

# RANCANGAN SISTEM REKOMENDASI KATEGORI PERUSAHAAN PRAKTIK INDUSTRI MENGGUNAKAN ALGORITMA *NAÏVE BAYES*

Robi Wahyudi<sup>1)</sup>, Aji Prasetya Wibawa<sup>2)</sup>, Utomo Pujiyanto<sup>3)</sup>

<sup>1), 2), 3)</sup>Teknik Elektro Universitas Negeri Malang

Jl Semarang No. 5 Malang 65145

Email : [robii.wahyudi@gmail.com](mailto:robii.wahyudi@gmail.com)<sup>1)</sup>, [aji.prasetya.ft@um.ac.id](mailto:aji.prasetya.ft@um.ac.id)<sup>2)</sup>, [utomo.pujiyanto@yahoo.co.id](mailto:utomo.pujiyanto@yahoo.co.id)<sup>3)</sup>

## Abstrak

*Mengembangkan kompetensi mahasiswa melalui praktik dapat dilakukan dimanapun, salah satunya di industri. Melalui langkah ini diharapkan kompetensi yang dimiliki mahasiswa sesuai dengan harapan industri. Universitas sebagai mitra industri menjadikan praktik ini matakuliah wajib, yaitu Praktik Industri (PI). Dari hasil observasi yang dilakukan kepada mahasiswa Pendidikan Teknik Informatika Universitas Negeri Malang yang telah menempuh PI menyimpulkan adanya kebingungan pada saat pemilihan lokasi PI. Hal ini dikarenakan banyaknya pilihan industri yang tersedia serta tidak ada informasi kompetensi yang dibutuhkan perusahaan. Akibatnya, mahasiswa yang tidak sesuai dengan kompetensi yang dibutuhkan perusahaan mengalami penolakan pada proses permohonan PI ke perusahaan. Penolakan permohonan PI ke perusahaan yang berulang dapat menghambat waktu pelaksanaan PI. Kategori perusahaan menentukan kompetensi yang dibutuhkan pada saat PI. Pada artikel ini akan dibahas penerapan algoritma Naïve Bayes untuk menentukan kategori lokasi PI yang paling sesuai dengan kompetensi mahasiswa sehingga dapat mengurangi jumlah penolakan dari perusahaan. Pada tahap selanjutnya akan dilanjutkan ke tahapan pengembangan Sistem Pendukung Keputusan (SPK).*

**Kata kunci:** *Praktik Industri(PI), Naïve Bayes, Sistem Pendukung Keputusan (SPK)*

## 1. Pendahuluan

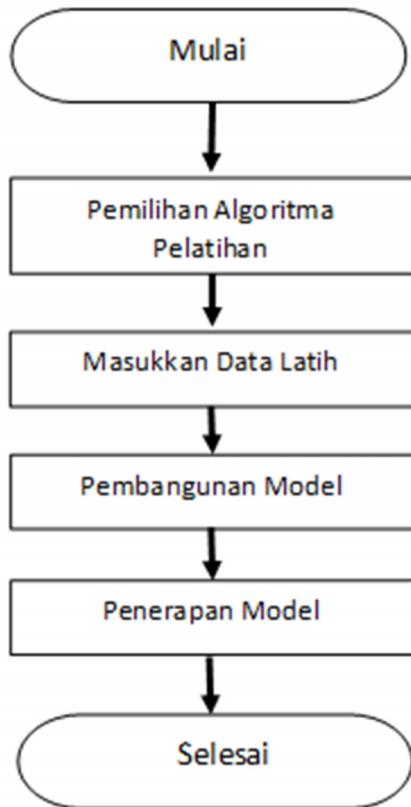
Universitas sebagai lembaga pendidikan tinggi berperan meningkatkan kompetensi lulusan yang sesuai dengan kebutuhan industri salah satunya melalui Praktik Industri. Praktik Industri (PI) merupakan matakuliah wajib bagi mahasiswa untuk menerapkan teori secara langsung di industri. Prosedur PI diawali dengan pengajuan surat permohonan kepada perusahaan tujuan lokasi PI. Pada proses ini, mahasiswa memiliki kemungkinan diterima atau ditolak oleh perusahaan. Apabila mahasiswa mendapat keputusan ditolak oleh perusahaan dapat mengajukan permohonan kembali. Proses ini dapat menjadi hambatan pelaksanaan PI apabila mahasiswa ditolak perusahaan beberapa kali.

Agar hambatan tersebut dapat diminimalkan, dibutuhkan sebuah mekanisme yang dapat membantu mahasiswa menyesuaikan kompetensi dengan perusahaan. Sistem rekomendasi kategori perusahaan PI merupakan Sistem Pendukung Keputusan (SPK) yang dirancang untuk membantu menyelesaikan permasalahan pemilihan kategori perusahaan PI. Pemilihan kategori perusahaan PI adalah permasalahan yang berkaitan dengan klasifikasi. *Naïve Bayes* dapat digunakan untuk menyelesaikan permasalahan klasifikasi dengan probabilitas sederhana. Sehingga dengan penerapan algoritma tersebut, dapat memberikan rekomendasi kategori lokasi PI yang sesuai.

## 2. Pembahasan

Sistem Pendukung Keputusan (SPK) adalah sistem informasi interaktif berbasis komputer yang digunakan untuk membantu pengambilan keputusan dalam situasi yang semiterstruktur dan situasi yang tidak terstruktur dengan berbagai metode yang dapat digunakan [3]. Pemilihan metode didasarkan pada faktor (karakteristik) yang ditemukan pada permasalahan. Berdasarkan karakteristik permasalahan pemilihan kategori lokasi PI *Naïve Bayes* salah satu metode yang dapat digunakan. Metode ini merupakan salah satu metode sederhana untuk klasifikasi yang berdasar pada konsep teorema Bayes [1]. Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya kedalam kelas tertentu dari sejumlah kelas yang tersedia [2]. Kelas hasil klasifikasi dalam sistem ini merupakan kategori perusahaan. Adapun kategori tersebut ada tiga yakni : (a) Instansi Pemerintah; (b) Perusahaan Swasta; (c) *Softwarehouse*.

Secara garis besar mekanisme kerja dari klasifikasi terdapat empat langkah utama, meliputi : (a) pemilihan algoritma pelatihan; (b) memasukkan data latih; (c) pembangunan model; (d) penerapan model. Pembahasan secara detail mekanisme klasifikasi dapat dilihat pada gambar 1 serta penjelasan berikut ini.



Gambar 1 Diagram Alir Klasifikasi [1]

a. Pemilihan Algoritma Pelatihan

Naive Bayes dipilih sebagai metode pelatihan dalam sistem ini. rumus Naive Bayes untuk klasifikasi (1) dapat dituliskan sebagai berikut :

$$P(Y | X) = \frac{P(Y) \prod_{i=1}^n P(X_i|Y)}{P(X)} \dots \dots (1)$$

Keterangan :

- P(Y|X) : Probabilitas data dengan vector X pada kelas Y
- P(Y) : Probabilitas awal kelas Y.
- $\prod_{i=1}^n P(X_i|Y)$  : Probabilitas independen kelas Y dari semua atribut dalam vector X.
- P(X) : Nilai P(X) selalu tetap ,hasil tertinggi digunakan untuk pemilihan kelas.

Jika X adalah vector masukkan yang berisi atribut dan Y adalah label kelas, notasi penulisan Naive Bayes adalah P(Y|X). Penggunaan notasi P(Y|X)

menunjukkan probabilitas label kelas Y yang didapatkan setelah atribut-attribut X diamati. Pada sistem ini label kelas Y adalah kategori perusahaan, sedangkan atribut-attribut X adalah faktor yang mempengaruhi pemilihan lokasi PI. Proses untuk memodelkan data dari atribut kombinasi X dan Y didapatkan dari data latih. Data latih merupakan data pengalaman yang digunakan untuk membangun model klasifikasi. Data ini didapatkan dari pengalaman mahasiswa yang telah menempuh PI. Dalam Naive Bayes, setiap atribut dan kelas dihitung nilai probabilitasnya. Hal ini dimaksudkan untuk mendapatkan nilai probabilitas akhir pada setiap kelas menggunakan data uji. Nilai probabilitas akhir P(Y|X) pada setiap kelas yang paling tinggi merupakan kelas yang dijadikan hasil prediksi

b. Menyiapkan Data Latih

Model probabilitas didapatkan dari analisis data latih (training set), Sedangkan data uji (test set) digunakan untuk mengetahui tingkat akurasi dari model probabilitas yang telah dihasilkan [4]. Sebelum pembangunan data latih harus ditentukan nilai atribut yang akan digunakan. Atribut-attribut dan kelas yang digunakan pada sistem ini dijabarkan sebagai berikut:

1. Atribut Index Prestasi Kumulatif (IPK)

Index Prestasi Kumulatif didapatkan dari perhitungan persamaan (2).

$$IPK = \frac{(\sum N_{ki} * Ni)}{(\sum N_{ki})} \dots (2)$$

Keterangan :

$\sum N_{ki}$  : Jumlah kredit matakuliah (SKS)

Ni : Nilai Matakuliah

IPK merupakan atribut bertipe numerik sehingga diperlukan langkah deskretisasi dengan mengganti nilai atribut kontinu tersebut dengan nilai interval diskret [1]. Atribut IPK dapat dinyatakan dengan huruf A,A-,B+,B-,C,D,E yang ekivalen dengan angka 4-3.71; 3.70-3.52; 3.30-3.00; 2.99-2.7; 2.69-2.30; 2.29-2.0; 1.99-1.0; 0.99-0 [2]. Adapun contoh deskretisasi nilai kontinu IPK dengan nilai huruf ditunjukkan pada tabel 1.

Tabel 1 Tabel Contoh Dekretisasi Nilai IPK

Mahasiswa	IPK (Angka)	IPK (Huruf)
Mhs Satu	3.43	A-
Mhs Dua	3.75	A
Mhs Tiga	2.70	B

## 2. Atribut Jarak Lokasi PI (J)

Nilai untuk atribut jarak ada dua yaitu kategori jarak dekat dan jarak jauh. Atribut jarak didasarkan pada tempat tinggal mahasiswa dengan lokasi PI. Nilai interval untuk jarak dekat antara 0 sampai 10 Km, sedangkan untuk nilai interval jarak jauh lebih dari 10 Km.

## 3. Atribut Waktu Kerja (WK)

Berdasarkan waktu kerja normal 8-10 jam nilai atribut waktu kerja dibagi menjadi tiga kategori, yaitu: (a) kurang dari 8 jam; (b) 8-10 jam; (c) lebih dari 10 jam. Waktu kerja mahasiswa PI disesuaikan dengan kebijakan dari perusahaan.

## 4. Atribut Beban Penugasan (BP)

Atribut ini berdasarkan penilaian mahasiswa terhadap beban penugasan yang diberikan pada saat PI. Penentuan nilai pada atribut beban penugasan meliputi: (a) ringan; (b) sedang; (c) berat. Penugasan selama PI diberikan oleh pembimbing yang ditunjuk oleh industri sehingga dapat disesuaikan dengan kompetensi mahasiswa.

## 5. Atribut Tujuan Bekerja (TB)

Atribut tujuan bekerja dibagi dalam skala 1 sampai 5 berdasarkan skala *likert*. Skala tersebut menunjukkan pendapat persetujuan tentang lokasi PI sebagai tujuan bekerja. Berikut adalah skala untuk atribut tujuan bekerja: (a) 5= sangat setuju (SS); (b) 4= setuju (S); (c) 3= netral (N); (d) 2= tidak setuju (TS); (e) 1= sangat tidak setuju (STS).

## 6. Kelas

Pada sistem ini kelas yang akan digunakan sebagai klasifikasi merupakan kategori perusahaan. Kategori perusahaan tersebut didapatkan dari hasil observasi mahasiswa yang telah menempuh PI. Dari hasil observasi tersebut didapatkan tiga kategori perusahaan yaitu : (a) Instansi Pemerintah (IP); (b) Perusahaan Swasta (PS); (c) *Softwarehouse* (SH).

### c. Pembangunan Model

Tahapan ini bertujuan untuk mendapatkan model dari data latih. Dari data ini kemudian dihitung nilai probabilitas setiap nilai pada setiap atribut dan kelasnya. Probabilitas tersebut dapat dinotasikan dengan  $P(X_i|Y_i)$ . Setiap jumlah dari nilai atribut ditambah 1 untuk menghindari hasil probabilitas pembagian nol [4]. Sebagai contoh pembangunan model pada label Kelas ditunjukkan pada tabel 2.

Tabel 2 Contoh Pembangunan Model Label Kelas

Kelas	Jumlah
IP	10
PS	5
SH	5

Dari tabel 2 dapat diamati probabilitas  $P(X)$  untuk setiap kelas IP,PS,SH masing-masing 0.5,0.25,0.25. Nilai  $P(X)$  didapatkan dari pembagian jumlah dalam kelas tersebut dibagi dengan jumlah kelas seluruhnya. Langkah perhitungan ini disesuaikan dengan formula (1).

### d. Penerapan Model

Untuk mendapatkan akurasi metode harus dilakukan penerapan model probabilitas yang telah dibuat. Tahapan ini bertujuan untuk mendapatkan kelas hasil prediksi. Kelas hasil prediksi akan dibandingkan dengan kelas sebenarnya. Dimana hasil prediksi tersebut merupakan perbandingan nilai probabilitas kelas data uji yang memiliki nilai paling tinggi. Setelah diketahui kelas hasil prediksi selanjutnya dilakukan *filtering* penampilan data nama perusahaan PI berdasarkan kategori tersebut. Langkah ini merupakan rekomendasi sistem yang selanjutnya dijadikan pembantu keputusan pemilihan lokasi PI mahasiswa.

### Pengujian

Adapun pengujian yang akan digunakan adalah *precision, recall* dan akurasi [5]. Data sampel yang digunakan didapatkan dari mahasiswa angkatan 2011-2013 yang telah menempuh PI. Jumlah data yang digunakan sebanyak 50 data dengan kelas Instansi Pemerintah (IP),Perusahaan Swasta (PS) dan *Softwarehouse* (SH) masing-masing sejumlah 17,16 dan 17.

Pada tahap ini data akan dibagi menjadi dua, yaitu data latih dan data uji. Dimana data latih digunakan untuk membentuk model probabilitas pada metode *Naïve Bayes* sedangkan data uji digunakan untuk menguji tabel probabilitas tersebut. Percobaan dilakukan sebanyak tiga kali dengan masing-masing data latih sejumlah 20,30 dan 40 sedangkan data uji sejumlah 10 data. Adapun hasil pengujian akan dijabarkan berikut ini.

#### 1. Percobaan ke-1

Pada percobaan ini digunakan data latih sejumlah 20 data sedangkan untuk data uji digunakan sejumlah 10 set data. Data ini diambil dari sampel mahasiswa angkatan 2011-2013 yang telah menempuh PI. Adapun hasil percobaan ke-1 didapatkan matrik konfusi seperti pada tabel 3.

**Tabel 3** Tabel Matrik Konfusi Percobaan ke-1

$F_{gh}$		Predicted Class		
		IP	PS	SH
Actual Class	IP	0	1	2
	PS	1	1	0
	SH	1	0	4

Hasil pengujian menggunakan 20 data latih dan 10 data uji menunjukkan nilai *precision*, *recall* dan akurasi masing-masing 43.3%, 36.6% dan 50%.

2. Percobaan ke-2

Pada percobaan ke-2 data yang digunakan sama yakni dari sampel mahasiswa angkatan 2011-2013 yang telah menempuh PI. Data latih yang digunakan sejumlah 30 data sedangkan untuk data uji sejumlah 10 set data. Adapun matrik konfusi yang dihasilkan pada percobaan ke-2 ditunjukkan pada tabel 4.

**Tabel 4** Tabel Matrik Konfusi Percobaan ke-2

$F_{gh}$		Predicted Class		
		IP	PS	SH
Actual Class	IP	1	1	1
	PS	1	1	0
	SH	1	0	4

Hasil pengujian menggunakan 30 data latih dan 10 data uji menunjukkan nilai *precision*, *recall* dan akurasi masing-masing 53.3%, 53.3% dan 60%.

3. Percobaan ke-3

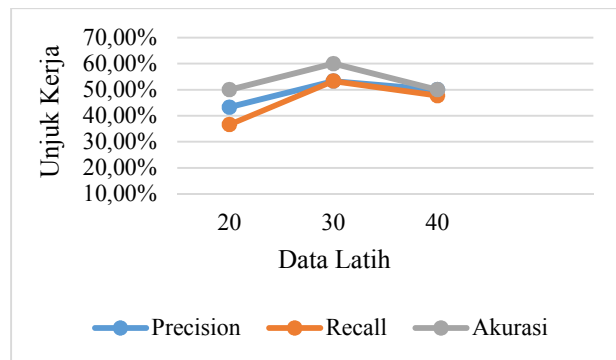
Menggunakan data yang sama, data latih yang digunakan pada percobaan ini sejumlah 40 data dan data uji yang digunakan sejumlah 10 set data. Hasil dari data percobaan ke-3 tersebut ditunjukkan pada tabel 5:

**Tabel 5** Tabel Matrik Konfusi Percobaan ke-3

$F_{gh}$		Predicted Class		
		IP	PS	SH
Actual Class	IP	1	1	1
	PS	1	1	0
	SH	2	0	3

Hasil pengujian menggunakan 40 data latih dan 10 data uji menunjukkan nilai *precision*, *recall* dan akurasi masing-masing 50%, 47.7% dan 50%.

Dari hasil percobaan-percobaan tersebut didapatkan grafik perbandingan dari nilai *precision*, *recall* dan akurasi. Adapun grafik perbandingan tersebut ditunjukkan pada gambar 2.



**Gambar 2.** Grafik Hasil Pengujian

Secara umum hasil perbandingan yang ditunjukkan pada gambar 2 tidak terlalu signifikan. Hal ini dikarenakan data yang digunakan terlalu sedikit jumlahnya. Apabila mahasiswa bertambah pada setiap angkatan maka data yang didapatkan akan semakin banyak sehingga rekomendasi dari SPK yang akan dikembangkan semakin tepat.

3. Kesimpulan

Berdasarkan hasil dan pembahasan, dapat disimpulkan bahwa algoritma *Naïve Bayes* dapat digunakan sebagai metode sistem rekomendasi kategori perusahaan praktik industri bagi mahasiswa. Selanjutnya, metode ini akan diimplementasikan pada SPK menggunakan data latih yang lebih banyak sehingga unjuk kerja sistem dapat meningkat signifikan.

Daftar Pustaka

- [1] Prasetyo, Eko. 2012. *Data Mining Konsep dan Aplikasi menggunakan Matlab*. Yogyakarta: Andi.
- [2] Teknik Elektro Universitas Negeri Malang. 2014. *Katalog Praktik Industri (PI) TE-UM*. Malang: Universitas Negeri Malang.
- [3] Kusri. 2007. *Konsep dan Aplikasi Sistem Pendukung Keputusan*. Yogyakarta: Andi.
- [4] Kusri and E. T. Luthfi. 2009. *Algoritma Data Mining* 1st ed. Yogyakarta, Indonesia: Andi.
- [5] Marina Sokolova and Guy Lapalme, "A systematic analysis of performance measures for classification tasks," *In Information Processing and Management*, vol. 45, pp. 427-437, March 2009.

Biodata Penulis

**Robi Wahyudi**, mahasiswa tingkat akhir Jurusan Teknik Elektro Universitas Negeri Malang.

**Aji Prasetya Wibawa, S.T, M.M.T, Ph.D**, memperoleh gelar Sarjana Teknik (S.T) dari Jurusan Teknik Elektro Universitas Brawijaya, lulus tahun 2004. Memperoleh gelar Magister Manajemen Teknologi (M.MT) dari Institut Teknologi Sepuluh Nopember, lulus tahun 2007. Memperoleh gelar Doctor of Philosophy (PhD) dari School of Engineering, University of South Australia, pada tahun 2014. Saat ini menjadi Dosen di Universitas Negeri Malang sekaligus anggota dari Institute of Electrical and Electronics Engineers (IEEE).

*Utomo Pujianto, S.Kom, M.Kom* , memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Matematika Universitas Gadjah Mada, lulus tahun 2003. Memperoleh gelar Magister Komputer (M.Kom) Program Pasca Sarjana Magister Komputer Informasi Institut Teknologi Sepuluh Nopember, lulus tahun 2010. Saat ini menjadi Dosen di Universitas Negeri Malang.

