

# Model Aplikasi Penentuan Jenis Beasiswa Berbasis Algoritma *K-NN* Termodifikasi

Soegiarto<sup>1</sup>, Bahar<sup>2</sup>

Program Studi Teknik Informatika, STMIK Banjarbaru

Jl. A. Yani Km. 33,3 Banjarbaru

<sup>1</sup>ttsoegiarto.com, <sup>2</sup>baharahman@gmail.com

## Abstrak

Salah satu model penunjang keputusan yang biasa digunakan oleh organisasi adalah prediksi dalam rentang waktu diskrit (klasifikasi). Dalam klasifikasi ada dua pekerjaan utama yang dilakukan, yaitu (1) pembangunan model basis pengetahuan sistem, dan (2) penggunaan model untuk melakukan pengenalan/prediksi kelas suatu objek data baru. Pada algoritma klasifikasi yang standar, umumnya model basis pengetahuan dihasilkan dari algoritma model (algoritma pelatihan), yang memiliki kelemahan yaitu data yang dimodelkan berdasarkan suatu teknik pemodelan tertentu cenderung dipaksa untuk masuk pada kelas tertentu yang sudah diset sebelumnya, sehingga sering terjadi sebuah objek kelas tidak mencerminkan sepenuhnya sifat dari kelas yang ada (tidak alami), berdampak pada proses klasifikasi menjadi kurang akurat.

Pada penelitian ini diujicoba model Klasifikasi yang menggunakan Algoritma Klustering untuk pembentukan Kelas Model secara alami sebagai Basis Pengetahuan Sistem, sehingga dapat menyelesaikan permasalahan pada teknik pemodelan yang menghasilkan objek klas yang cenderung dipaksakan, sehingga menghasilkan proses klasifikasi yang lebih akurat.

Kata Kunci: *Klasifikasi, Klustering, Algoritma k-Nearest Neighbor Termodifikasi, Basis Pengetahuan*

## Abstract

*One decision support models used by the organization is in the range of discrete time prediction (classification). In the classification there are two main work is done, namely (1) the development model of the knowledge base system, and (2) the use of the model to perform recognition / prediction of the class a new data object. In the classification algorithms are standard, general model of a knowledge base generated from the algorithm models (algorithm training), which has the disadvantage of data modeled by a modeling technique particular tend to be forced to enter a particular class set up already, so often an object class is not fully reflects the nature of the existing classes (unnatural), have an impact on the classification process will be less accurate.*

*In this study tested the model classification using clustering algorithm for the formation of classes naturally as a Model System Knowledge Base, so that it can resolve the problems in the modeling technique produces an object class that tends imposed, resulting in more accurate classification process.*

*Keywords: classification, clustering, k-Nearest Neighbor algorithm Modified, 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 perilaku bisnis di masa mendatang, klasifikasi penerima beasiswa, penetapan skala prioritas pembiayaan, dan lain-lain [1][2][3][4][5]. Model-model penunjang keputusan tersebut

memanfaatkan konsep data mining untuk menggali informasi yang berguna, yang tersembunyi dalam tumpukan data suatu organisasi dari tahun ke tahun[6].

Salah satu model penunjang keputusan yang biasa digunakan oleh organisasi adalah prediksi, baik prediksi dalam rentang waktu diskrit (klasifikasi) maupun yang bersifat kontinu (regresi). Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas target yang tersedia. Dalam klasifikasi ada dua pekerjaan utama yang dilakukan, yaitu (1) pembangunan model berdasarkan data latih sebagai prototip untuk disimpan sebagai memori atau basis pengetahuan sistem, dan (2) penggunaan model atau pengetahuan tersebut untuk melakukan pengenalan/klasifikasi/prediksi suatu objek data lain agar diketahui di kelas mana objek data tersebut dalam model yang sudah disimpannya [7]. Ada dua cara yang umumnya digunakan pada tahap pembangunan model berdasarkan data latih, yaitu: (1) belajar dari kasus atau kejadian ril yang pernah terjadi di alam nyata berdasarkan hubungan sebab akibat / fakta, dan (2) model dihasilkan dari algoritma model (algoritma pelatihan)[8]. Kedua cara tersebut masing-masing memiliki kelebihan dan kekurangan. Kelebihan dari cara pertama adalah model dihasilkan dengan cara yang cepat karena tidak diperlukan proses pembelajaran (*lazy learner*) [7], namun kelemahannya adalah tidak semua pengetahuan dapat diperoleh berdasarkan kejadian dari alam nyata (fakta), sehingga proses klasifikasi tidak dapat dilakukan pada kasus seperti ini. Kelebihan cara kedua adalah model secara pasti dapat dihasilkan dari suatu teknik pemodelan tertentu (misal: model logika standar If... Then ...), namun kelemahannya adalah data yang dimodelkan berdasarkan model ini cenderung dipaksa untuk masuk pada klas tertentu yang sudah diset sebelumnya, sehingga sering terjadi sebuah objek klas tidak mencerminkan sepenuhnya sifat dari kelas yang ada [2].

Artikel ini menyajikan sebuah model aplikasi Klasifikasi berbasis algoritma *K-NN Classification* menggunakan Algoritma Klastering (*K-Means Clustering*) untuk pembentukan kelas data pembelajaran secara alami sebagai Basis Pengetahuan sistem, sehingga disamping dapat menyelesaikan masalah tidak tersedianya fakta sebagai Basis Pengetahuan sistem, juga dapat menyelesaikan permasalahan pada teknik pemodelan yang menghasilkan objek kelas yang cenderung dipaksakan. Ujicoba dilakukan dalam kasus penentuan jenis beasiswa yang dapat diterima oleh mahasiswa.

## 2. METODOLOGI

### 2.1 Model Algoritma

Algoritma *K-NN classification* yang menggunakan algoritma Klastering *K-Means* untuk pembentukan kelas data pelatihan dengan pencarian jarak menggunakan rumus *Euclidian*, sebagai berikut:

1. Menentukan Nilai n buah data awal sebagai basis pengetahuan awal sistem
2. Menentukan Nilai K (tetangga) terdekat
3. Mempersiapkan data training berupa nilai kriteria suatu data baru yang belum diketahui statusnya.
4. Menentukan status (kelas data) setiap data training menggunakan formula *K-Mean Clustering* [2] sebagai basis pengetahuan sistem
5. Menghitung jarak setiap sampel data training 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<sub>i</sub> = sampel data
- y<sub>i</sub> = data uji
- i = variabel data
- n = dimensi data

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

**2.2 Model Interface**

*Interface* sistem aplikasi penentuan jenis beasiswa menggunakan algoritma K-NN termodifikasi disajikan pada gambar 1 dan 2.

STMIK BANJARBARU  
APLIKASI PENENTUAN JENIS BEASISWA

---

Data Training

No Data	IPK	Penghasilan	Keaktifan
1			
2			
3			
4			

---

Posisi Klaster

Data	Berprestasi	Kurang Mampu
1	*	
2		*
3		*
4		*

Gambar 1. Model Antarmuka Data Traning Aplikasi Penentuan Jenis Beasiswa Berbasis K-NN Termodifikasi

Pada gambar 1, user mengundu data melalui tombol Brouse dan menampilkan data pada tabel data. User menekan tombol proses untuk memberikan label (beasiswa berprestasi atau beasiswa kurang mampu) sesuai prosedur kerja algoritma klastering. Data yang telah diberi status/label ini akan menjadi basis pengetahuan bagi algoritma K-NN. Setelah sistem memiliki data traning, user selanjutnya dapat menggunakan aplikasi untuk menentukan status data baru yang belum diketahui statusnya (penerima beasiswa berprestasi atau beasiswa kurang mampu) melalui *interface* gambar 2.

STMIK BANJARBARU  
APLIKASI PENENTUAN JENIS BEASISWA

---

Nomor Induk

Nama Mahasiswa

Indek Prestasi Kumulatif

Penghasilan Kepala Keluarga Per Bulan

Nilai Keaktifan Berorganisasi

---

Sarn Jenis Beasiswa Untuk Mahasiswa tersebut Adalah:

Gambar 2. Model Antarmuka Data Uji Aplikasi Penentuan Jenis Beasiswa Berbasis K-NN Termodifikasi

Pada gambar 2, pengguna menginputkan 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 (penerima beasiswa Berprestasi atau penerima beasiswa Kurang Mampu). Status data baru tersebut juga akan menambah perbendaharaan

pengetahuan sistem klasifikasi sebagai pengetahuan baru. Hasil proses penentuan jenis beasiswa juga dapat ditampilkan atau dicetak berbentuk Daftar Kelompok penerima beasiswa berdasarkan jenis beasiswa.

### 3. HASIL DAN PEMBAHASAN

Misalkan hasil proses data pada interface gambar adalah seperti disajikan pada table 1.

Tabel 1. Hasil Penentuan Kelas Data Menggunakan Algoritma K-MeanClustering

(1)	(2)	(3)	(4)	(5)
1	3,55	2.400.000	80	Berprestasi
2	2,75	900.000	75	Kurang Mampu
3	3,15	950.000	80	Kurang Mampu
4	2,75	700.000	70	Kurang Mampu
5	3,45	750.000	85	Kurang Mampu
6	3,24	500.000	70	Kurang Mampu
7	3,02	500.000	80	Kurang Mampu
8	3,38	1.500.000	75	Berprestasi
9	3,18	850.000	75	Kurang Mampu
10	3,25	900.000	70	Kurang Mampu

Keterangan:

(1) : Mahasiswa

(2) : Indeks Prestasi Kumulatif (IPK)

(3) : Penghasilan Kepala Keluarga (PKK)

(4) : Nilai Keaktifan dalam Berorganisasi (NKB)

(5) : Jenis Beasiswa Yang Diperoleh (JB)

Data pada table 2 akan menjadi Basis Pengetahuan bagi algoritma *K-NN Classification* dalam proses penentuan kelas data bagi data baru (data uji) yang belum diketahui status kelasnya.

Misalkan terdapat sebuah data baru (mahasiswa ke-11) yang belum diketahui kelasnya: IPK=2,90; PKK=1.650.000 dan NKB=77,5; akan ditetapkan jenis beasiswa yang akan diperoleh. Dengan menetapkan nilai K tetangga terdekat (K=3), dapat dihitung jarak data uji terhadap setiap data training menggunakan persamaan 1.

Jarak data uji terhadap data training 1 adalah:

$$d1 = \sqrt{(3,55 - 2,90)^2 + (2.400.000 - 1.650.000)^2 + (80 - 77,5)^2}$$

$$= 750.000$$

Jarak data uji terhadap data training 2 adalah:

$$d2 = \sqrt{(2,75 - 2,90)^2 + (900.000 - 1.650.000)^2 + (75 - 77,5)^2}$$

$$= 750.000$$

Demikian seterusnya untuk menghitung jarak data uji terhadap data training ke-3 hingga data training ke-10. Hasil selengkapnya disajikan pada table 2.

Tabel 2. Hasil Perhitungan Jarak Data Uji terhadap Data Training

Mhs	IPK	PKK	NKB	JB	Jarak
1	3,55	2.400.000	80	Berprestasi	750,000
2	2,75	900.000	75	Kurang Mampu	750,000
3	3,15	950.000	80	Kurang Mampu	700,000
4	2,75	700.000	70	Kurang Mampu	950,000
5	3,45	750.000	85	Kurang Mampu	900,000
6	3,24	500.000	70	Kurang Mampu	1,150,000
7	3,02	500.000	80	Kurang Mampu	1,150,000
8	3,38	1.500.000	75	Berprestasi	150,000
9	3,18	850.000	75	Kurang Mampu	800,000
10	3,25	900.000	70	Kurang Mampu	750,000

Dengan merujuk pada nilai  $K=3$ , maka dapat ditetapkan 3 data dengan jarak terpendek, yaitu:

- (1) mahasiswa nomor 3 (jarak=700.000) dengan jenis beasiswa (JB) adalah "Kurang Mampu";
- (2) mahasiswa nomor 1 (jarak = 750.000) dengan jenis beasiswa "Berprestasi";
- (3) mahasiswa nomor 2 (jarak = 750.000) dengan jenis beasiswa "Kurang Mampu".

Dari 3 data dengan jarak terpendek, terdapat 2 data yang berstatus JB="Kurang Mampu" dan 1 data berstatus JB="Berprestasi". Dengan demikian, status data uji (mahasiswa ke-11) adalah mendapatkan beasiswa "Kurang Mampu".

#### 4. Kesimpulan

Penggunaan model Klastering untuk menentukan status kelas pada data pelatihan algoritma Klasifikasi, dapat menyelesaikan permasalahan pemodela data pelatihan yang menghasilkan objek kelas yang cenderung dipaksakan.

Pada artikel ini baru diujicoba satu model klastering, yaitu algoritma *K-Means* Klastering untuk menentukan status kelas pada data pelatihan algoritma klasifikasi. Masih perlu dilakukan ujicoba menggunakan model-model Klastering lainnya dan membandingkan hasilnya dengan penggunaan model *K-Mean* Klastering, agar diperoleh hasil yang lebih baik dalam proses penentuan status kelas data pelatihan pada model Klasifikasi.

**Daftar Pustaka**

- [1] Rosmawanti N., Bahar, *Model K-Nearest Neighbor Menggunakan Kombinasi Basis Aturan dan Basis Pengetahuan*, Prosiding, Seminar Nasional Ilmu Komputer 2014, Yogyakarta, 2014.
- [2] Kusriani, Amha T.L., *Algoritma Datamining*, Penerbit ANDI: Yogyakarta, 2008.
- [3] Tedy R., Ardhyta W.I., Wahyu P., Sri Kusumadewi, "Sistem Pendukung Keputusan Berbasis Pocket PC Sebagai Penentu Status Gizi Menggunakan Metode k\_Nearest Neighbor", *Jurnal Teknoin*, vol. 13, no. 2, hal.:1-6, 2008.
- [4] Chusnul I., *Penerapan Case Based Reasoning Dengan Algoritma Nearest Neighbor Untuk Analisis Pemberian Kredit di Lembaga Pembiayaan*, Jurnal Manajemen Informatika, Vol. 2, No. 1, hal.:11-21, 2013.
- [5] Nobertus K., Helmi, Bayu P., *Algoritma k-Nearest Neighbor Dalam Klasifikasi Data Produksi Kelapa Sawit Pada PT. Minamas Kecamatan Parindu*, Jurnal Mimaster, Vol.2, No.1, Hal.:33-38, 2013.
- [6] Prasetyo E., *Data Mining – Mengelolah Data Menjadi Informasi*, Penerbit ANDI: Yogyakarta, 2014.
- [7] Santosa B., *Data Mining – Teknik Pemanfaatan Data Untuk Keperluan Bisnis*, Graha Ilmu: Yogyakarta, 2007.