

Perancangan Sistem Klasifikasi Mahasiswa Drop Out Menggunakan Algoritma K-Nearest Neighbor

I Putu Ramayasa

STMIK STIKOM Bali
Jl. Raya Puputan No 86 Renon Denpasar-Bali, (0361)244445
e-mail: ramayasa@stikom-bali.ac.id

Abstrak

Perguruan tinggi merupakan penyelenggara akademik bagi para mahasiswa. Perguruan tinggi dituntut untuk mampu menyelenggarakan pendidikan yang berkualitas bagi para mahasiswa sehingga mampu menghasilkan sumber daya manusia yang berkualitas. Salah satu permasalahan yang dihadapi perguruan tinggi termasuk STIKOM Bali adalah mahasiswa yang berhenti studi (*drop out*). Tingginya jumlah mahasiswa *drop out* pada STIKOM Bali dapat diminimalisir dengan pengambilan keputusan yang tepat dalam upaya mencegah mahasiswa *drop out*. Untuk menggali sebuah informasi yang penting dalam database untuk membantu pengambilan keputusan dalam meminimalisir mahasiswa *drop out* dibutuhkan suatu teknologi seperti data mining yang dapat menggali informasi berdasarkan pola yang telah ditentukan. Salah satu algoritma dalam data mining yang dapat digunakan untuk menggali informasi adalah algoritma K-Nearest Neighbor. Hasil perancangan berupa data flow diagram, konseptual database dan basis model. Perancangan tersebut menampilkan aliran data yang digunakan untuk kebutuhan input, proses dan output. Hasil pengujian menunjukan akurasi algoritma K-Nearest Neighbor sangat baik.

Kata kunci: Drop out, K-Nearest Neighbor, Data mining

1. Pendahuluan

Data Mining adalah proses penggalian data dari tumpukan database yang berukuran besar yang digunakan untuk menentukan knowledge berupa informasi penting dan bermanfaat bagi suatu organisasi, seperti perguruan tinggi. Pada perguruan tinggi pengimplementasian data mining algoritma K-Nearest Neighbor digunakan untuk mengetahui pola tingkat lulusan mahasiswa sehingga dapat digunakan untuk bahan dalam mengambil sebuah keputusan. STIKOM Bali merupakan salah satu perguruan tinggi yang bergerak di bidang ICT (*Information and Communication Technology*) yang berada di Bali. STIKOM Bali didirikan pada tanggal 10 Agustus 2002 dengan izin dari mendiknas RI No. 157/D/O/2002. STIKOM Bali

Perguruan tinggi merupakan penyelenggara akademik bagi para mahasiswa. Perguruan tinggi dituntut untuk mampu menyelenggarakan pendidikan yang berkualitas bagi para mahasiswa sehingga mampu menghasilkan sumber daya manusia yang berkualitas. Informasi tentang tingkat kelulusan dari mahasiswa STIKOM Bali sangat diperlukan dalam upaya pengambilan keputusan atau kebijakan sehingga dapat meningkatkan pelayanan pendidikan khususnya dibidang ICT. Pemanfaatan data mining dapat digunakan sebagai bahan pertimbangan dalam mengambil keputusan atau kebijakan tentang tingkat kelulusan mahasiswa.

Salah satu permasalahan yang dihadapi perguruan tinggi termasuk STIKOM Bali adalah mahasiswa yang berhenti studi (*drop out*). Tingginya jumlah mahasiswa *drop out* pada STIKOM Bali dapat diminimalisir dengan pengambilan keputusan yang tepat dalam upaya mencegah mahasiswa *drop out* [5]. Untuk menggali sebuah informasi yang penting dalam database untuk membantu pengambilan keputusan dalam meminimalisir mahasiswa *drop out* dibutuhkan suatu teknologi seperti data mining yang dapat menggali informasi berdasarkan pola yang telah ditentukan. Salah satu algoritma dalam data mining yang dapat digunakan untuk menggali informasi adalah algoritma K-Nearest Neighbor.

K-Nearest Neighbor merupakan salah satu algoritma yang digunakan dalam pengklasifikasian. Prinsip kerja K-Nearest Neighbor adalah mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga (*neighbor*) terdekatnya dalam data pelatihan [1][2]. Algoritma K-Nearest Neighbor adalah sebuah algoritma untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data sample yang digunakan adalah data mahasiswa angkatan 2009 sampai 2015 dan data uji yang digunakan adalah data mahasiswa angkatan 2016 dan 2017. Diharapkan dengan pemanfaatan data mining dalam menggali informasi kelulusan mahasiswa dengan menggunakan algoritma K-Nearest Neighbor dapat meminimalisir tingkat *drop out* mahasiswa khususnya di STIKOM Bali.

K-Nearest Neighbor merupakan salah satu algoritma yang digunakan dalam pengklasifikasian. Prinsip kerja K-Nearest Neighbor adalah mencari jarak terdekat antara

data yang akan dievaluasi dengan K tetangga (*neighbor*) terdekatnya dalam data pelatihan. Algoritma *K-Nearest Neighbor* adalah sebuah algoritma untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. *K-Nearest Neighbor* menggunakan algoritma *supervised* dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada *K-Nearest Neighbor*. Tujuan dari algoritma ini adalah mengklasifikasikan obyek baru berdasarkan atribut dan *training sample*. Algoritma *K-Nearest Neighbor* menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari data obyek yang baru. Berikut adalah rumus perhitungan untuk mencari jarak terdekat antara nilai yang lain dengan d sebagai simbol dari jarak [2][3] :

$$d = \sqrt{\sum_i (x_2 - x_1)^2} \dots \dots \dots (1)$$

Berikut adalah penjelasan simbol dari rumus di atas :

x_1 = data uji

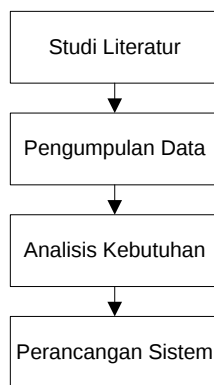
x_2 = sampel data

i = variabel data

d = jarak

2. Pembahasan

Penelitian dilakukan di Kampus STMIK STIKOM Bali. Adapun alur penelitian yang digunakan dapat dilihat pada Gambar 1.

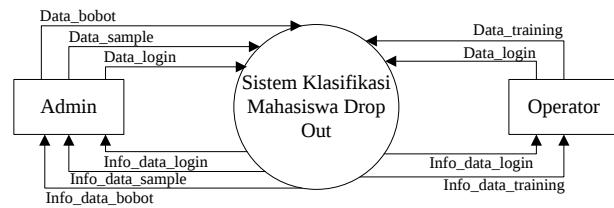


Gambar 1. Alur Penelitian

Alur penelitian diawali dengan melakukan studi literatur yaitu mengumpulkan sumber bacaan yang terkait dengan penelitian seperti algoritma *K-Nearest Neighbor*, studi mahasiswa yang bersumber dari buku, jurnal dan sumber lainnya. Tahap kedua dilakukan pengumpulan data studi mahasiswa angkatan 2009 sampai 2015. Tahap selanjutnya dilakukan analisis kebutuhan sistem yang beracuan pada pengumpulan data dan studi literatur. Tahap terakhir dilakukan perancangan sistem menggunakan *data flow diagram*, konseptual *database* dan basis model.

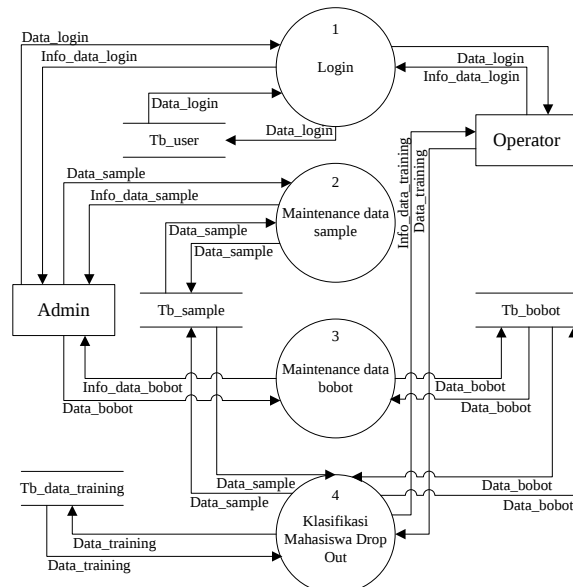
Diagram konteks merupakan diagram yang menggambarkan kegiatan keseluruhan pada [4]. Dalam

diagram konteks terdapat dua aktor yaitu admin dan operator yang masing-masing memiliki hak akses yang berbeda. Gambar 2 merupakan diagram konteks dari sistem klasifikasi mahasiswa *drop out*.



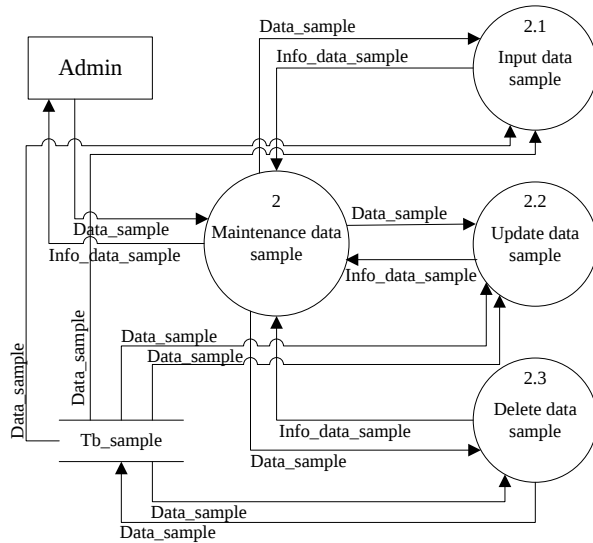
Gambar 2. Diagram Konteks

DFD Level 0 merupakan hasil pemecahan dari diagram konteks. Pada DFD level 0 terdapat empat proses yaitu proses *login*, proses *maintenance data sample*, *maintenance data bobot* dan proses klasifikasi mahasiswa *drop out*. DFD level 0 dapat dilihat pada Gambar 3.



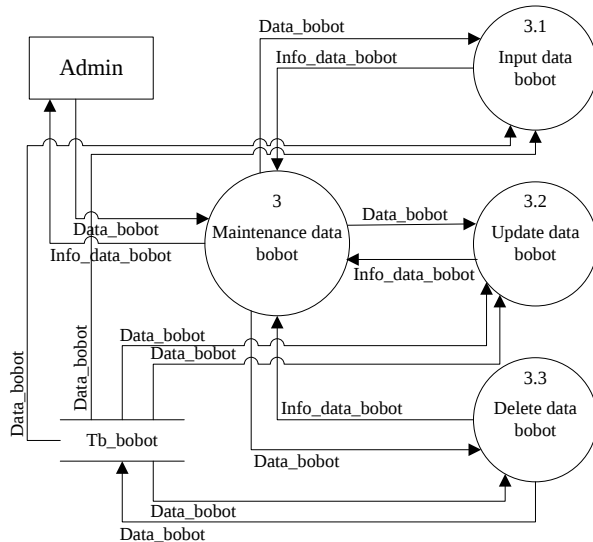
Gambar 3. DFD Level 0

DFD Level 1 Proses *maintenance data sample* merupakan pengembangan dari proses *maintenance data sample*. Pada DFL level 1 ini terdapat tiga proses yaitu *input data sample*, *update data sample* dan *delete data sample*. Proses ini digunakan oleh admin untuk mengelola *data sample*. Gambar 4 merupakan gambaran dari DFD level 1 proses *maintenance data sample*.



Gambar 4. DFD Level 1 Proses Maintenance data sample

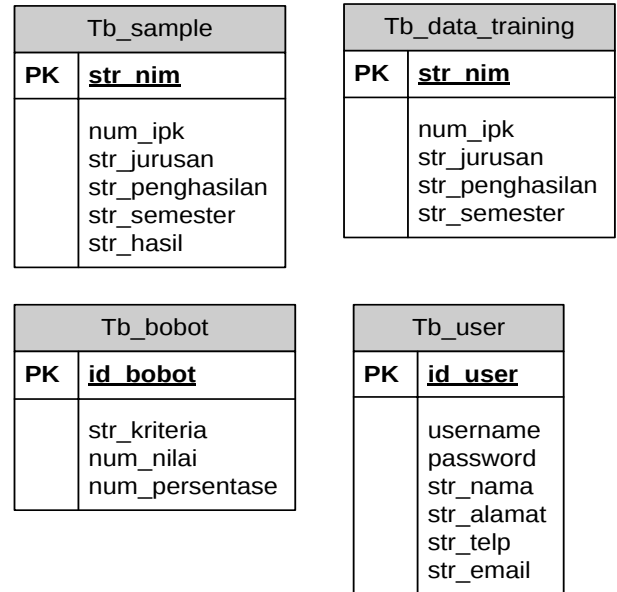
DFD Level 1 Proses *maintenance data* bobot merupakan pengembangan dari proses *maintenance data* bobot. Pada DFL level 1 ini terdapat tiga proses yaitu *input data* bobot, *update data* bobot dan *delete data* bobot. Proses ini digunakan oleh admin untuk mengelola *data* bobot. Gambar 5 merupakan gambaran dari DFD level 1 proses *maintenance data* bobot.



Gambar 5. DFD Level 1 Proses Maintenance data bobot

2.3 Konseptual Database

Konseptual *database* merupakan gambaran rancangan tabel yang digunakan pada klasifikasi mahasiswa *drop out*. Pada sistem ini terdapat empat tabel yaitu tabel *sample*, tabel *bobot*, tabel *data training*, tabel *user*. Gambar 6 merupakan gambaran dari konseptual *database* yang digunakan.



Gambar 6. Konseptual Database

Model pembobotan setiap kriteria yang digunakan dalam perancangan basis model bertujuan untuk menentukan prioritas kriteria dalam mengeluarkan output sistem. Berikut adalah tahapan – tahapan yang dilakukan :

1. Input nilai kriteria
2. Input bobot kriteria
3. Hitung nilai kriteria

$$NK = \frac{\sum_{i=1}^n (SBK) \times BBT \%}{n} \dots\dots\dots (2)$$

$$Nilai Akhir = \frac{\sum NK}{n} \dots\dots\dots (3)$$

Dimana :

SBK : Kriteria

BBT : Bobot Kriteria

NK : Nilai Kriteria

Rancangan model untuk klasifikasi mahasiswa *drop out* adalah sebagai berikut :

a. Kriteria IPK

Model kriteria IPK bertujuan untuk menentukan besaran IPK yang di dapat mahasiswa serta besaran nilai dari masing – masing kriteria IPK. Rentangan nilai untuk kriteria IPK yaitu 40 – 100. Pembobotan dari kriteria IPK dapat dilihat pada Tabel 1.

Tabel 1. Pembobotan IPK

NO	IPK	Nilai	Bobot
1	IPK < 2	100	
2	IPK >= 2 and IPK < 3	80	80%
3	IPK >= 3 and IPK < 3,5	60	
4	IPK >= 3,5	40	

b. Kriteria Jurusan

Model kriteria Jurusan bertujuan untuk menentukan jurusan yang dipilih mahasiswa serta besaran nilai dari masing – masing kriteria jurusan. Rentangan nilai untuk kriteria jurusan yaitu 60 – 100. Pembobotan dari kriteria jurusan dapat dilihat pada Tabel 2.

Tabel 2. Pembobotan Jurusan

NO	Jurusan	Nilai	Bobot
1	Sistem Informasi	100	70%
2	Sistem Komputer	80	
3	Manajemen Informatika	60	

c. Kriteria Penghasilan Orang Tua

Model kriteria Penghasilan orang tua bertujuan untuk menentukan besaran pendapatan orang tua mahasiswa serta besaran nilai dari masing – masing kriteria penghasilan. Rentangan nilai untuk kriteria penghasilan yaitu 40 – 100. Pembobotan dari kriteria penghasilan dapat dilihat pada Tabel 3.

Tabel 3. Pembobotan Penghasilan Orang Tua

NO	Penghasilan	Nilai	Bobot
1	Gaji < 2 jt	100	60%
2	Gaji >= 2 jt & Gaji < 4 jt	80	
3	Gaji >= 4 jt & Gaji < 6 jt	60	
4	Gaji >= 6 jt	40	

c. Kriteria Semester

Model kriteria semester bertujuan untuk mengetahui semester mahasiswa serta besaran nilai dari masing – masing kriteria semester. Rentangan nilai untuk kriteria semester yaitu 60 – 100. Pembobotan dari kriteria semester dapat dilihat pada Tabel 4.

Tabel 4. Pembobotan Semester

NO	Semester	Nilai	Bobot
1	Semester 1	100	50%
2	Semester 2	90	
3	Semester 3	80	
4	Semester 4	70	
5	Semester 5	60	

Struktur tabel merupakan rincian dari konseptual *database*. Masing-masing tabel menampilkan *field-field* yang ada pada tabel tersebut. Selain menampilkan *field* yang ada pada tabel, struktur tabel juga menampilkan tipe data, ukuran dan keterangan dari masing-masing *field* yang ada pada tabel. Struktur tabel yang digunakan pada sistem klasifikasi mahasiswa *drop out* dapat dilihat pada Tabel 5, Tabel 6, Tabel 7 dan Tabel 8.

Tabel 5. Tabel sample

Nama Field	Tipe	Ukuran
str_nim	varchar	10
num_ipk	integer	-
str_jurusan	varchar	20
str_penghasilan	integer	-
str_semester	integer	-
str_hasil	varchar	50

Tabel 6. Tabel data training

Nama Field	Tipe	Ukuran
str_nim	varchar	10
num_ipk	integer	-
str_jurusan	varchar	20
str_penghasilan	integer	-
str_semester	integer	-

Tabel 7. Tabel bobot

Nama Field	Tipe	Ukuran
id_bobot	varchar	10
str_kriteria	varchar	50
num_nilai	integer	-
num_persentase	integer	-

Tabel 8. Tabel user

Nama Field	Tipe	Ukuran
id_user	varchar	10
username	varchar	50
password	varchar	50
str_nama	varchar	100
str_alamat	varchar	200
str_telp	varchar	15
str_email	varchar	50

d. Proses kerja algoritma K-NN :

1. Masukkan data training
2. Tentukan nilai k
3. Hitung jarak dengan rumus euclidean distance
4. Urutkan nilai jarak
5. Cari jarak terdekat sesuai nilai k
6. Menghasilkan prediksi

e. Data Training

NO	NIM	Bobot IPK	Bobot Jurusan	Bobot Penghasilan Orang Tua	Bobot Semester	Keterangan
1	100030244	80	100	80	100	Dropout
2	100030019	100	80	100	90	Dropout
3	100030211	80	100	100	90	Dropout
4	100030335	80	100	80	80	Dropout
5	100030122	60	60	60	70	Lulus
6	100030198	40	60	60	60	Lulus
7	100030196	60	80	40	60	Lulus
8	100030189	60	60	60	70	Lulus
9	100030294	40	80	40	60	Lulus
10	100030272	40	60	60	60	Lulus
11	100030281	100	80	80	80	Dropout
12	100030330	80	100	80	90	Dropout
13	100030203	80	80	100	80	Dropout
14	100030269	40	60	60	70	Lulus
15	100030314	60	80	60	70	Lulus
16	100030251	100	80	100	100	Dropout
17	100030117	100	80	100	100	Dropout
18	100030092	100	100	80	80	Dropout
19	100030038	80	100	80	90	Dropout
20	100030101	100	60	100	90	Dropout

Gambar 7. Data Training

f. Data Uji

NO	NIM	Bobot IPK	Bobot Jurusan	Bobot Penghasilan Orang Tua	Bobot Semester	Keterangan
1	100030116	80	100	100	90	Dropout
2	100030118	60	60	40	60	Lulus
3	100030250	80	60	100	100	Dropout
4	100030166	60	60	60	70	Lulus
5	100030326	60	80	40	60	Lulus

Gambar 8. Data Testing

Masing-masing proses pengujian tersebut menggunakan Indeks Prestasi Kumulatif (IPK), jurusan, penghasilan orang tua, semester dengan menggunakan nilai k yang berbeda. Pengujian ini dilakukan dengan menggunakan data training yang berjumlah sebanyak 20 data dan 5 data serta data uji. Hasil prediksi data uji akan dibandingkan dengan data asli dan dicari kecocokannya. Pengujian ini juga berguna untuk mengetahui apakah nilai k yang digunakan adalah nilai k yang terbaik dengan hasil tingkat keberhasilannya tinggi atau tidak untuk memprediksi kelulusan mahasiswa.

3. Kesimpulan

Berdasarkan uraian dari perancangan sistem klasifikasi mahasiswa *drop out* dapat ditarik kesimpulan bahwa *Data Flow Diagram*, Konseptual *database*, Basis Model dan Struktur tabel dapat digunakan untuk merancang sistem klasifikasi mahasiswa *drop out*. Pada

perancangan yang sudah dilakukan terlihat aliran data yang digunakan untuk kebutuhan *input*, proses dan *output* sistem. Perancangan ini dapat diimplementasikan sehingga dapat meminimalisir tingkat *drop out* mahasiswa khususnya di STIKOM Bali.

Daftar Pustaka

- [1]. A.G. Asmara. "Implementasi Metode Simple Additive Weighting (SAW) Dalam Memprediksi Calon Mahasiswa Dropout STMIK STIKOM Bali", Jurnal Sistem dan Informatika, Vol. 11, No. 1, pp.11-18, 2016.
- [2]. I. W. Supriana, "Sistem Pendukung Keputusan Dalam Pemilihan Tempat Kost Dengan Metode Pembobotan (Studi Kasus : Sleman Yogyakarta) Program Pascasarjana Ilmu Komputer Fakultas MIPA Universitas Gadjah Mada", Jurnal Ilmu Komputer, Vol. 5, No. 2, pp.11-16, 2012.
- [3]. S. Wahyun, "Implementasi Rapidminer Dalam Menganalisa Data Mahasiswa Drop Out", Jurnal Abdi Ilmu, Vol. 10, No.2 , pp.1899-1902, 2018.
- [4]. R. S. Pressman. "Rekayasa Perangkat Lunak Pendekatan Praktisi Buku1 dan 2", Yogyakarta: Andi. 2012.
- [5]. L. Selvia. W. B. Ginting. A.D. Zarman, "Teknik Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighborhood", Jurnal Teknik Komputer Unikom, Vol. 3, No. 2, pp. 29-34, 2014.