SKRIPSI

SISTEM REKOMENDASI PENGAMBILAN KATEGORI SKRIPSI DENGAN PERHITUNGAN NAIVE BAYES

Oleh:

Muhammad Amdanibik

11.1.03.0.0257

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS NUSANTRA PGRI KEDIRI

2015

SKRIPSI

SISTEM REKOMENDASI PENGAMBILAN KATEGORI SKRIPSI DENGAN PERHITUNGAN NAÏVE BAYES

Telah disetujui untuk diajukan Kepada Panitia Ujian/Sidang Skripsi

Program Studi Teknik Informatika Fakultas Teknik

Universitas Nusantara PGRI Kediri

Mengetahui/Menyetujui

Dosen Pembimbing

Tanggal, 17 Agustus 2015

Pembimbing I	Pembimbing II
Drs.Zaenal Arifin, M.M	Patmi Kasih, M.Kom
NIDN. 0023086809	NIDN. 0701107802

iii

Skripsi Oleh :

MUHAMMAD AMDANIBIK
NPM: 11.1.03.02.0257

Judul :

SISTEM REKOMENDASI PENGAMBILAN KATEGORI SKRIPSI DENGAN PERHITUNGAN NAÏVE BAYES

Telah dipertahankan di depan Panitia Ujian/Sidang Skripsi
Program Studi Teknik Informatika UNP Kediri
Pada tanggal : 28 September 2015

Dan Dinyatakan telah memenuhi persyaratan

Panitia Penguji:

1. Ketua : Dr.Suryo Widodo, M.Pd
2. Penguji 1 : Dr.Atrup, M.Pd., M.M
3. Penguji 2 : Resty Wulaningrum, M.Kom

Mengetahui,
Dekan Fakultas Teknik

Dr.Suryo Widodo, M.Pd
NIP. 19640202 199103 002

iv

Sistem Rekomendasi Pengambilan Kategori Skripsi Dengan Perhitungan Naïve Bayes

Muhammad Amdanibik
11.1.03.02.0257
Fakultas Teknik - Program Studi Teknik Informatika
Email : dnamdan@gmail.com
Drs.Zaenal Arifin, M.M dan Patmi Kasih, M.Kom
UNIVERSITAS NUSANTARA PGRI KEDIRI

ABSTRAK

Skripsi atau Tugas Akhir adalah hal terbesar yang dilalui mahasiswa untuk memperoleh gelar sarjana, dalam hal ini banyak mahasiswa yang ragu dalam pengambilan kategori skripsi atau tugas akhir. Disisi lain nilai mata kuliah dari mahasiswa yang telah didapat belum secara maksimal digunakan sepenuhnya untuk membantu pengambilan kategori skripsi. Oleh karena itu nilai dari mata kuliah bisa digunakan sebagai acuan untuk melakukan pengambilan kategori skripsi. Sistem rekomendasi pengambilan kategori skripsi menggunakan perhitungan naïve bayes dan dibangun dengan bahasa pemrograman php dan database mysql.

Kata Kunci: Sistem Rekomendasi, Skripsi, Naïve Bayes, Nilai Mata Kuliah, Kategori Skripsi.

1. Latar Belakang

Perkuliahan merupakan jenjang pendidikan yang ditempuh di perguruan tinggi, mahasiswa dituntut untuk aktif belajar dan berperan dalam perkuliahan. Diantara kegiatan sosial yang wajib diikuti mahasiswa adalah kegiatan sosial masyrakat seperti KKN dan melakukan praktek langsung sesuai dengan disiplin ilmu yang sesuai dengan jurusan masing-masing yang biasa disebut dengan PKL (Praktek Kerja Lapangan). Selain itu kegiatan perkuliahan yang diikuti juga menentukan hasil akhir yang akan dicapai.

Tugas akhir (Skripsi) merupakan sebuah puncak untuk mencapai kelulusan di tingkat perkuliahan. Pada tugas akhir mahsiswa dituntut untuk melakukan penelitian serta melakukan penerapan ilmu sesuai dengan disiplin ilmu yang diambil pada jurusan yang sudah dipilih. Tugas akhir terkadang juga menjadi sebuah tembok besar yang harus diloncati untuk mencapai kelulusan

Pemilihan kategori pada tugas akhir (skripsi) juga merupakan salah satu hal yang paling rentan dalam tugas akhir. Pemilihan kategori tugas akhir yang tepat sangat berpengaruh dalam menentukan keberhasilan. Selain itu, faktor pendukung seperti nilai dapat membantu kita dalam menentukan pemilihan kategori dan juga tingkat keyakinan pada kategori yang akan dipilih juga berpengaruh pada cepat atau lambatnya dalam menyelesaikan skripsi.

2. Metode

Naïve bayes merupakan salah satu algoritma penerapan dari teorema bayes. Perhitungan pada naïvebayes berdasarkan pada asusmsi penyederhanaan bahwa nilai atribut secara conditional saling bebas jika diberikan nilai output. Probabilitas mengamati secara bersama produk dari probabilitas individu. (Budi Santosa, 2007)

Secara garis besar naïve bayes terdiri dari 4 langkah yaitu :

1) Mencari nilai probabilitas hipotesa (P(H))
2) Mencari nilai probabilitas posterior evidence (P(X|H))
3) Mencari nilai keseluruah dari probabilitas posterior evidence (π(P(X|H)))
4) Mencari nilai probabilitas posterior dari hipotesa (P(H|X))

Kasus dalam penelitian adalah menentukan kategori skripsi (tugas akhir) yang akan direkomendasikan oleh sistem kepada user. Didalam naïve bayes field atau column data akan dibagi menjadi 2 bagian utama yaitu :

1) H = data yang menjadi hipotesa/ attribute/ class, biasa ditulis dengan C

2) X = data evidence/ syarat/ field/ attribute yang akan mempengaruhi data H

Pencarian kategori akan dibagi menjadi 2 bagian besar, pertama data testing dari user termasuk dalam rekomendasi pengambilan kategori SPK atau tidak, jika benar maka proses mining akan berhenti. Jika tidak maka proses mining akan berlanjut. Sesuai dengan dataset yang telah didapat, peneliti membagi class (H) menjadi 7 kategori yaitu AI, Citra, Game, Jaringan, Pakar, SI, dan SPK.

Mata kuliah yang akan diambil untuk menjadi evidence (X) adalah mata kuliah yang mempunyai keterkaitan dengan mata kuliah inti di program studi teknik informatika. Mata kuliah tersebut yaitu Algoritma Pemrograman I, Prak. Algoritma Pemrograman I, Pengantar Teknologi Informasi, Prak. Pengantar Teknologi Informasi, Algoritma Pemrograman II, Prak. Algoritma Pemrograman II, Struktur Data, Prak.Struktur Data, Basis Data, Prak.Basis Data, Dasar Multimedia, Pemrograman Berorientasi Objek, Prak. Pemrograman Berorientasi Objek, Arsitektur Komputer, Sistem Informasi, Jaringan Komputer, Prak. Jaringan Komputer, Pemrograman Web, Prak. Pemrograman Web, Rekayasa Perangkat Lunak, Interaksi Manusia dan Komputer, Kecerdasan buatan, Metodologi Penelitian, Sistem Terdistribusi, Prak. Sistem Terdistribusi, Grafika Komputer, Sistem Berbasis Pengetahuan, Data Mining, Sistem Operasi dan Pengolahan Citra Digital.

Mata kuliah tersebut diatas merupakan mata kuliah inti program studi teknik informatika dari semester 1-6. Pemelihan range semester didasarkan pada bahwa pendaftaran dan pengambilan judul proposal skripsi pada program studi teknik informatika dilakukan pada awal semester 7. Penjelasan nilai mata kuliah yang digunakan dalam data adalah sebagai berikut :

1) Nilai A atau nilai 4 diubah menjadi angka 8
2) Nilai B+ atau nilai 3.5 diubah menjadi angka 7
3) Nilai B atau nilai 3 diubah menjadi angka 6
4) Nilai C+ atau nilai 2.5 diubah menjadi angka 5
5) Nilai C atau nilai 2 diubah menjadi angka 4
6) Nilai D atau nilai 1 diubah menjadi angka 2
7) Nilai E atau nilai 0 diubah menjadi angka 0

Untuk melakukan testing kita perlu memasukkan data testing ke dalam dataset yang berupa nilai mata kuliah dari mahasiswa yang ingin mencari rekomendasi skripsi yang akan diambil. Penambahan ini

dilakukan karena naïve bayes melakukan learning bersamaan dengan masuknya data testing jika tidak ada data testing yang masuk maka naïve bayes tidak akan melakukan learning

Tabel Sampel Data Testing

No	Mata Kuliah	Singkatan	Nilai
1	Algoritma Pemrograman I	Alpro I	8
2	Pengantar Teknologi Informasi	PTI	7
3	Prak. Pengantar Teknologi Informasi	Prak.PTI	7
4	Prakt. Algoritma Pemrograman I	Prak.Alpro I	8
5	Algoritma Pemrograman II	Alpro II	7
6	Sistem Informasi	SI	7
7	Arsitektur dan Organisasi Komputer	Arsikom	7
8	Basis Data	Basis data	8
9	Prak. Basis Data	Prak.Basis Data	8
10	Prak. Algoritma Pemrograman II	Prak.Alpro II	7
11	Sistem Operasi	SO	8
12	Pemrograman Berorientasi Obyek	PBO	8
13	Struktur Data	Struktur Data	8
14	Dasar Multimedia	Dasmul	6
15	Prak. Pemrograman Berorientasi Obyek	Prak.PBO	7
16	Prak. Struktur Data	Prak.Struktur Data	8
17	Rekayasa Perangkat Lunak	RPL	7
18	Jaringan Komputer	Jarkom	8
19	Pemrograman Web	P.Web	8
20	Prakt. Jaringan Komputer	Prak.Jarkom	8

21	Prak. Pemrograman Web	Prak.P.Web	8
22	Grafika Komputer	Grafkom	8
23	Interaksi Manusia dan Komputer	IMK	7
24	Kecerdasan Buatan	AI	7
25	Sistem Terdistribusi	Sister	8
26	Prak. Sistem Terdistribusi	Prak.Sister	7
27	Metodologi Penelitian	Metpen	7
28	Data Mining	Damin	8
29	Sistem Berbasis Pengetahuan	SBP	7
30	Pengolahan Citra	PCD	7

Adapun perhitungan naïve bayes pada data testing diatas adalah

1) Langkah Pertama : Mencari nilai Probabilitas Prior Hipotesa/ P (H)

Langkah ini digunakan untuk mencari nilai probabilitas class terhadap data, class yang dimaksud adalah kategori skripsi yang sudah masuk kedalam data training, terdapat 2 tahap yaitu menentukan data testing masuk dalam kategori SPK atau tidak jika tidak maka akan menentukan data testing termasuk dalam kategori AI, Citra, Game, Jaringan, Pakar, atau SI. Dataset yang didapat berjumlah 174 data. Jumlah data dengan kategori SPK 112 data dan jumlah data NonSPK 62 data.

Nilai P(H) yang didapat adalah

a) P (Kategori = SPK) : 112 / 174 = 0.643678161
b) P (Kategori = Non SPK) : 62 / 174 = 0.356321839

2) Langkah Kedua : Mencari nilai Probabilitas Posterior dari Evidence / P(X|H)

Langkah ini digunakan untuk mencari nilai probabilitas dari data testing terhadap data training untuk masing-masing kategori yang ada.

Nilai P(X|H) yang didapat adalah

a) Alpro I

P(Alpro I = 8 | Kategori = SPK) : 7 / 112 = 0.0625

P(Alpro I = 8 | Kategori = NonSPK) : 5 / 62 = 0.080645161

b) PTI

P(PTI = 7 | Kategori = SPK) : 41 / 112 = 0.366071429

P(PTI = 7 | Kategori = NonSPK) : 33 / 62 = 0.532258065

c) Prak. PTI

P(Prak.PTI = 7 | Kategori = SPK) : 37 / 112 = 0.330357143

P(Prak.PTI = 7 | Kategori = NonSPK) : 31 / 62 = 0.5

d) Prak. Alpro I

P(Prak.Alpro I = 8 | Kategori = SPK) : 7 / 112 = 0.0625

P(Prak.Alpro I = 8 | Kategori = NonSPK) : 5 / 62 = 0.080645161

e) Alpro II

P(Alpro II = 7 | Kategori = SPK) : 35 / 112 = 0.3125

P(Alpro II = 7 | Kategori = NonSPK) : 27 / 62 = 0.435483871

f) SI

P(SI = 7 | Kategori = SPK) : 34 / 112 = 0.303571429

P(SI = 7 | Kategori = NonSPK) : 16 / 62 = 0.258064516

g) Arsikom

P(Arsikom = 7 | Kategori = SPK) : = 60 / 112 = 0.535714286

P(Arsikom = 7 | Kategori = NonSPK) : 35 / 62 = 0.564516129

h) Basis Data

P(Basis Data = 8 | Kategori = SPK) : 2 / 112 = 0.017857143

P(Basis Data = 8 | Kategori = NonSPK) : 1 / 62 = 0.016129032

i) Prak. Basis Data

P(Prak.Basis Data = 8 | Kategori = SPK) : 2 / 112 = 0.017857143

P(Prak.Basis Data = 8 | Kategori = NonSPK) : 3 / 62 = 0.048387097

j) Prak. Alpro II

P(Prak.Alpro II = 7 | Kategori = SPK) : 28 / 112 = 0.25

P(Prak.Alpro II = 7 | Kategori = NonSPK) : = 25 / 62 = 0.403225806

k) SO

P(SO = 8 | Kategori = SPK) : 15 / 112 = 0.133928571

P(SO = 8 | Kategori = NonSPK) : 11 / 62 = 0.177419355

l) PBO

P(PBO = 8 | Kategori = SPK) : 7 / 112 = 0.0625

P(PBO = 8 | Kategori = NonSPK) : 2 / 62 = 0.032258065

m) Struktur Data

P(Struktur Data = 8 | Kategori = SPK) : 4 / 112 = 0.35714286

P(Struktur Data = 8 | Kategori = NonSPK) : 4 / 62 = 0.064516129

n) Dasmul

P(Dasmul = 6 | Kategori = SPK) : 48 / 112 = 0.428571429

P(Dasmul = 6 | Kategori = NonSPK) : 9 / 62 = 0.14516129

o) Prak. PBO

P(Prak.PBO = 7 | Kategori = SPK) : 15 / 112 = 0.133928571

P(Prak.PBO = 7 | Kategori = NonSPK) : 19 / 62 = 0.306451613

p) Prak. Struktur Data

P(Prak.Struktur Data = 8 | Kategori = SPK) : 3 / 112 = 0.026785714

P(Prak.Str.Data = 8 | Kategori = NonSPK) : 4 / 62 = 0.064516129

q) RPL

P(RPL = 7 | Kategori = SPK) : 19 / 112 = 0.169642857

P(RPL = 7 | Kategori = NonSPK) : 13 / 62 = 0.209677419

r) Jarkom

P(Jarkom = 8 | Kategori = SPK) : 12 / 112 = 0.107142857

P(Jarkom = 8 | Kategori = NonSPK) : 8 / 62 = 0.129032258

s) P.Web

P(P.Web = 8 | Kategori = SPK) : 5 / 112 = 0.044642857

P(P.Web = 8 | Kategori = NonSPK) : 7 / 62 = 0.112903226

t) Prak. Jarkom

P(Prak.Jarkom = 8 | Kategori = SPK) : 26 / 112 = 0.23214857

P(Prak.Jarkom = 8 | Kategori = NonSPK) : 11 / 62 = 0.177419355

u) Prak. P.Web

P(Prak.P.Web = 8 | Kategori = SPK) : 6 / 112 = 0.053571429

P(Prak.P.web = 8 | Kategori = NonSPK) : 7 / 62 = 0.1129032266

v) Grafkom

P(Grafkom = 8 | Kategori = SPK) : 3 / 112 = 0.026785714

P(Grafkom = 8 | Kategori = NonSPK) : 3 / 62 = 0.048387097

w) IMK

P(IMK = 7 | Kategori = SPK) : 30 / 112 = 0.2678157143

P(IMK = 7 | Kategori = NonSPK) : 24 / 62 = 0.387096774

x) AI

P(AI = 7 | Kategori = SPK) : 32 / 112 = 0.285714286

P(AI = 7 | Kategori = NonSPK) : 33 / 62 = 0.532258065

y) Sister

P(Sister = 8 | Kategori = SPK) : 5 / 112 = 0.044642857

P(Sister = 8 | Kategori = NonSPK) : 10 / 62 = 0.161290323

z) Prak. Sister

P(Prak.Sister = 7 | Kategori = SPK) : 46 / 112 = 0.410714286

P(Prak.Sister = 7 | Kategori = NonSPK) : 23 / 62 = 0.370967742

aa) Metpen

P(Metpen = 7 | Kategori = SPK) : 31 / 112 = 0.276785714

P(Metpen = 7 | Kategori = NonSPK) : 15 / 62 = 0.241935484

bb) Damin

P(Damin = 8 | Kategori = SPK) : 8 / 112 = 0.071428571

P(Damin = 8 | Kategori = NonSPK) : 5 / 62 = 0.080645161

cc) SBP

P(SBP = 7 | Kategori = SPK) : 35 / 112 = 0.3125

P(SBP = 7 | Kategori = NonSPK) : 31 / 62 = 0.5

dd) PCD

P(PCD = 7 | Kategori = SPK) : 31 / 112 = 0.276785714

P(PCD = 7 | Kategori = NonSPK) : 26 / 62 = 0.419354839

3) Langkah ketiga : mencari nilai Perkalian Probabilitas Posterior Evidence / πP(X|H)

Nilai probabilitas ini digunakan untuk mendapatkan nilai perkalian dari masing-masing nilai probabilitas kategori, nilai πP(X|H) yang didapat adalah

πP(X | Kategori = SPK) = 0.0625 * 0.366071429 * 0.330357143 * 0.0625 * 0.3125 * 0.303571429 * 0.535714286 * 0.017857143 * 0.017857143 * 0.25 * 0.133928571 * 0.0625 * 0.035714286 * 0.4288571429 * 0.133928571 * 0.026785714 * 0.169642857 * 0.107142857 * 0.044642857 * 0.232142857 * 0.053571429 * 0. 026785714 * 0.267857143 * 0.285714286 * 0.044642857 * 0.410714286 * 0.276785714 * 0.071428571 * 0.3125 * 0.27685714 = 5.70547E-28

πP(X | Kategori = NonSPK) = 0.080645161 * 0.532258065 * 0.5 * 0.080645161 * 0.435483871 * 0.258064156 * 0.564516129 * 0.016129032 * 0.048387097 * 0.177419355 * 0.032258065 * 0.064516129 * 0.14516129 * 0.306451613 * 0.064516129 * 0.209677419 * 0.129032258 * 0.112903226 * 0.177419355 * 0.112903226 * 0.048387097 * 0.387096774 * 0.532258065 * 0.161290323 * 0.370967742 * 0.241935484 * 0.080645161 * 0.5 * 0.419354839 = 5.46776E-24

4) Langkah keempat : mencari nilai probabilitas posterior hipotesa / P(H|X)

Hasil pada langkah keempat akan digunakan untuk menentukan kategori data testing masuk dalam kategori SPK atau NonSPK, nilai P(H|X) yang didapat adalah

a) P(Kategori SPK | X) = πP(X|Kategori = SPK) * P(Kategori = SPK) : 5.70547E-28 * 0.643678161 = 3.67249E-28

b) P(Kategori NonSPK | X) = πP(X|Kategori = NonSPK) * P(Kategori = NonSPK) : 5.46776E-24 * 0.356321839 = 1.94828E-24

Dari hasil perhitungan diatas, didapat hasil maksimal adalah 1.94828E-24 maka

data testing yang dimasukkan disarankan untuk mengambil kategori skripsi **NonSPK**.

Karena pada proses yang pertama menghasilkan kategori NonSPK maka proses akan dilakukan kembali dengan data nilai yang sama dari mahasiswa tersebut dengan proses perhitungan yang sama untuk kategori skripsi NonSPK yaitu AI, Citra, Game, Jaringan, Pakar, atau SI hingga diperoleh satu kategori rekomendasi paling dominan untuk mahasiswa.

3. Hasil dan Kesimpulan

Setelah melakukan beberapa uji coba terhadap sistem yang telah dibuat maka diperoleh beberapa kesimpulan sebagai berikut :

1. Berdasarkan hasil penelitian didapatkan bahwa nilai mata kuliah dari mahasiswa dapat digunakan sebagai acuan untuk melakukan pengambilan kategori skripsi sehingga mahasiswa tidak bingung lagi dalam mencari rekomendasi yang cocok untuk pengambilan kategori skripsi.
2. Nilai mata kuliah dari mahasiswa juga dapat digunakan secara optimal dan tidak hanya mendapat nilai saja tetapi juga dapat digunakan untuk memberikan rekomendasi untuk pengambilan kategori skripsi dari mahasiswa tersebut nanti.
3. Perhitungan naïve bayes dapat diimplentasikan kedalam sistem untuk memberikan kategori skripsi pada mahasiswa dengan hasil yang memuaskan. Dengan perhitungan ini mahasiswa mendapatkan hasil rekomendasi terbaik dari nilai mata kuliah yang dimasukkan kedalam sistem.

4. Daftar Pustaka

Clarentia, Claudia Ciptohartono. 2014. *Algoritma Klasifikasi Naïve Bayes untuk Menilai Kelayakan Kredit*. Semarang : Skripsi. Skripsi Program Studi Teknik Informatika Universitas Dian Nuswantoro.

Hamzah, Amir. 2012. *Klasifikasi Teks Dengan Naïve Bayes Classifier(NBC) untuk Pengelompokan Teks Berita dan Abstract Akademis*. Yogyakarta : Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) Periode III.

http://www.id.wikipedia.org/wiki/Universitas_Nusantara_PGRI_Kediri/diakses pada 15 Agustus 2015

http://www.twoh.co/2013/05/sekilas-tentang-sistem-rekomendasi-recommender-system/ diakses pada 10 December 2014.

http://www.uajy.ac.id/mahasiswa/administrasi-akademik/perkuliahan/kegiatan-perkuliahan/ diakses pada 18 mei 2015

Indranandita, Amalia. 2008. *Sistem Klasifikasi dan Pencarian Jurnal Dengan Menggunakan Metode Naïve Bayes dan Vector* . Jurnal Informatika Volume 4 Nomor 2.

Kusrini dan Emha Taufiq Luthfi. 2009. *Algoritma Data Mining.* Yogyakarta : Andi.

Nurani, Atri dkk. 2007. *Implementasi Naïve Bayes Classifier pada Progam Bantu Penentuan Buku Referensi Mata Kuliah*. Jurnal Informatika Volume 3 Nomor 2.

Santosa, Budi. 2007. *Data Mining Teknik dan Pemanfaatan Data untuk Keperluan Bisinis teori dan aplikasi*. Yogyakarta : Graha Ilmu.

Sidik, Betha. 2012. *Pemrograman Web PHP Edisi Revis*i. Bandung : Informatika.

Susanto, Sani dan Dedy suryadi. 2010. *Pengantar Data Mining menggali pengetahuan dari bongkahan data*. Yogyakarta : Andi.