

JURNAL ARMADA INFORMATIKA

STMIK Methodist Binjai

jurnal.stmikmethodistbinjai.ac.id/jai

Data Mining

Algoritma K-Nearest Neighbours untuk Pemberian Kredit KPR

Ahmad Yusri Azhari ¹, Raudah ², Riki Wanto ³^{1,2,3}Sistem Informasi, STMIK Logika, Medan

INFORMASI ARTIKEL

Diterima Redaksi: 05 Mei 2023

Revisi Akhir: 25 Mei 2023

Diterbitkan Online: 01 Juni 2023

KATA KUNCI

Pemberian; Algoritma K-Nearest Neighbours;
Kredit KPR;

KORESPONDENSI

Phone: +62 852-6253-6314

E-mail: ahmdayusriazharilogika@gmail.com

A B S T R A K

Secara garis besar, telah banyak nasabah yang sudah menjalin kerja sama dengan BRI. Sampai saat ini, belum ada sistem yang dibangun untuk membuat klasifikasi nasabah yang lulus untuk mendapatkan KPR. Secara garis besar, saat ini klasifikasi belum terpikirkan, padahal dengan adanya sistem klasifikasi ini akan memudahkan PT. Bank Rakyat Indonesia melakukan penilaian terhadap nasabah baru berdasarkan data yang sudah tersimpan sebelumnya (model *supervised learning*). Penelitian ini menyelesaikan masalah tersebut dengan kriteria yang digunakan sebagai dasar penilaian klasifikasi adalah kewarganegaraan, usia, pendapatan bulanan dan lama kerja menerapkan Metode yang digunakan dalam aplikasi ini adalah metode K-Nearest Neighbours.

PENDAHULUAN

Salah satu sumber utama pendapatan bagi sebuah bank adalah memberikan pinjaman/kredit kepada nasabah yang memenuhi ketentuan perbankan, hal ini karena sebagian besar biaya operasional bank yang ada dikelola dalam bentuk pinjaman. PT. Bank Rakyat Indonesia (Persero) atau yang dikenal dengan BRI merupakan sebuah perseroan terbatas yang bergerak dibidang penyedia jasa perbankan, pada perjalanan bisnisnya PT. Bank Rakyat Indonesia (Persero) lebih dikenal dengan Bank yang menyalurkan kredit perumahan (KPR) khususnya pada kredit perumahan bersubsidi. KPR BRI subsidi adalah kredit pemilikan rumah program kerja sama dengan Kementerian Pekerjaan Umum dan Perumahan Rakyat dengan suku bunga rendah dan cicilan ringan dan tetap sepanjang jangka waktu kredit, terdiri atas KPR untuk pembelian rumah tapak dan rumah susun.

Adapun manfaat dari penelitian ini adalah mengetahui bagaimana aplikasi yang berjalan di institusi khususnya dalam proses analisis nasabah dalam mendapatkan KRP di PT. Bank Rakyat Indonesia (Persero). membantu pihak PT. Bank Rakyat Indonesia (Persero) dalam memberikan solusi hasil analisis berupa klasifikasi nasabah yang mendapatkan KPR. Sehingga dari permasalahan yang terjadi pada sistem yang sedang berjalan yang telah diuraikan sebelumnya penulis kemudian berkeinginan untuk merancang sebuah sistem baru berbasis komputerisasi tentang pembelian dan order produk. Dalam pembuatan sistem yang baru ini penulis menggunakan bahasa pemrograman yaitu *pemrograman PHP* dengan menggunakan data testing berdasarkan data yang sudah ada yang akan digunakan dalam pembandingan dalam data pengujian. Dengan adanya sistem yang baru ini dapat membantu pekerjaan yang ada pada PT. Bank Rakyat Indonesia, Persero Cabang Putri Hijau.

TINJAUAN PUSTAKA

Algoritma *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. KNN memiliki beberapa kelebihan, yaitu ketangguhan terhadap *training data* yang memiliki banyak *noise* dan efektif apabila *training data*-nya besar. Atau dengan kata lain KNN merupakan algoritma yang menggunakan seluruh data latih untuk melakukan proses klasifikasi (*complete storage*)

Menurut Yunita, D., (2017), dikatakan bahwa *K-Nearest Neighbor* (KNN) merupakan kelompok *instance based learning*. Algoritma KNN juga termasuk salah satu teknik dalam *lazy learning*. KNN bekerja dengan cara mencari k-objek yang terdapat di dalam data training terdekat dengan objek yang ada pada data *testing* (data baru). Algoritma KNN adalah sebuah metode klasifikasi suatu objek berdasar data pembelajaran yang jaraknya terdekat dengan objek tersebut. KNN (*K-Nearest Neighbor*) merupakan teknik data mining yang paling sering digunakan pada proses klasifikasi yang diperkenalkan oleh Fix dan Hodges pada tahun 1951. KNN melakukan klasifikasi terhadap objek baru berdasarkan data latih yang jaraknya paling dekat dengan objek tersebut.

Prinsip kerja KNN (*K-Nearest Neighbor*) adalah mencari jarak terdekat antara data yang akan dievaluasi dengan k tetangga (*neighbor*) terdekatnya dalam pelatihan. Nilai *k* yang terbaik untuk algoritma ini tergantung pada data. Secara umum, nilai *k* yang tinggi akan mengurangi efek *noise* pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi semakin kabur. Teknik ini sangat sederhana dan mudah diimplementasikan. Mirip dengan teknik *clustering*, yakni dengan mengelompokkan suatu data baru berdasarkan jarak data baru tersebut ke beberapa data/tetangga (*neighbor*) terdekat. Misalkan ditentukan $k = 5$, maka setiap data *testing* dihitung jaraknya terhadap data *training* dan dipilih 5 data *training* yang jaraknya paling dekat ke data *testing*.

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

METODOLOGI

Agar penelitian ini dapat terlaksana dan berjalan sesuai dengan langkah-langkah dan prosedur penelitian, maka penulis merancang suatu proses tahapan penelitian yaitu:

- 1) Menentukan Objek Penelitian
- 2) Mengidentifikasi Masalah Yang Ada
- 3) Mengumpulkan data dan mengurutkannya
- 4) Menganalisis data
- 5) Merancang sistem
- 6) Melakukan pengujian, jika sesuai maka sistem dapat diimplementasikan.
- 7) Implementasi sistem
- 8) Penerapan terhadap sistem yang telah dibangun.

HASIL DAN PEMBAHASAN

Proses atau tahapan dari perhitungan algoritma *K-Nearest Neighbor* adalah melakukan proses normalisasi dan menentukan *Euclidean Distance* dan memasukkan nilai *k* guna memperoleh hasil terdekat dari nilai itu sendiri (umumnya 1 atau 2). Dimisalkan ada data pengujian (*testing*) sebagai berikut

Tabel 1. Data Transaksi

No.	Nama	Lama Kerja	Usia	Pendidikan	Pendapatan Bulanan	Kewarganegaraan
1	Alan ramadan	8	38	4	3500000	2
2	Bambang	9	40	5	3100000	2
3	Citra	10	35	5	3700000	2
4	Danang	12	45	5	4500000	2
5	Erling	10	40	4	5200000	1
6	Fadli	8	38	3	3000000	1
7	Galang	6	27	4	2500000	1

No.	Nama	Lama Kerja	Usia	Pendidikan	Pendapatan Bulanan	Kewarganegaraan
8	Hartati	15	60	4	2500000	1
9	armin	7	40	4	2500000	2
10	irfan	10	30	3	3000000	1
11	Susan	15	40	4	3500000	1
12	Daniel	12	38	4	2700000	1
13	Raisa	8	35	4	3200000	1
14	Sinta	5	25	4	3500000	1
15	Amanah	8	30	4	2700000	1

Langkah selanjutnya adalah mencari nilai normalisasi. Dengan normalisasinya adalah:

$$Normalisasi = \frac{data_x - data_{min}}{data_{max} - data_{min}}$$

Dimana $data_x$ adalah baris pertama dari satu atribut data latih, kemudian data min adalah nilai terkecil dari seluruh data pada atribut x berlaku juga untuk data max yang mencari nilai terbesar dari semua atribut data x.

Normalisasi sel total Lama Kerja (Alan Ramadhan) $(2,3) = (8 - 5)/(15 - 5) = 0.3$

Normalisasi sel total Lama Kerja (Bambang) $(3,3) = (9 - 5)/(15 - 5) = 0.4$

Normalisasi sel total Lama Kerja (Citra) $(3,3) = (10-5)/(15-5)=0.5$

Normalisasi sel total Lama Kerja (Danang) $(3,3)=(12-5)/(15-5)=0.7$

Dan seterusnya untuk kolom lama kerja. Sedangkan untuk kolom Usia, dapat dilakukan dengan bentuk:

Normalisasi sel total Usia (Alan Ramadan) $(2,4)=(38-25)/(60-25)=0.3714$

Normalisasi sel total Usia (Bambang) $(3,4)=(40-25)/(60-25)=0.4286$

Normalisasi sel total Usia (Citra) $(4,4)=(35-25)/(60-25)=0.2857$

Dan seterusnya hingga hasil yang didapat dapat dilihat pada tabel berikut ini:

Tabel 2. Hasil Normalisasi Data

No.	Nama	Lama Kerja	Usia	Pendidikan	Pendapatan Bulanan	Kewarganegaraan
1	Alan ramadan	0.3	0.3714	0.5	0.3704	
2	Bambang	0.4	0.4286	1	0.2222	
3	Citra	0.5	0.2857	1	0.4444	
4	Danang	0.7	0.5714	1	0.7407	
5	Erling	0.5	0.4286	0.5	1	
6	Fadli	0.3	0.3714	0	0.1852	
7	Galang	0.1	0.0571	0.5	0	
8	Hartati	1	1	0.5	0	
9	armin	0.2	0.4286	0.5	0	
10	irfan	0.5	0.1429	0	0.1852	
11	Susan	1	0.4286	0.5	0.3704	
12	Daniel	0.7	0.3714	0.5	0.0741	
13	Raisa	0.3	0.2857	0.5	0.2593	
14	Sinta	0	0	0.5	0.3704	
15	Amanah	0.3	0.1429	0.5	0.0741	

Langkah selanjutnya adalah pengolahan untuk data uji dengan data sebagai berikut:

Tabel 3. Data Uji

No	Nama	Lama Kerja	Usia	Pendidikan	Pendapatan Bulanan	Kewarganegaraan
----	------	------------	------	------------	--------------------	-----------------

1	admin	8	40	4	3500000	2
2	admin	10	35	4	2700000	2
3	admin	15	27	5	3700000	1
4	admin	13	37	4	2300000	2
5	admin	12	39	5	3300000	2

Langkah selanjutnya adalah normalisasi data uji dengan proses yang sama seperti data latih, sehingga didapat:

Tabel 4. Hasil Normalisasi Data Uji

No	Nama	Lama Kerja	Usia	Pendidikan	Pendapatan Bulanan	Kewarganegaraan
1	admin	0.2222	0.3939	0.5	0.3704	1
2	admin	0.4444	0.2424	0.5	0.0741	1
3	admin	1	0	1	0.4444	0
4	admin	0.7778	0.303	0.5	-0.0741	1
5	admin	0.6667	0.3636	1	0.2963	1

Langkah selanjutnya adalah menghitung jarak *euclidean*, dengan asumsi data yang diuji adalah sebagai berikut:

Tabel 5. Nilai Jarak Euclidean

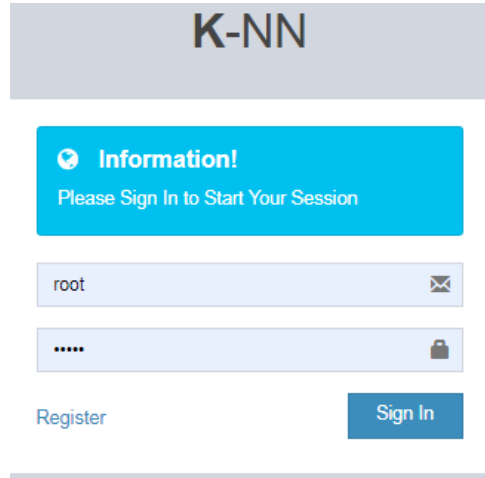
No.	Nama	Jarak Euclidean	Label
1	Alan ramadan	0.080988	LULUS
2	armin	0.372684	LULUS
3	Bambang	0.552069	LULUS
4	Citra	0.586819	LULUS
5	Danang	0.804314	LULUS
6	Raisa	1.01494	TIDAK LULUS
7	Amanah	1.075568	TIDAK LULUS
8	Sinta	1.097511	TIDAK LULUS
9	Galang	1.124973	TIDAK LULUS
10	Fadli	1.136159	TIDAK LULUS
11	Daniel	1.147429	LULUS
12	irfan	1.193513	LULUS
13	Erling	1.214402	LULUS
14	Susan	1.26735	LULUS
15	Hartati	1.452421	LULUS

Maka langkah terakhir adalah menentukan lulus atau tidak berdasarkan nilai rata-rata *euclidean* berdasarkan label didapat

Tabel 6. Hasil Pengujian Di Dapat Nilai X (K=2)

No	Nama	Lama Kerja	Usia	Pendidikan	Pendapat an Bulanan	Kewargane garaan	Label	X
1	admin	0.2222	0.3939	0.5	0.3704	1	LULUS	0.08099
2	admin	0.6667	0.3636	1	0.2963	1	LULUS	0.26027
3	admin	1	0	1	0.4444	0	LULUS	0.72348
4	admin	0.4444	0.2424	0.5	0.0741	1	TIDAK LULUS	0.35396
5	admin	0.7778	0.303	0.5	-0.0741	1	TIDAK LULUS	0.65616

Demonstrasi program merupakan suatu prosedur yang dilaksanakan untuk menampilkan hasil dari sistem yang diusulkan, dimana hasilnya telah dijalankan dengan judul Analisis Algoritma K-Nearest Neighbour Untuk Pemberian Kredit KPR Pada PT. Bank Rakyat Indonesia, Persero Cabang Putri Hijau Medan. Berikut ini tampilan demonstrasi program sistem yang diusulkan. Sebelum menjalankan program, diwajibkan kepada *user* untuk mengisi *form login* terlebih dahulu.



Gambar 1. Login

Setelah *login* berhasil dijalankan maka tampil menu utama sistem yang diusulkan.



Gambar 2. Utama

Berikut ini adalah tampilan halaman data testing sistem yang diusulkan pada tampilan gambar dibawah ini.

No	Nama	Lama Kerja	Ukua	Pendidikan	Pendapatan Bulanan	Kewarganegaraan	Label	Aksi
1	Alan ramadhan	8	38	4	3000000	2	LULUS	Edit Hapus
2	Bambang	9	40	5	3100000	2	LULUS	Edit Hapus
3	Cika	10	35	3	3700000	2	LULUS	Edit Hapus
4	Darung	12	45	3	4500000	2	LULUS	Edit Hapus
5	Ering	10	40	4	5200000	1	LULUS	Edit Hapus
6	Fadi	8	36	2	3000000	1	TIDAK LULUS	Edit Hapus
7	Galing	6	27	4	2000000	1	TIDAK LULUS	Edit Hapus
8	Hartal	15	60	4	2000000	1	TIDAK LULUS	Edit Hapus
9	amin	7	40	4	2000000	2	TIDAK LULUS	Edit Hapus
10	ifen	10	30	3	3000000	1	TIDAK LULUS	Edit Hapus
11	Suan	15	40	4	3500000	1	LULUS	Edit Hapus
12	Daral	12	38	4	2750000	1	LULUS	Edit Hapus

Gambar 3. Data Test

Berikut ini adalah tampilan hasil pengujian jarak dapat dilihat pada gambar berikut ini:

No.	Nama	Jarak	Label
1	Alian	0,00000	LULUS
2	Amir	0,27204	TIDAK LULUS
3	Bambang	0,45200	LULUS
4	Cika	0,28819	LULUS
5	Denang	0,80434	LULUS
6	Rasa	1,01494	LULUS
7	Anasari	1,07088	LULUS
8	Sita	1,09711	LULUS
9	Galang	1,12497	TIDAK LULUS
10	Fatih	1,13618	TIDAK LULUS
11	Dewi	1,16726	LULUS
12	Ida	1,18812	TIDAK LULUS
13	Elang	1,21402	LULUS
14	Sulan	1,25735	LULUS
15	Natali	1,40201	TIDAK LULUS

Gambar 4. Tampilan Hasil Pengujian Jarak Euclidean

KESIMPULAN

1. Sistem mampu melakukan pengelompokan atau klasifikasi terhadap pemberian kredit kepada nasabah dengan melabialisasi dengan Lulus atau Tidak Lulus berdasarkan data latih ataupun testing dalam pengujiannya
2. Sistem mampu melakukan proses normalisasi yang terbentuk merupakan nilai bentuk atau gap antara nilai terendah dengan nilai tertinggi yang dibandingkan.
3. Sistem mampu mengolah data uji berdasarkan data testing yang telah dikumpulkan.

DAFTAR PUSTAKA

- Anhar. 2016. Kumpulan Source Code Visual Basic 6.0 untuk Skripsi. Jakarta: PT. Elex Media Komputindo
- Anisya. 2013. Aplikasi Sistem Database Rumah Sakit Terpusat Pada Rumah Sakit Umum (RSU) 'Aisyiyah Padang Dengan Menerapkan Open Source (PHP-MYSQL). Padang.: Jurnal Momentum .Vol.15 No.2.
- Ardiana, P.Y., et al, 2021. Sistem Basis Data Lanjutan. Penerbit Yayasan Kita Menulis, Medan
- Eriyanto, 2013. Analisis Isi: Pengantar Metodologi Untuk Penelitian Ilmu Komunikasi dan Ilmu-Ilmu Sosial Lainnya. Edisi Ketiga. Penerbit PT. Fajar Interpatama Mandiri, Jakarta.
- Hadi, et al. 2017. Sistem Informasi Konsultasi Kesehatan Online Dengan MySQL Di Balai Pengobatan Azzainiyah Nurul Jadid Paiton Probolinggo. Jurnal Teknik Informatika, 5(02).
- Harun, M., 2015, Rekayasa Perangkat Lunak Untuk Sistem Informasi, Graha Ilmu, Yogyakarta
- Hidayat, R. 2014. Sistem Informasi Ekspedisi Barang Dengan Metode E-CRM Untuk Meningkatkan Pelayanan Pelanggan. Jurnal Sisfotek Global, 4(2).
- Indriyanto, J., 2021. Algoritma K-Nearest Neighbor Untuk Prediksi Nasabah Asuransi. Penerbit NEM. Jakarta
- Juansyah. 2015. Pembangunan Aplikasi Child Tracker Berbasis Assisted - Global Positioning System (A-Gps) Dengan Platform Android. Jurnal KOMPUTA. Vol 1 No. 1, Agustus 2015. ISSN : 2089-9033.
- Kadir, A., 2014, Pengenalan Sistem Informasi Edisi Revisi, PT. Andi Offset, Yogyakarta
- Kurniawan, R., et al, 2019. Cara Mudah Belajar Statistik: Analisis Data & Eksplorasi. Edisi Pertama. Penerbit Kencana. Jakarta
- Maimunah et al. 2017. Aplikasi Sistem Order Online Berbasis Mobile Android Pada Outlet Pizza Hut Delivery. Jurnal SEMNASTEKNOMEDIA ONLINE Vol 5 No 1.
- Maudi, et al. 2014. Desain Aplikasi Sistem Informasi Pelanggan PDAM Berbasis Web GIS (Studi Kasus: Kota Demak). Jurnal Geodesi Undip Vol.3 No.3.
- Mulyani, S., 2016, Metode Analisis dan Perancangan Sistem, Bandung: Abdi Sistematika
- Musrifah, Ai, dan Ega Mutia. 2017. Pembuatan Aplikasi Pengelolaan Proposal Di Unit Kegiatan Pengelola (UPK) Kecamatan Mande Berbasis Desktop. Jurnal MJI. Vol. 9 No. 1, Juni 2017.
- Pahlevi, S.M., 2013. Tujuh Langkah Praktis Pembangunan Basis Data: Pendekatan Praktis Analisis, Perancangan dan Implementasi Basis Data Dengan Banyak Contoh, Gambar dan Mindmap. PT. Elex Media Komputindo, Jakarta.
- Pratama, B.I., 2021. Metode Analisis Isi (Metode Penelitian Populer Ilmu-Ilmu Sosial). Penerbit Unisma Press. Malang.
- Rahardi, M., et al. 2016. Perancangan Sistem Group Tracking Pada Aktivitas Touring Berbasis Mobile. Universitas Gadjah Mada. Seminar Nasional Teknologi Informasi dan Multimedia 2016. ISSN: 2302-3805
- Rosa, A.S., dan Shalahuddin, M., 2013., Rekayasa Perangkat Lunak: Terstruktur dan Berorientasi Objek, Graha Ilmu, Yogyakarta
- Satia, O., & Sujadi, H. 2016. Sistem Informasi Geografis Budaya dan Ekowisata Menggunakan Algoritma Dijkstra Dengan Google Maps API Dan Java Server Page (DISPORABUDPAR Majalengka). STIMA
- Supardi, Y., 2015, Semua Bisa Menjadi Programmer Visual Basic 2012 – Basic, PT. Elex Media Komputindo, Jakarta.