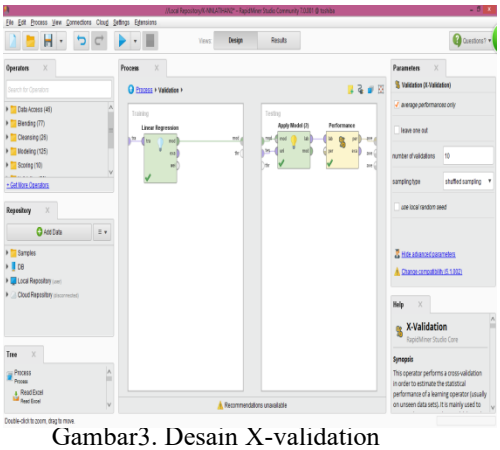
Gambar1. Desain Data RapidMiner



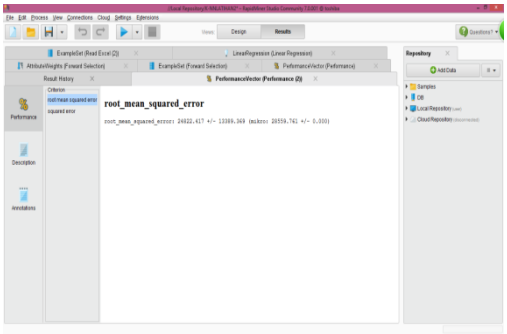
Gambar2. Desain Forward Selection

ISSN Cetak 2745-925X
ISSN Online 2722-273X

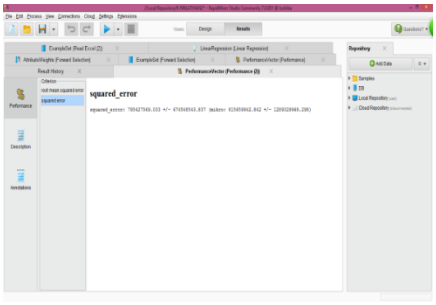
JESSI Volume 01 Nomor 2 November 2020



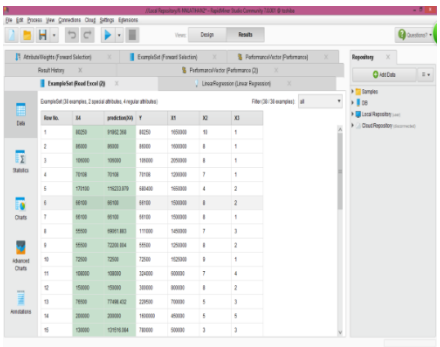
Gambar3. Desain X-validation



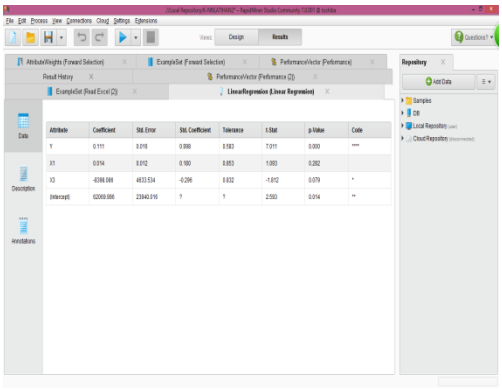
Gambar4. MSE



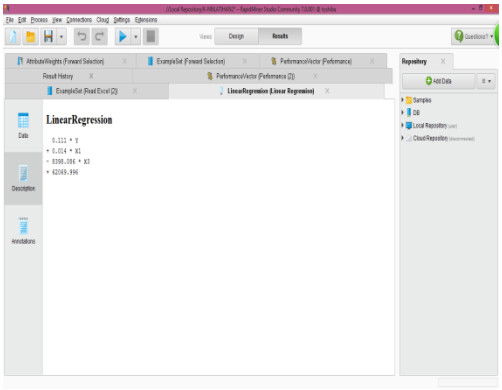
Gambar5. SE



Gambar6. Olah Data



Gambar7. Linear Regression



Gambar8. Hasil Liner Regression

4. Kesimpulan dan Saran

Dari hasil eksperiment yang dilakukan prediksi harga komoditi lada dengan menggunakan algoritma KNN berbasis forward selection berhasil dilakukan. Berdasarkan hasil eksperiment yang telah dilakukakan peneliti, menunjukan bahwa fitur seleksi yakni forward selection merupakan model yang lebih baik dalam melakukan seleksi variabel yang signifikan dibandingkan dengan bacward elimination.

Daftar Pustaka

[1] W. A. Sari, R. R. Hidayat, F. I. Administrasi, and U. Brawijaya, “ANALISIS SISTEM DAN PROSEDUR PEMBERIAN KREDIT KEPEMILIKAN RUMAH (KPR) BTN BERSUBSIDI DALAM USAHA MENDUKUNG PENGENDALIAN MANAJEMEN KREDIT (Studi Pada PT . Bank Tabungan Negara (Persero) Tbk . Kantor Cabang Kediri),” *J. Adm. Bisnis*, vol.

41, no. 1, p. 92, 2016.

[2] L. R. Angga Ginanjar Mabrur, "Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit," *J. Komput. dan Inform.*, vol. 1, no. 1, pp. 53–57, 2012.

[3] M. Nanja and P. Purwanto, "Metode K-Nearest Neighbor Berbasis Forward Selection Untuk Prediksi Harga Komoditi Lada," *Pseudocode*, vol. 2, no. 1, pp. 53–64, 2015.

[4] H. Leidiyana, "Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor," *J. Penelit. Ilmu Komputer, Syst. Embed. Log.*, vol. 1, no. 1, pp. 65–76, 2013.

[5] H. F. Takalamingan, D. P. E. Saerang, and M. Y. B. Kalalo, "Analisis Penerapan Sistem Dan Prosedur Pemberian Kredit Pemilikan Rumah (Kpr) Subsidi Pada Bank Tabungan Negara Cabang Manado," *Going Concern J. Ris. Akunt.*, vol. 14, no. 1, pp. 830–840, 2018.