

PROPOSAL PROGRAM PENELITIAN

Jenis Program / Kluster

PREDIKSI PENERIMA BEASISWA MENGGUNAKAN K-NEAREST NEIGHBOR



Nama Peneliti (Muafi, S.Ag., M.Kom)

NIDN :07050776009

UNIVERSITAS NURUL JADID

FAKULTAS TEKNIK

MEI 2018

Judul Penelitian : PREDIKSI PENERIMA BEASISWA MENGGUNAKAN
K-NEAREST NEIGHBOR
Bidang Fokus : Fakultas Teknik Universitas Nurul Jadid

Peneliti

- a. Nama Lengkap : Muafi, S.Ag., M.Kom
- b. NIDN : 07050776009
- c. Jabatan Fungsional : -
- d. Program Studi : Teknik Informatika
- e. Nomor Hp : 085231713772
- f. Alamat surel (Email) : muafiumar514@gmail.com

Lama Penelitian : 1 Tahun

Usulan Penelitian Tahun ke- : 2018

Total Biaya Penelitian : 4800000

Asal Biaya Penelitian : -

- Dana Internal (Tunjangan)
- Dana eksternal (hibah)
- Lain-lain.....

Disahkan pada.....

Di.....

Mengetahui,

Dekan Fakultas

Peneliti,

Moh. Furqan, M.Kom

NIDN:.....

Muafi, S.Ag., M.Kom

NIDN : 07050776009

PREDIKSI PENERIMA BEASISWA MENGUNAKAN K-NEAREST NEIGHBOR

ABSTRAK

Pendidikan adalah sebuah sarana untuk membentuk moral dan etika, agar siswa tidak hanya berorientasi pada proses wawasan keilmuan, tapi bagaimana dalam proses belajar mengajar bisa mengubah tingkah laku untuk menuju kearah yang lebih baik. Sehingga perlu sarana penunjang dalam menentukan pemberian beasiswa sebagai motivasi.

Data yang akan digunakan dalam penelitian ini adalah data penerimaan beasiswa tahun 2015/2016 dari kelas 7 sampai kelas 8 pada semester ganjil, adapun metode yang akan digunakan untuk *menyelesaikan* persoalan yang telah dipaparkan sebelumnya adalah menggunakan teknik data mining yaitu klasifikasi, tujuannya adalah untuk mengekstrak data lebih dalam lagi sehingga tidak hanya informasi yang didapat tapi juga sebuah pengetahuan baru yang didapat dari sebuah data. Sedangkan metode yang akan digunakan dalam penelitian ini adalah KNN, KNN merupakan kepanjangan dari K-Nearest Neighbor lebih detailnya KNN adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut.

Keyword : Klasifikasi, data mining KNN

BAB I

PENDAHULUAN

1.1 Latar Belakang

Lembaga pendidikan adalah lembaga atau tempat berlangsungnya proses pendidikan atau belajar mengajar yang dilakukan dengan tujuan untuk mengubah tingkah laku individu menuju ke arah yang lebih baik melalui interaksi dengan lingkungan sekitar, Lembaga pendidikan yang ada di Indonesia terbagi menjadi 3 yaitu Pendidikan formal, pendidikan nonformal, dan pendidikan informal.

Tidak dapat dipungkiri bahwa pendidikan sebuah sarana untuk membentuk moral, cendekiawan, sastrawan, dan bidang-bidang yang lain, tanpa adanya pendidikan mustahil sebuah bangsa bisa berkembang, SMK Al Ishlah Besuki merupakan sekolah menengah kejuruan yang berada di kawasan Kabupaten Situbondo Kecamatan Besuki yang berbasis Informatika dan Multimedia dan sudah memiliki banyak alumni yang bisa dikatakan sukses, untuk meningkatkan mutu pendidikan yang ada di SMK Al Ishlah Besuki salah satunya adalah memberikan beasiswa bagi siswa yang berprestasi, tentunya penerimaan beasiswa ini harus tepat sasaran agar tidak siswa yang dirugikan dalam hal ini. Salah satu cara untuk menentukan penerimaan beasiswa ini adalah melihat secara objektif data siswa namun hal ini rentan dengan kesalahan, tidak jarang cenderung pada nepotisme dan kedekatan semata, untuk mencegah hal tersebut dibutuhkan sebuah *algoritma data mining* untuk mengekstrak data lebih dalam lagi sehingga bisa membuat pihak sekolah lebih bijak dalam menentukan siswa yang berhak menerima beasiswa.

Data yang akan digunakan dalam penelitian ini adalah data penerimaan beasiswa tahun 2016/2017 dari kelas 7 sampai kelas 8 pada semester ganjil, adapun metode yang akan digunakan untuk menyelesaikan persoalan yang telah dipaparkan sebelumnya adalah menggunakan teknik data mining yaitu klasifikasi, tujuannya adalah untuk mengekstrak data lebih dalam lagi sehingga tidak hanya informasi yang didapat tapi juga sebuah pengetahuan baru yang didapat dari sebuah data. Sedangkan metode yang akan digunakan dalam penelitian ini adalah KNN, KNN merupakan kepanjangan dari K-Nearest Neighbor lebih detailnya KNN adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut.

Pada fase pembelajaran, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi dari data pembelajaran. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk data test (yang klasifikasinya tidak diketahui). Jarak dari vektor yang baru ini terhadap seluruh vektor data pembelajaran dihitung, dan sejumlah k buah yang paling dekat

diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut.

Ketepatan algoritma KNN ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. Riset terhadap algoritma ini sebagian besar membahas bagaimana memilih dan memberi bobot terhadap fitur, agar performa klasifikasi menjadi lebih baik.

Sesuai dengan permasalahan diatas, maka untuk itu peneliti sangat tertarik untuk membahasnya dalam penelitian yang berjudul : ” **APLIKASI UNTUK MENENTUKAN PENERIMA BEASISWA MENGGUNAKAN K-NEAREST NEIGHBOR**” . Agar dapat menentukan penerima beasiswa kedepannya, sehingga bisa menaikkan persentase penerima beasiswa yang ada dilembaga pendidikan khususnya SMK Al Ishlah Besuki, adapun atribut yang digunakan untuk memprediksi penerima beasiswa yaitu Kehadiran, Ekstra, Ranking, Nilai raport semester akhir, Penghasilan orang tua dan Pelanggaran siswa.

1.2 Tujuan Penelitian

Tujuan dari penelitian ini adalah menentukan dan membangun system pemberian beasiswa menggunakan teknik *datamining* dengan metode *KNN* untuk mengektrak data lebih dalam sehingga tidak hanya informasi yang didapat. Sehingga pemberian beasiswa sesuai kriteria dengan *atribut* dalam penerima beasiswa.

1.3 Rumusan Masalah

Bagaimana membuat system untuk memprediksi penerima beasiswa dengan mengektrak data menggunakan Algoritma Klasifikasi KNN.

BAB II

2.1 Tinjauan Pustaka

Pada jurnal penelitian yang di pilih oleh Mohammad Syaifudin menjelaskan tentang penggunaan Algoritma KNN dengan studi kasus penerimaan siswa miskin di SDN Pagung 1, dengan attribute tanggungan orang tua, Pekerjaan orang tua, Penghasilan orang tua, adapun sample data sebagai berikut: (M.Syaifudin:2009)

Tabel 2.1 Sample Data Training

Siswa	Tanggungan Orang Tua	Pekerjaan Orang Tua	Penghasilan Orang Tua	Status BSM
Amin	3	Swasta	<1.500.000	Dapat
Bella	2	Petani	<750.000	Dapat
Cindy	1	Wiraswasta	<1.500.000	Tidak
Desi	2	PNS	>2.500.000	Tidak

Pada tabel 2.1 masing masing di normalisasi dengan tingkat pengaruh yang berbeda - beda untuk menentukan kelayakan menerima BSM, adapun data normalisasi nampak pada tabel 2.2

Tabel 2.2 Data Normalisasi

Siswa	Tanggungan Orang Tua	Pekerjaan Orang Tua	Penghasilan Orang Tua	Status BSM
Amin	3	3	2	Dapat
Bella	2	1	1	Dapat
Cindy	1	2	2	Tidak
Desi	2	4	3	Tidak

Pada tabel 2.2 kita sudah menormalisasi data menjadi angka setelah itu kita hitung dengan menggunakan rumau KNN. Sebelum kita menghitung kita tentukan dulu data testing yang mau kita hitung. Berikut data testing yang harus kita hitung namapak pada tabel 2.3

Tabel 2.3 Data Testing

Siswa	Tanggungan Orang Tua	Pekerjaan Orang Tua	Penghasilan Orang Tua	Status BSM
Santi	2	2	2	?

Berikut ini adalah contoh proses perhitungan pada penelitian ini:

$$\begin{aligned}
D(a,b)_1 &= \sqrt{\sum_{k=1}^d (a - b)^2} \\
&= \sqrt{(2 - 3)^2 + (2 - 3)^2 + (3 - 2)^2} \\
&= \sqrt{(-1)^2 + (-1)^2 + (1)^2} \\
&= \sqrt{3} = 1,7
\end{aligned}$$

$$\begin{aligned}
D(a,b)_2 &= \sqrt{\sum_{k=1}^d (a - b)^2} \\
&= \sqrt{(2 - 2)^2 + (2 - 1)^2 + (3 - 1)^2} \\
&= \sqrt{(0)^2 + (1)^2 + (2)^2} \\
&= \sqrt{5} = 2,2
\end{aligned}$$

$$\begin{aligned}
D(a,b)_3 &= \sqrt{\sum_{k=1}^d (a - b)^2} \\
&= \sqrt{(2 - 1)^2 + (2 - 2)^2 + (3 - 2)^2} \\
&= \sqrt{(1)^2 + (0)^2 + (1)^2} \\
&= \sqrt{2} = 1,4
\end{aligned}$$

$$\begin{aligned}
D(a,b)_1 &= \sqrt{\sum_{k=1}^d (a - b)^2} \\
&= \sqrt{(2 - 2)^2 + (2 - 4)^2 + (3 - 3)^2} \\
&= \sqrt{(0)^2 + (-2)^2 + (0)^2} \\
&= \sqrt{4} = 2
\end{aligned}$$

Produk software yang dihasilkan dalam penelitian ini diharapkan dapat mempercepat proses pemilihan calon siswa yang berhak mendapatkan BSM. Dengan menerapkan metode KNN dalam aplikasi penerima BSM ini dapat menghindari terjadinya kesalahan dalam menentukan calon siswa penerima BSM. Sehingga aplikasi ini dapat membantu kinerja guru untuk menjadi lebih cepat dan efisien

Abdul Rohman pada jurnal penelitian ini menjelaskan tentang memprediksi kelulusan mahasiswa . Dalam sistem pendidikan mahasiswa adalah aset penting bagi sebuah institusi pendidikan dan untuk itu perlu diperhatikan tingkat kelulusan mahasiswa tepat pada waktunya. Prosentase naik turunnya kemampuan mahasiswa untuk menyelesaikan studi tepat waktu merupakan salah satu elemen penilaian akreditasi universitas. Untuk itu perlu adalah pemantauan dan evaluasi terhadap kelulusan mahasiswa dengan menggunakan klasifikasi data mining. Dengan mengolah data mahasiswa akan memperoleh hal penting

untuk keperluan perguruan tinggi terutama untuk prediksi kelulusan mahasiswa. Dalam penelitian ini dalam memprediksi kelulusan mahasiswa dengan menggunakan algoritma klasifikasi data mining *K-Nearest Neighbor* dengan mengklaster *data k=1, k=2, k=3, k=4, dan k=5*. Hasil yang diperoleh dengan cluster data *k=5 accuracy* adalah 85,15% dan nilai *AUC* adalah 0.888 adalah akurasi paling tinggi. (Abdul Rohman:2007)

Tabel 2.4 Skema Teknik KNN



Tabel 2.4 menggambarkan urutan dari penelitian yang telah dilakukan oleh rohman untuk mendapatkan hasil dari penelitiannya, dimulai dari masalah kemudian metode yang akan digunakan untuk menyelesaikan masalah tersebut yaitu KNN, setelah metode lanjut dengan pengembangan yang menggunakan rapidminer, selanjutnya ada penerapannya pada sebuah data mahasiswa yang lulus tepat waktu dan tidak tepat waktu, untuk keberhasilan prediksinya indikatornya menggunakan confusion matrix dan kurva ROC, dari indicator tersebut terbentuklah sebuah hasil yang berupa akurasi ketepatan yang dihasilkan dari

sebuah data menggunakan metode KNN, Adapun hasil akurasi dalam penelitian ini seperti tabel 2.5

Tabel 2.5
Konfusión Matrix Penelitian Abdul Rohman

Cluster	Akurasi
<i>K1</i>	82,25 %
<i>K2</i>	79,45 %
<i>K3</i>	83,95 %
<i>K4</i>	82, 62 %
<i>K5</i>	85,15 %

Tabel 2.5 menggambarkan tentang teori. Cluster K1. K2, K3 dan seterusnya itu adalah hasil dari perhitungan akurasi dari sistem. Jadi kita bisa tahu data itu validnya berapa persen jikan digunakan untuk prediksi.

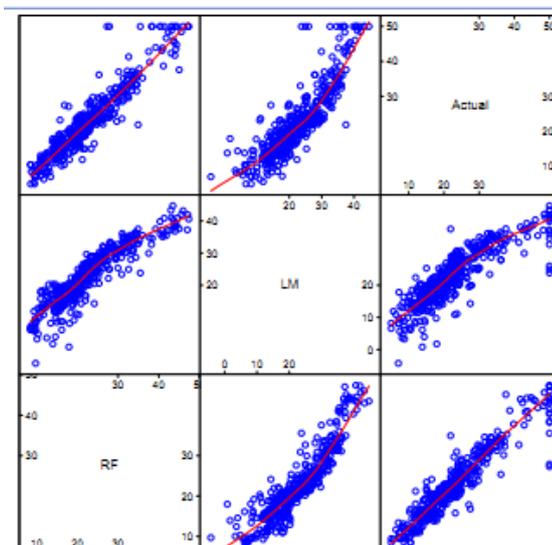
Penelitian ini menjelaskan bagaimana menggunakan tekni Klasifikasi dan regresi dengan menggunakan RandomForest, pada penelitian ini menjelaskan tentang Klasifikasi dan regresi dengan tapi dengan metode yang sama yaitu randomForest

Berikut ini adalah hasil akurasi dari teknik random forest pada penelitian ini:

Confusion Matrix :

	WinF	WinNF	Veh	on	Tabl	Head	Class.error
WinF	63	6	1	0	0	0	0.1000000
WinNF	9	62	1	2	2	0	0.1842105

Gambar 2.1 merupakan representasi dari hasil perhitungan rando, forest dengan linear model yang datanya didapat dari response of the Boston Housing data.



Gambar 2.1 Hasil Comparasi antara Random Forest dengan Linear Model dengan

aktual response of The Boston Housing data

2.2 Kontribusinya

Penerapan datamining dalam memprediksi pemberian beasiswa dengan menggunakan KNN

BAB 3 METODE

3.1 Rancangan Penelitian

Dalam penelitian ini memerlukan suatu metode untuk mengumpulkan data yang dibutuhkan, karena metode atau teknik merupakan salah satu syarat dalam sebuah penelitian. Disamping adanya obyek yang sistematis dalam memilih dan menggunakan metode ini harus relevan dengan pokok permasalahan agar nantinya diperoleh hasil yang dapat dipertanggung jawabkan kebenarannya. Pada penelitian ini metode penelitian yang digunakan adalah metode kualitatif.

Metode penelitian kualitatif adalah metode penelitian yang berlandaskan pada filsafat *postpositivisme*, digunakan untuk meneliti pada kondisi obyek yang alamiah, (sebagai

lawannya adalah eksperimen) dimana peneliti adalah sebagai instrumen kunci, teknik pengumpulan data dilakukan secara *triangulasi* (gabungan), analisis data bersifat induktif atau kualitatif, dan hasil penelitian kualitatif lebih menekankan makna dari pada *generalisasi*. (Sugiyono, 2014).

Metode penelitian ini menggunakan metode kualitatif, karena berupaya memahami dan mengkaji fenomena atau peristiwa, orang, objek, atau proses yang terkait dengan sistem yang berada di SMK Al Ishlah Besuki. Penelitian kualitatif ditujukan untuk memahami fenomena-fenomena sosial orang-orang yang diajak berwawancara, diobservasi, diminta memberikan data, pendapat, dan pemikiran agar penelitian dapat dilakukan dengan sesuai dari pengamatan semua objek yang diteliti serta mengungkapkan semua fakta yang ada pada saat penelitian berlangsung mengenai sistem penentuan penerimaan beasiswa di SMK Al Ishlah Besuki.

3.2 Subjek Penelitian

Subjek penelitian adalah subjek yang dituju untuk diteliti. Dalam penelitian ini yang menjadi subjek utama penelitian adalah SMK Al Ishlah Besuki mengenai proses penentuan Beasiswa di subjek pendukung adalah data wali siswa sebagai untuk menentukan kelayakan penerimaan beasiswa tersebut. Yang mana SMK Al Ishlah Besuki berada di daerah Kabupaten Situbondo Jawa Timur

3.3 Instrument Penelitian

Pada instrument penelitian terdapat dua hal utama yang mempengaruhi kualitas hasil penelitian yaitu, instrument pengumpulan data dan Instrument analisis data.

3.4 Instrument Pengumpulan Data

Instrumen dalam pengumpulan data dari penelitian ini terdapat 3 bagian berupa:

3.4.1 Observasi

Adapun bahan observasi dalam instrument pengumpulan data ini adalah sistem penentuan selama ini menggunakan manual. Sistem manual ini sudah dikatakan objektif tapi masih rawan manipulasi data karena masih aplikasi excel yang datanya bisa dimanipulasi oleh oknum yang tidak bertanggung jawab nantinya, maka dibutuhkan sebuah aplikasi yang bisa menentukan penentuan penerimaan beasiswa secara otomatis, dan solusi yang cocok dalam kasus ini adalah data mining dengan algoritma KNN, karena perhitungan KNN bisa dipertanggung jawabkan.

3.4.2 Wawancara

Wawancara di lakukan dengan Petugas TU Lembaga SMK Al Ishlah Besuki sebagai Subjek penelitian.

3.4.3 Studi Literatur

Studi literature adalah mencari referensi teori yang relevan dengan kasus atau permasalahan yang ditemukan .Referensi tersebut berisi tentang:

- SMK Al Ishlah Besuki Secara Umum
- Beasiswa
- Sistem –sistem yang berada di Penentuan penerima Beasiswa
- Teori yang mendukung penelitian

Referensi ini dapat di cari dari buku, jurnal, artikel, laporan penelitian, dan situs – situs internet.Hasil dari studi literature ini adalah terkoleksinya referensi yang relevan.Tujuannya adalah untuk memperkuat permasalahan serta sebagai dasar teori dalam melakukan studi literatur.

3.5 Instrument Analisis Data

Instrument analisis data merupakan komponen penting yang harus diperhatikan dalam melakukan suatu penelitian. Berikut ini rincian alat dan bahan yang digunakan dalam penelitian ini.Adapun kebutuhan yang dipakai dalam penelitian ini meliputi perangkat keras (*Hardware*) dan perangkat lunak (*Software*) komputer.

1. Kebutuhan Perangkat Keras

Dalam menerapkan dari rancangan yang telah dijelaskan sebelumnya, dibutuhkan beberapa perangkat keras untuk membuat aplikasi penentu penerimaan beasiswa di SMK Al Ishlah Besuki adalah:

Laptop digunakan untuk membangun sistem informasi dengan spesifikasi sebagai berikut :

- Memory 2.00 GB
- Hardisk 500 GB
- Processor Intel (R)DualCore
- CPU B820 @1.70GHz
- Printer
- Flashdisk

2. Kebutuhan Perangkat Lunak

Dalam menerapkan rancangan yang telah dibuat, dibutuhkan beberapa *software* untuk membuat program pada aplikasi penentuan Beasiswa di SMK Al Ishlah Besuki yaitu:

- a. Netbeans
- b. Database PHP MySQL (XAMPP)
- c. Power Designer
- d. Sistem Operasi Microsoft Windows 7.
- e. Microsoft Office 2010 untuk menyusun laporan.

3.6 Instrument pengujian program

3.6.1 Unit Testing Black Box

Dalam pengujian program pada aplikasi penentuan penerima beasiswa ini, peneliti menggunakan teknik pengujian Black Box karena Metode Black Box memungkinkan perekayasa perangkat lunak mendapatkan serangkaian kondisi input yang sepenuhnya menggunakan semua persyaratan fungsional untuk suatu program.

Adapun alasannya menggunakan metode Black Box adalah dapat menemukan kesalahan dalam kategori berikut :

- Fungsi-fungsi yang tidak benar atau hilang
- Kesalahan interface
- Kesalahan dalam struktur data atau akses basis data eksternal
- Inisialisasi dan kesalahan terminasi
- Validitas fungsional
- Kesensitifan sistem terhadap nilai input tertentu
- Batasan dari suatu data

3.6.2 Sistem Testing

Unit Testing merupakan pengujian terhadap program yang nantinya akan dibuat. Untuk melakukan pengujian pada program tersebut peneliti lebih menekankan pada pengujian secara *Black Box*, karena pengujian berdasarkan apa yang dilihat dan fokus terhadap fungsionalitas dan output.

3.7 Teknik Pengumpulan Data

Teknik pengumpulan data merupakan langkah yang paling strategis dalam penelitian, karena tujuan utama dari penelitian adalah mendapatkan data. Tanpa mengetahui teknik

pengumpulan data, maka peneliti tidak akan mendapatkan data yang memenuhi standar data yang ditetapkan.

Teknik pengumpulan data yang dipergunakan dalam penelitian ini adalah:

1. Observasi

Penelitian ini mengadakan pengamatan langsung terhadap objek yang dijadikan tempat penelitian yaitu di SMK Al Ishlah Besuki bersama dengan petugas TU ini dilakukan selama 2 bulan. Setelah meneliti objek yang di lihat yaitu sistem yang berada pada lembaga pendidikan ini merupakan sistem penentuan penerimaan beasiswa yang kurang efisien dan memerlukan waktu yang relative lebih banyak.

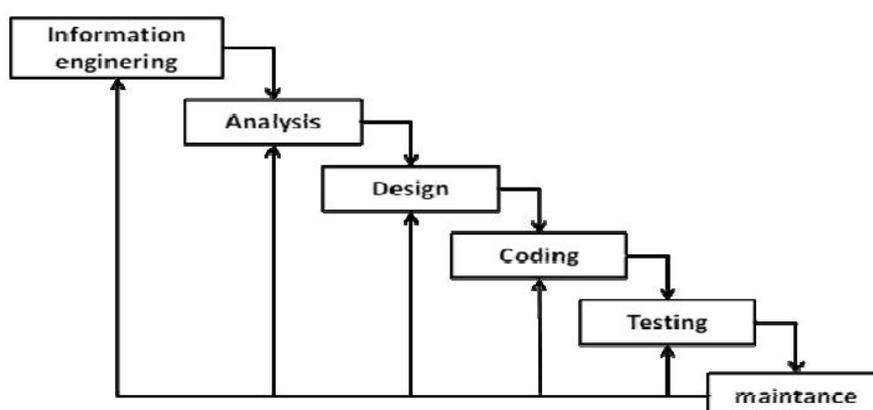
2. Wawancara

Pengumpulan data informasi dilakukan dengan cara wawancara secara langsung dengan Kepala Sekolah SMK Al Ishlah Besuki dilakukan pada saat pelaksanaan penelitian hari sabtu tanggal 10 Desember 2016 siang hari pukul 10.45 wib

Hasil dari wawancara ini menyimpulkan bahwa sistem penentuan penerimaan beasiswa ini membutuhkan kebutuhan-kebutuhan baru yang belum terakomodasi pada sistem lama. Seperti perlunya sistem informasi Pendataan Siswa di SMK Al Ishlah Besuki berbasis dengan MySQL agar penyampaian informasi terkait dengan data siswa secara detail mengingat pentingnya data siswa dalam penentuan beasiswa di SMK Al Ishlah Besuki.

3.8 Teknik Analisis Data Dan Pengembangan Sistem

Metode *waterfall* ini merupakan metode yang sering digunakan oleh penganalisa sistem pada umumnya. Metode waterfall adalah paradikma rekayasa lunak yang paling luas dipakai dan paling tua. Model waterfall mengusulkan sebuah pendekatan kepada perkembangan perangkat lunak yang sistematis dan sekuensial yang mulai pada tingkat dan kemajuan sistem pada analisis, desain, kode, pengujian dan pemeliharaan. Setelah setiap tahap didefinisikan, tahap tersebut diakhiri (signed off) dan pengembangan berlanjut ketahap berikutnya.(Firman, 2010)



G

ambar 3.1 Model Waterfall. (Firman Kemal P ,2010)

Secara garis besar metode waterfall mempunyai langkah-langkah sebagai berikut :

1. *Information Engineering*(Rekayasa Sistem)

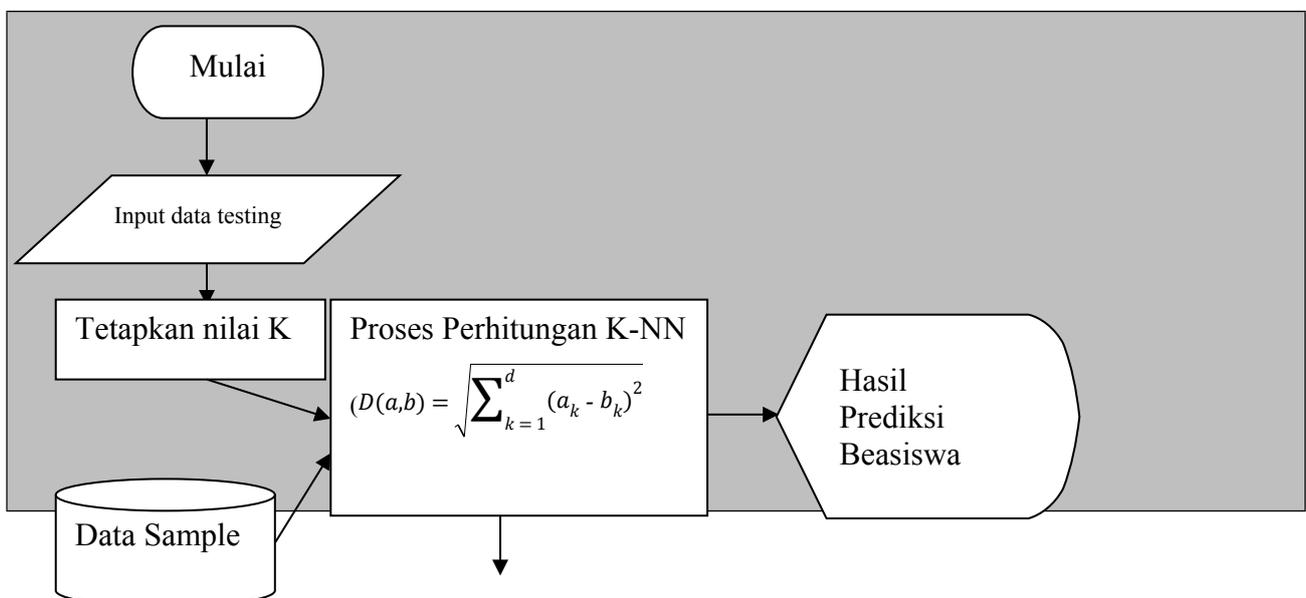
Sistem baru yang akan dibuat membutuhkan data dan informasi dari sistem lama. Data dan informasi dalam sistem lama dapat membantu dalam pembuatan sistem baru, informasi tersebut diperoleh dari metode pengumpulan data yang dilakukan di SMK Al Ishlah Besuki.

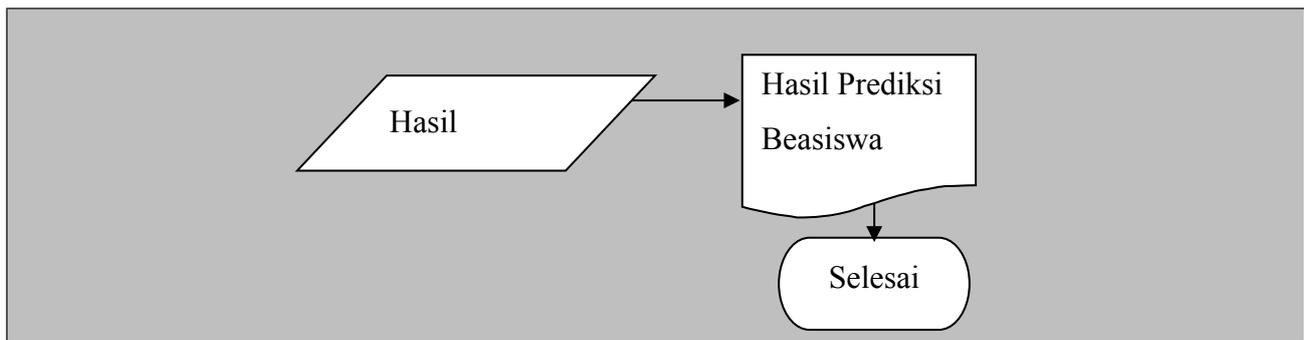
2. *Analysis* (Analisis Kebutuhan)

Analisis sistem dilakukan dengan wawancara, observasi dan study literatur. Hasil wawancara dan observasi yang diperoleh adalah sistem lama yang digunakan masih manual dan perlu ada perbaikan sistem. Sistem baru tersebut adalah sistem yang diharapkan dapat membantu pihak lembaga sekolah.

3. *Design* (Desain)

Setelah menganalisa permasalahan yang ada dalam pada penentuan Beasiswa di SMK Al Ishlah Besuki berbasis web dan MySQL, selanjutnya digunakan perancangan sistem dengan menggunakan flowchart yang merekam persyaratan fungsional sebuah sistem, selanjutnya Class Diagram merupakan himpunan dari objek-objek yang sejenis digunakan untuk alur sistem yang akan dikembangkan, selain *Flowchart* dan Class Diagram perancangan sistem juga membutuhkan *database* atau basis data, karena suatu basis data adalah tempat penyimpanan file data. Sebagai file data, suatu basis data tidak menyajikan informasi secara langsung kepada pengguna.





Gambar 3.2 Flow chart Proses KNN

Dilihat dari gambar 3.2 proses prediksi siswa yang akan menerima beasiswa diawali dengan menginputkan data penerima beasiswa terdahulu untuk dijadikan data testing sebagai acuan untuk memprediksi siswa berikutnya yang akan menerima beasiswa dengan menggunakan algoritma KNN, Kemudian masukan data yang akan di prediksi kemudian hitung dengan menggunakan Algoritma kemudian hasilnya tampil pada layar, dan selanjutnya adalah proses print yang kemudian akan menjadi sebuah laporan hasil prediksi beasiswanya.

4. *Coding* (Penulisan Program)

Dalam tahap ini yaitu pengkodean merupakan menerjemahan dari desain ke bahasa yang bisa dikenal oleh komputer yang dilakukan oleh programmer untuk mengerjakan suatu sistem. Setelah pengkodean selesai maka akan dilakukan testing terhadap sistem yang telah dibuat.

5. *Testing* (Uji Coba)

Setelah sistem baru selesai dibuat maka dilakukan ujicoba terlebih dahulu sebelum diberikan kepada user, dalam hal ini adalah pihak dari instansi SMK Al Ishlah Besuki bagian system informasi (sinfo) yang menangani sistem tersebut apakah sudah sesuai dengan yang diinginkan atau sebaliknya. Kalau tidak sesuai dengan apa yang diinginkan user sistem tersebut direvisi dan dilakukan ujicoba kembali setelah sistem tersebut direvisi.

6. *Maintenance* (Pemeliharaan)

Setelah melakukan pengujian terhadap sistem baru yang dibuat, tahap berikutnya adalah perawatan sistem. Ketika dijalankan mungkin saja masih ada *errors* yang tidak ditemukan sebelumnya, atau ada penambahan fitur-fitur pada software yang dibuat. Maka dari itu perlu adanya perawatan dan pengembangan sistem, validasi data dan updating data

3.9 K- NEAREST NEIGHBOR (KNN)

K- Nearest Neighbor merupakan salah satu metode yang digunakan dalam pengklasifikasian. Prinsip kerja K-Nearest Neighbor adalah mencari jarak terdekat antara data yang akan dievaluasi dengan KNN terdekatnya dalam data pelatihan.

Berikut rumus pencarian jarak terdekat menggunakan rumus Euclidin.

$$D(a,b)_1 = \sqrt{\sum_{k=1}^d (a - b)^2}$$

Dengan :

a : data uji

b : sampel data

k : variable data

dist : jarak

d : dimensi data

Contoh Soal

1. Sample Data Training

Siswa	Ekstra	Alpha	Sakit	Ijin	Penghasilan Orang Tua	Pelanggan	Ranking	Raport	Status BSM
Adelia	1	8	1	2	41	1	2	73	Tidak
Ageng	2	8	4	3	16	5	2	76	Tidak
Putri	2	5	1	2	16	1	1	98	Dapat
Riky	2	7	1	1	41	0	1	97	Dapat

2. Data Testing

Siswa	Ekstra	Alpha	Sakit	Ijin	Penghasilan Orang Tua	Pelanggan	Ranking	Raport	Status BSM
Joni	1	12	6	2	16	5	2	70	?

Cara perhitungannya

$$D(a,b)_1 = \sqrt{\sum_{k=1}^d (a - b)^2}$$

=

$$\begin{aligned} & \sqrt{(1 - 1)^2 + (12 - 8)^2 + (6 - 1)^2 + (2 - 2)^2 + (16 - 41)^2 + (5 - 1)^2 + (2 - 2)^2 + (70 - 73)^2} \\ & = \sqrt{(0)^2 + (4)^2 + (5)^2 + (0)^2 + (-25)^2 + (4)^2 + (0)^2 + (-3)^2} \\ & = \sqrt{691} = 26,2 \end{aligned}$$

$$D(a,b)_2 = \sqrt{\sum_{k=1}^d (a - b)^2}$$

$$\begin{aligned}
&= \\
&\quad \sqrt{(1-2)^2 + (12-8)^2 + (6-4)^2 + (2-3)^2 + (16-16)^2 + (5-5)^2 + (2-2)^2 + (70-73)^2} \\
&= \sqrt{(-1)^2 + (4)^2 + (2)^2 + (-1)^2 + (0)^2 + (0)^2 + (0)^2 + (-6)^2} \\
&= \sqrt{58} = 7,6
\end{aligned}$$

$$D(a,b)_3 = \sqrt{\sum_{k=1}^d (a-b)^2}$$

$$\begin{aligned}
&= \\
&\quad \sqrt{(1-2)^2 + (12-5)^2 + (6-1)^2 + (2-2)^2 + (16-16)^2 + (5-1)^2 + (2-1)^2 + (70-98)^2} \\
&= \sqrt{(-1)^2 + (7)^2 + (5)^2 + (0)^2 + (0)^2 + (4)^2 + (1)^2 + (-28)^2} \\
&= \sqrt{876} = 29,5
\end{aligned}$$

$$D(a,b)_3 = \sqrt{\sum_{k=1}^d (a-b)^2}$$

$$\begin{aligned}
&= \\
&\quad \sqrt{(1-2)^2 + (12-7)^2 + (6-11)^2 + (2-1)^2 + (16-41)^2 + (5-0)^2 + (2-1)^2 + (70-97)^2} \\
&= \sqrt{(-1)^2 + (5)^2 + (5)^2 + (1)^2 + (-25)^2 + (5)^2 + (1)^2 + (-27)^2} \\
&= \sqrt{1591} = 39,5
\end{aligned}$$

Jadi nilai yang terdekat adalah 7,6

BAB 4
LUARAN DAN TARGET CAPAIAN

Memprediksi siswa lebih awal sehingga bisa siswa didik lebih awal dengan menggunakan algoritma K-NN dengan cara mencari nilai/arak terdekat dari data sebelumnya, membantu pihak sekolah dalam menentukan penerima beasiswa secara adil dengan kriteria kriteria dari data sebelumnya

BAB 5
RENCANA ANGGARAN BIAYA

No	Keterangan	Nominal
1	Pengumpulan data training	1.300.000
2	Pengodingan dengan tenaga progremer free line	3.000.000
3	Training kepada petugas	500.000
Total		4.800.000

DAFTAR PUSTAKA

1. Mohammad Syaifudin, Penggunaan Algoritma KNN dengan Studi Kasus Penerimaan Siswa Miskin di SDN Pagung (2009)
2. Abdul Rahman, Prediksi Kelulusan Mahasiswa dengan Menggunakan Algoritma Klasifikasi Data Mining *K-Nearest Neighbor* (2009)
3. Muafi, Prediksi Kelulusan Mahasiswa Menggunakan Algoritma C4.5 dengan Metode Forward Selection (2012)
4. Hanang Wijayanto, Klasifikasi batik menggunakan metode k-nearest neighbour berdasarkan *Gray Level Co-Occurrence Matrices (GLCM)*. 2016
5. Faid dkk, Praktek dan Teori Data Mining dengan Weka dan NNetbeans. 2017. Duta Media