

Model Penentuan Jenis Beasiswa Menggunakan Algoritma K-NN Kombinasi Basis Aturan Dan Basis Pengetahuan

Bahar, Nidia Rosmawanti

STMIK Banjarbaru

Jl. A. Yani Km. 33,3 Banjarbaru

baharahman@gmail.com, nidiabjb@yahoo.com

Abstrak

Pengambil keputusan sering menggunakan intuisi dalam proses pengambilan keputusan, padahal dengan intuisi banyak memiliki kekurangan sehingga perlu dikembangkan suatu model pengambilan keputusan yang diambil berdasarkan informasi yang telah diolah dan disajikan dengan dukungan model penunjang keputusan. Model k-Nearest Neighbor (k-NN) adalah suatu model penunjang keputusan yang tidak berbasis aturan, melainkan menggunakan algoritma pembelajaran terawasi, hasil dari data masukan yang baru diklasifikasi berdasarkan kedekatan dengan data pada kasus lama (fakta) sebagai data pelatihan. Model k-NN cukup sederhana dan mudah diaplikasikan dalam kasus nyata. Namun demikian model k-NN memiliki kelemahan, yaitu membutuhkan data kasus yang telah pernah terjadi (fakta) sebagai basis pengetahuan sistem. Dengan demikian, model k-NN tidak dapat memutuskan permasalahan baru tanpa belajar dari kasus yang telah pernah terjadi. Pada penelitian ini dirumuskan sebuah model k-NN yang menggunakan basis aturan untuk menghasilkan data pelatihan sebagai basis pengetahuan sistem.

Kata Kunci: Sistem Penunjang Keputusan, k-NN (k-Nearest Neighbor), Basis Aturan, Basis Pengetahuan

Abstract

Decision makers often use intuition in decision making processes, but with intuition has many shortcomings that need to develop a model of making decisions based on the information that has been processed and presented with the support of decision support models. Model k-Nearest Neighbor (k-NN) is a model-based decision support is not a rule, but uses a supervised learning algorithm, the results of the new input data are classified based on proximity to the data on the old case (facts) as training data. K-NN model is quite simple and easy to apply in real cases. However, k-NN models have a weakness, which requires data on cases that have never happened (facts) as a knowledge base system. Thus, k-NN models can not be decided without learning a new problem from a case that has ever happened. In this study formulated a model of k-NN that uses the rule base to generate the training data as a knowledge base system.

Keywords: Decision Support Systems, k-NN (k-Nearest Neighbor), Rule Base, Knowledge Base

1. Pendahuluan

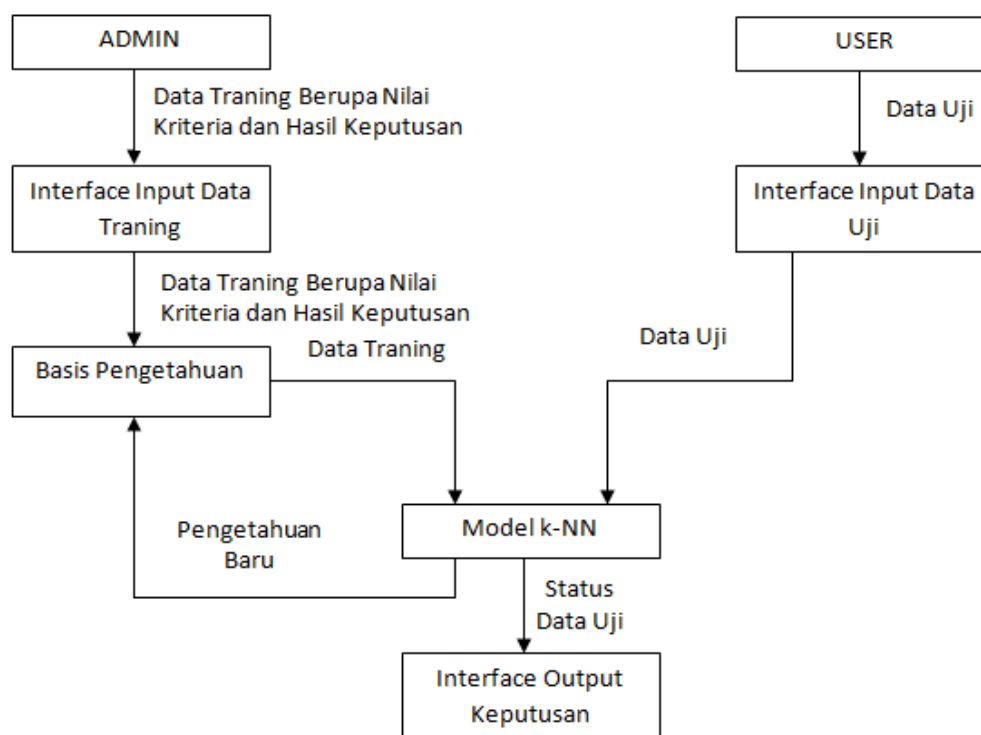
Ada banyak model penunjang keputusan dalam dunia komputasi yang disesuaikan dengan bidang atau masalah yang akan diselesaikan, misalnya: masalah penetapan skala prioritas, masalah prediksi, masalah klasifikasi dan pengelompokan, masalah asosiasi, dan masalah-masalah lainnya. Umumnya masalah-masalah tersebut berkaitan dengan permasalahan bisnis, misalnya segmentasi pelanggan/produk, asosiasi produk, klasifikasi transaksi bisnis, pengelompokan pelanggan berdasarkan kesamaan karakteristik, prediksi kelakuan bisnis di masa mendatang, klasifikasi penerima beasiswa, penetapan skala prioritas pembiayaan, dan lain-lain [1][2].

Salah satu model penunjang keputusan dalam bidang klasifikasi yang cukup sederhana dan mudah diaplikasikan dalam kasus nyata adalah model *k-Nearest Neighbor* (k-NN). Model k-NN adalah suatu model penunjang keputusan yang tidak berbasis aturan, melainkan menggunakan algoritma pembelajaran terawasi, dimana hasil dari data masukan yang baru diklasifikasi berdasarkan kedekatan dengan data pada kasus lama (fakta) sebagai data pelatihan [3]. Namun demikian model k-NN memiliki kelemahan, yaitu membutuhkan data kasus yang telah pernah terjadi (fakta) sebagai basis pengetahuan sistem. Dengan demikian, model k-NN tidak dapat memutuskan permasalahan baru tanpa belajar dari kasus yang telah pernah terjadi [4][5][6].

Paper ini merumuskan sebuah model k-NN yang menggunakan basis aturan untuk menghasilkan data pelatihan sebagai basis pengetahuan sistem, sehingga lebih fleksibel dalam menyelesaikan masalah-masalah klasifikasi.

2. Metode Penelitian

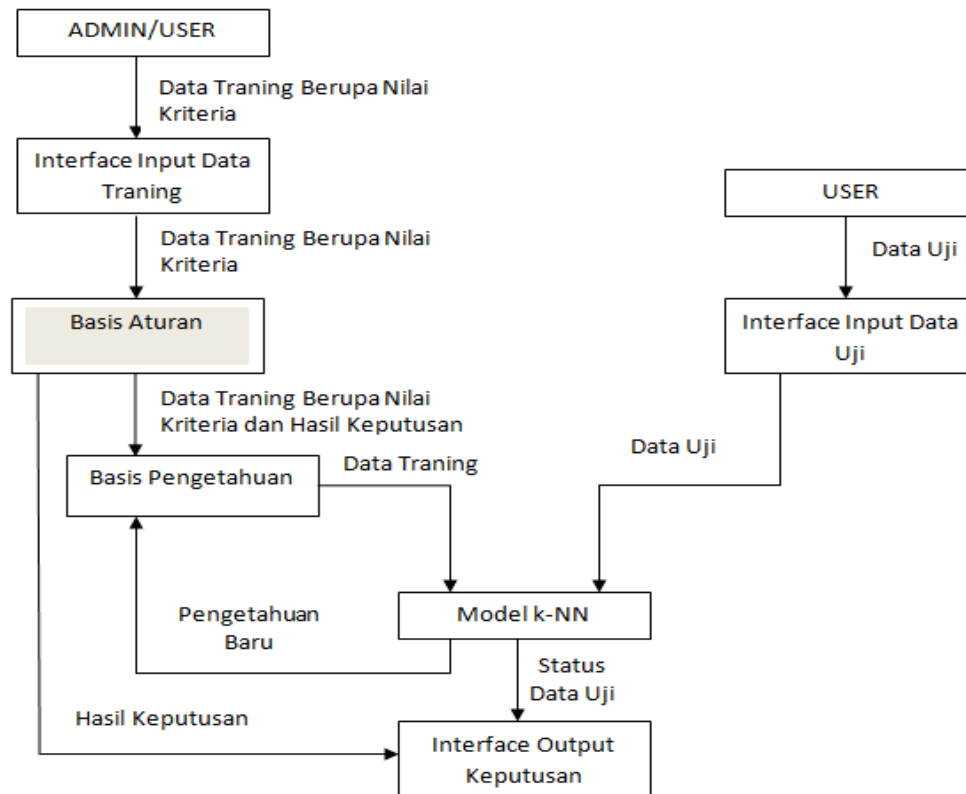
Secara umum, struktur Program Aplikasi Sistem yang menggunakan model *k-Nearest Neighbor* (k-NN) seperti pada gambar [1] [2]:



Gambar 1. Struktur Program Aplikasi Berbasis k-NN Secara Umum

Admin menginputkan data kejadian masa lalu sebagai data pembelajaran/*traning*, data ini tersimpan di dalam Basis Pengetahuan Sistem melalui interface data traning. Pengguna menginputkan data baru yang akan diketahui statusnya melalui interface data uji. Model k-NN memproses data dan menghasilkan output keputusan berupa mengenai status data baru yang diuji. Status data baru tersebut juga akan menambah perbendaharaan pengetahuan sistem sebagai pengetahuan baru.

Struktur model *k-Nearest Neighbor* (k-NN) yang mengkombinasikan Basis Aturan dan Basis Pengetahuan seperti pada gambar 2:



Gambar 2. Struktur Program Aplikasi Berbasis k-NN yang Menggunakan Kombinasi Basis Aturan dan Basis Pengetahuan

Pada awal proses, Admin atau Pengguna memasukkan Nilai Kriteria suatu kejadian baru melalui interface input data tranning. Data ini diproses pada Basis Aturan yang bertindak sebagai roole untuk menghasilkan suatu keputusan yang dapat digunakan sebagai keputusan akhir dan keputusan yang tersimpan sebagai pengetahuan pada Basis Pengetahuan. Setelah basis pengetahuan terisi sejumlah n data, Basis Aturan dapat di nonaktifkan, selanjutnya Pengguna dapat memasukkan data baru (data uji) yang akan diketahui statusnya melalui interface data uji. Model k-NN memproses data dan menghasilkan output keputusan berupa status data baru yang diuji. Status data baru tersebut juga akan menambah perbendaharaan pengetahuan sistem sebagai pengetahuan baru.

Berikut adalah langkah-langkah dalam algoritma *k-Nearest Neighbor* yang mengkombinasikan Basis Aturan dan Basis Pengetahuan, dengan pencarian jarak menggunakan rumus *Euclidian*:

1. Menentukan Nilai n buah data awal sebagai basis pengetahuan awal sistem
2. Menentukan Nilai K (tetangga) terdekat
3. Mempersiapkan data tranning berupa nilai kriteria suatu data baru yang belum diketahui statusnya.
4. Menentukan Status setiap data tranning berdasarkan suatu aturan (*roole*) tertentu untuk menghasilkan basis pengetahuan sistem
5. Menghitung jarak setiap sampel data tranning terhadap data yang akan diuji (data uji) berdasarkan persamaan:

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots (1)$$

Dengan:

- d = jarak
- x_i = sampel data
- y_i = data uji
- i = variabel data
- n = dimensi data

6. Menetapkan status data uji berdasarkan nilai rata-rata K buah sampel data training terdekat.

3. Hasil Dan Pembahasan

Misalkan di sebuah Perguruan Tinggi XX terdapat 11 mahasiswa calon penerima Beasiswa yang belum diketahui status jenis Beasiswa yang akan diperolehnya (seperti pada tabel 1):

Tabel 1 Data Mahasiswa Calon Penerima Beasiswa P.T. XX

Maha-siswa	IPK	Penghasilan Keluarga Rata2 per kepala (Rp/Bln)	Nilai Keaktifan Berorganisasi	Nilai Kualitas Proposal Riset	Jenis Beasiswa
1	3,55	2.400.000	80	75	?
2	2,75	900.000	75	70	?
3	3,15	950.000	80	80	?
4	2,75	1.150.000	70	85	?
5	3,45	750.000	85	80	?
6	3,24	500.000	70	70	?
7	3,02	500.000	80	75	?
8	2,38	1.100.000	75	85	?
9	3,18	1.250.000	75	75	?
10	3,25	900.000	70	75	?
11	3,27	1.800.000	75	70	?

Sebelumnya pada Perguruan Tinggi XX tersebut belum pernah menyelenggarakan pemberian beasiswa kepada mahasiswanya. Pihak Perguruan Tinggi menetapkan sebuah aturan secara umum untuk menetapkan jenis beasiswa berdasarkan nilai kriteria tertentu, yaitu:

- Jika $IPK \geq 3$, Rata-rata Penghasilan Keluarga ≥ 1 Juta, Nilai Keaktifan Berorganisasi ≥ 70 , Nilai Kualitas Proposal Riset ≥ 70 , maka Jenis Beasiswa yang diperoleh adalah Beasiswa BERPRESTASI
- Jika $IPK \geq 3$, Rata-rata Penghasilan Keluarga > 1 Juta, Nilai Keaktifan Berorganisasi ≥ 70 , Nilai Kualitas Proposal Riset ≥ 70 , maka Jenis Beasiswa yang diperoleh adalah beasiswa KURANG MAMPU
- Jika $IPK < 3$, Rata-rata Penghasilan Keluarga ≥ 1 Juta, Nilai Keaktifan Berorganisasi ≥ 70 , Nilai Kualitas Proposal Riset ≥ 70 , maka TIDAK DAPAT
- Jika $IPK < 3$, Rata-rata Penghasilan Keluarga < 1 Juta, Nilai Keaktifan Berorganisasi ≥ 70 , Nilai Kualitas Proposal Riset ≥ 70 , maka Jenis Beasiswa yang diperoleh adalah Beasiswa KURANG MAMPU

Keempat aturan tersebut di atas akan menjadi Basis Aturan bagi sistem. Misalkan ditentukan nilai $n = 10$, $k = 3$, maka:

1. Sistem (basis aturan) akan membaca dan memproses data mahasiswa ke 1 hingga mahasiswa ke 10 untuk menghasilkan status Jenis Beasiswa, hasilnya pada tabel 2:

Tabel 2. Status Jenis Beasiswa Berdasarkan Penalaran Basis Aturan

Maha-siswa	IPK	Penghasilan Keluarga Rata2 per kepala (Rp/Bln)	Nilai Keaktifan Berorganisasi	Nilai Kualitas Proposal Riset	Jenis Beasiswa Yang Diperoleh
1	3,55	1.400.000	80	75	Berprestasi
2	2,75	900.000	75	70	Kurang Mampu
3	3,15	950.000	80	80	Kurang Mampu
4	2,75	1.150.000	70	85	Tidak Dapat
5	3,45	750.000	85	80	Kurang Mampu
6	3,24	500.000	70	70	Kurang Mampu
7	3,02	500.000	80	75	Kurang Mampu
8	2,38	1.100.000	75	85	Tidak Dapat
9	3,18	1.250.000	75	75	Berprestasi
10	3,25	900.000	70	75	Kurang Mampu
11	3,27	1.800.000	75	70	

- Selanjutnya Sistem menonaktifkan basis aturan dan menampilkan status tersebut sebagai suatu hasil keputusan. Pada saat yang bersamaan, sistem juga akan menyimpan data status ke 10 mahasiswa tersebut sebagai Basis Pengetahuan Sistem.
- Untuk menentukan status data ke 11 dan seterusnya, user menggunakan *interface input* data uji untuk menginput nilai kriteria, selanjutnya sistem memproses data tersebut menggunakan formula standar model k-NN (persamaan 1) dengan merujuk pada data yang terdapat dalam basis pengetahuan sistem.

Jarak Data Uji terhadap data tranning 1 adalah:

$$\begin{aligned}
 d1 &= \sqrt{(3,55 - 3,27)^2 + (1.400.000 - 1.800.000)^2 + (80 - 75)^2 + (75 - 70)^2} \\
 &= 400.000
 \end{aligned}$$

Jarak Data Uji terhadap data tranning 2 adalah:

$$\begin{aligned}
 d2 &= \sqrt{(2,75 - 3,27)^2 + (900.000 - 1.800.000)^2 + (75 - 75)^2 + (70 - 70)^2} \\
 &= 900.000
 \end{aligned}$$

Demikian seterusnya untuk menentukan jarak data uji terhadap data tranning 3 hingga data tranning 10. Hasil perhitungan selengkapnya disajikan pada tabel 3:

Tabel 3. Hasil Perhitungan Jarak Data Uji terhadap Data Training

Maha-siswa	IPK	Penghasilan Keluarga Rata2 per kepala (Rp/Bln)	Nilai Keaktifan Berorgani-sasi	Nilai Kualitas Proposal Riset	Jenis Beasiswa Yang Diperoleh	Jarak
1	3,55	1.400.000	80	75	Berprestasi	400.000
2	2,75	900.000	75	70	Kurang Mampu	900.000
3	3,15	950.000	80	80	Kurang Mampu	850.000
4	2,75	1.150.000	70	85	Tidak Dapat	650.000
5	3,45	750.000	85	80	Kurang Mampu	1.050.000
6	3,24	500.000	70	70	Kurang Mampu	1.300.000
7	3,02	500.000	80	75	Kurang Mampu	1.300.000
8	2,38	1.100.000	75	85	Tidak Dapat	700.000
9	3,18	1.250.000	75	75	Berprestasi	550.000
10	3,25	900.000	70	75	Kurang Mampu	900.000

Berdasarkan 3 sampel data training terdekat, ada 2 data berlabel jenis Beasiswa BERPRESTASI dan 1 data berlabel TIDAK DAPAT Beasiswa, maka status data mahasiswa calon penerima Beasiswa yang diuji (data ke 11) adalah termasuk dalam kelompok penerima beasiswa BERPRESTASI.

4. Kesimpulan

Berdasarkan hasil analisis model yang dikembangkan pada penelitian ini, dapat disimpulkan bahwa model k-NN yang mengkombinasikan Basis Aturan dan Basis Pengetahuan lebih fleksibel dalam menetapkan status data uji, karena sistem tidak harus memerlukan data kejadian masa lalu (data histori) untuk difungsikan sebagai basis pengetahuan. Basis Pengetahuan tercipta berdasarkan hasil proses Basis Aturan. Hal ini berbeda dengan model k-NN standar yang memerlukan data kejadian masa lalu (data histori) yang difungsikan sebagai Basis Pengetahuan, karena tidak memiliki Basis Aturan untuk menciptakan Basis Pengetahuan.

Daftar Pustaka

- [1] Nidia, Bahar, 2013, *Penentuan Skala Prioritas Berbasis Algoritma AHP Termodifikasi*, Jurnal JUTISI, Vol. 2 No. 3.
- [2] Kusri, Amha T. L., 2008, *Algoritma Datamining*, PENERBIT ANDI, Yogyakarta
- [3] Tedy R., Ardhitya W.I., Wahyu P., Sri Kusumadewi, 2008, *Sistem Pendukung Keputusan Berbasis Pocket PC Sebagai Penentu Status Gizi Menggunakan Metode k-Nearest Neighbor*, Jurnal Teknoin, Vol. 13 No. 2, 1-6.
- [4] Chusnul Imama, 2013, *Penerapan Case Based Reasoning Dengan Algoritma Nearest Neighbor Untuk Analisis Pemberian Kredit di Lembaga Pembiayaan*, Jurnal Manajemen Informatika, Vol. 2 No. 1, 11-21
- [5] Ninki H., Sri Kusumadewi, 2008, *Sistem Pendukung Keputusan Berbasis SMS Untuk Menentukan Status Gizi Dengan Metode K-Nearest Neighbor*, Prosiding pada Seminar Nasional Aplikasi Teknologi Informasi, Yogyakarta, 21 Juni 2008.
- [6] Nobertus K., Helmi, Bayu P., 2013, *Algoritma k-Nearest Neighbor Dalam Klasifikasi Data Produksi Kelapa Sawit Pada PT. Minamas Kecamatan Parindu*, Jurnal Mimaster, Vol.2, No.1, 33-38.