

Implementasi Metode Bayes untuk Menentukan Potensi Diri Beserta Pengaruhnya Terhadap IPK Mahasiswa

Bayes Method to Determine Talents and its Impact on Grade Point Average

Agung Jasuma*¹, Kusri², M. Rudyanto Arief³

^{1,2,3} Magister Teknik Informatika, Universitas Amikom Yogyakarta

E-mail: *¹agung.jasuma@students.amikom.ac.id, ²kusri@amikom.ac.id,

³rudy@amikom.ac.id

Abstrak

Tingginya minat siswa untuk melanjutkan sekolah ke jenjang yang lebih tinggi berpengaruh pada tingginya jumlah mahasiswa yang terdaftar di Indonesia. Kesulitan dalam menyelesaikan perkuliahan menjadi masalah yang sering terjadi, salah satu penyebabnya adalah ketidak-sesuaian potensi diri dengan program studi yang dipilih. Penelitian ini menjadikan FIK Universitas Amikom Yogyakarta sebagai tempat studi kasus untuk mencari tahu korelasi potensi diri dengan IPK mahasiswa, mencari potensi diri terbaik pada masing-masing program studi serta mencari tahu akurasi metode naive bayes dalam mengklasifikasi potensi diri mahasiswa. Responden yang digunakan dalam penelitian ini berjumlah 50 orang yang terdiri dari mahasiswa minimal semester 4, alumni dan 1 pakar. Pengumpulan data menggunakan metode wawancara dan koesioner, sedangkan pengolahan data menggunakan metode naive bayes classifier, confusion matrix untuk pengujian, dan korelasi pearson product moment untuk mencari tahu ada tidaknya korelasi. Penelitian ini mendapatkan hasil bahwa potensi diri kemampuan logika, visual dan interpersonal berpengaruh terhadap tingginya IPK mahasiswa dimana nilai signifikansi logika=0.043<0.05, interpersonal=0,029<0.05 dan visual=0,05<0.05. Kemampuan logika cenderung akan berdampak baik pada IPK mahasiswa prodi S1-SI, S1-IF, serta S1-TK, sedangkan kemampuan visual berdampak baik pada program studi S1-TI, D3-MI dan D3-IF. Naive bayes juga diketahui memiliki tingkat akurasi sebesar 90,625% dalam mengklasifikasikan mahasiswa berdasarkan potensi diri.

Kata Kunci — bayes, IPK, mahasiswa

Abstract

The high interest of students to continue their education has an effect on the high number of students in Indonesia. Difficulties in completing lectures become a problem that often occurs, one of the reasons is incompatibility study program and student talents. This research made FIK Amikom University Yogyakarta as a case study to find out the correlation of Grade Poinr average (GPA) and student talents, best talents that's needed in each study program and the accuracy of naive bayes in classifying students' talents, 50 people consisting of students at least semester 4, alumni and 1 expert as respondents. Data collection uses interview and questionnaire methods, while data processing uses the naive bayes classifier, confusion matrix for testing, and Pearson product moment correlation to find out whether there is correlation. This study found that the self-potential logic, visual and interpersonal abilities influence the high GPA of students where the significance value of logic=0.043<0.05, interpersonal=0.029<0.05 and visual=0.05<0.05. Logical ability tends to have a good impact on the GPA of S1-SI, S1-IF, and S1-TK study program, while visual abilities have an impact on S1-TI, D3-MI and D3-IF. Naive Bayes is also known to have an accuracy rate of 90.625% in classifying students based on their talents.

Keywords — bayes, GPA, students

1. PENDAHULUAN

Pesatnya perkembangan pendidikan di Indonesia berpengaruh pada tingginya minat siswa SMA dan sederajat untuk melanjutkan sekolah ke jenjang yang lebih tinggi. Hal ini dibuktikan dengan data statistik yang dirilis oleh PDDikti Kemenristekdikti dalam statistik pendidikan tinggi tahun 2017 dimana jumlah mahasiswa Indonesia yang terdaftar sebesar 6.924.511 atau 21% lebih banyak dibanding penduduk Singapura yang berjumlah 5.747.886 jiwa. Yogyakarta yang merupakan kota pendidikan memiliki jumlah mahasiswa sebesar 373.478 dengan jumlah putus sekolah atau Drop Outs 12.425 orang atau 3,3% dari jumlah mahasiswa yang ada di provinsi Yogyakarta. Angka itu menjadikan provinsi Yogyakarta berada pada posisi ke 8 berdasarkan presentasi putus sekolah dari semua provinsi yang ada di Indonesia, dimana Bengkulu menjadi provinsi yang paling tinggi dalam jumlah putus sekolah yakni 8.2% dari 112.537 mahasiswa [1]. Banyak faktor yang menyebabkan mahasiswa kesulitan dalam menyelesaikan masa studi, salah satunya adalah ketidaksesuaian bidang ilmu yang dipelajari mahasiswa dan minat bakat mahasiswa tersebut [2][3][4]. Faktor tingkat stress juga menjadi hal yang dapat mempengaruhi mahasiswa dalam kegiatan belajar [5]. Hal tersebut yang menjadi dasar peneliti melakukan penelitian ini, dimana untuk membuktikan hal tersebut peneliti menjadikan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta sebagai tempat studi kasus.

Universitas Amikom Yogyakarta merupakan salah satu perguruan tinggi swasta di Yogyakarta yang unggul dalam cabang ilmu komputer. Fakultas Ilmu Komputer Universitas Amikom Yogyakarta sendiri terdiri dari beberapa program studi, yakni S1 Sistem Informasi, D3 Manajemen Informatika, S1 Teknologi Komputer, S1 Informatika, D3 Teknik Informatika, S1 Teknologi Informasi.

Tujuan penelitian ini adalah untuk mencari tahu ada tidaknya hubungan antara potensi diri atau bakat seseorang dengan IPK yang didapat selama perkuliahan, mencari tahu potensi diri paling baik untuk masing-masing program studi yang ada di Fakultas Ilmu Komputer Universitas Amikom Yogyakarta, serta mencari tahu akurasi metode naive bayes dalam mengklasifikasikan potensi diri mahasiswa.

Naive bayes digunakan untuk menghitung probabilitas terjadinya suatu peristiwa berdasarkan pengaruh yang didapat dari pengujian. Probabilitas Bayes menerangkan hubungan antara probabilitas terjadinya hipotesis H_1 jika terdapat fakta (evidence) E telah terjadi dan probabilitas terjadinya evidence E dengan syarat hipotesis H_1 telah terjadi [6]. Untuk mengetahui ada tidaknya korelasi atau hubungan antara potensi diri dengan IPK, metode yang digunakan adalah analisis korelasi Pearson product moment, yang merupakan salah satu teknik dalam mencari korelasi antara variabel yang berbentuk interval [7].

Penelitian yang mengangkat masalah terkait ketidaksesuaian antara minat bakat dengan jurusan maupun program studi yang dipilih sudah ada sebelumnya, seperti penelitian yang dilakukan oleh Liliana pada tahun 2013, dalam penelitiannya Swastina mengatakan bahwa pemilihan jurusan yang tidak sesuai dengan kemampuan, kepribadian, minat dan bakat dapat mempengaruhi mahasiswa dalam mengikuti perkuliahan. Pada penelitiannya ini Liliana melakukan eksperimen dan evaluasi menggunakan algoritma decision tree C4.5 untuk mengatasi permasalahan tersebut, dengan kesimpulan bahwa algoritma C4.5 dapat diterapkan untuk menentukan jurusan mahasiswa dengan tingkat akurasi sebesar 93,31% dan akurasi rekomendasi jurusan sebesar 82,64% [2].

Febryanti pada 2016 menerapkan metode Analytic Hierarchy Process (AHP) untuk menyelesaikan masalah peminatan bagi mahasiswa program studi sistem informasi Universitas Telkom. Kesimpulan dari penelitian ini mengatakan bahwa dengan menggunakan metode AHP diketahui bobot kriteria nilai matakuliah memiliki nilai sebesar 65% dan bobot potensi kecenderungan sebesar 35% dengan konsistensi rasio sebesar 10% [3].

Penelitian berikutnya yang dijadikan rujukan adalah penelitian yang dilakukan oleh Rozi dan Purnomo pada tahun 2017 yang membuat sistem pendukung keputusan untuk menentukan rekomendasi pemilihan minat studi menggunakan metode Mamdani. Data yang digunakan pada penelitian ini berasal dari data yang ada pada program studi sistem informasi FTI UMBY.

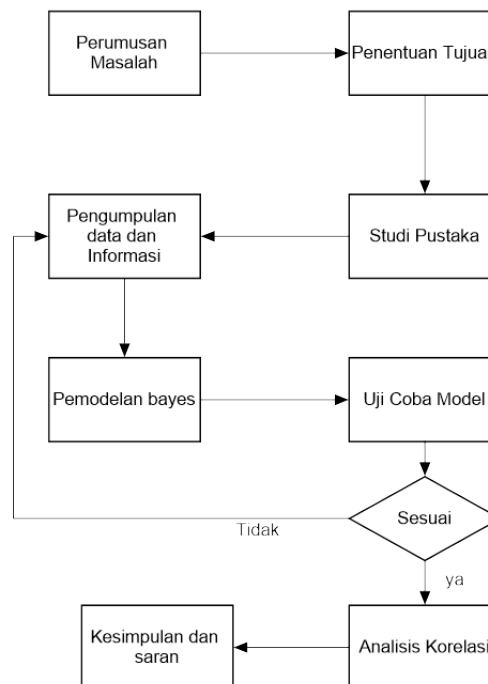
Berdasarkan eksperimen yang dilakukan terhadap 20 data yang dijadikan bahan testing, didapat 19 data yang sesuai dan 1 data yang tidak sesuai, sehingga menunjukkan nilai unjuk kerja sebesar 95% [4]. Syukri pada 2017 melakukan penelitian untuk mengklasifikasikan mahasiswa yang berpotensi lulus tepat waktu menggunakan metode naive bayes classifier dimana Syukri berkesimpulan bahwa semakin banyak data training yang digunakan, maka tingkat recall, precision dan accuracion sistem akan semakin baik, pada penelitian ini naive bayes mendapatkan tingkat akurasi sebesar 92.30% dalam hal mengklasifikasi mahasiswa yang akan lulus tepat waktu dan yang akan lulus terlambat [8]. Dalam dunia akademis metode bayes tidak hanya digunakan untuk mengklasifikasikan untuk prediksi siswa dapat lulus tepat waktu atau tidak, tapi juga digunakan untuk mendiagnosis siswa yang bermasalah untuk membantu pekerjaan guru BK seperti pada penelitian yang dilakukan oleh Bahar pada tahun 2017 lalu dengan nilai akurasi 100% [9].

Berdasarkan beberapa penelitian yang dijadikan rujukan, terdapat beberapa perbedaan antara penelitian yang telah dilakukan sebelumnya dengan penelitian ini. Metode yang digunakan pada penelitian ini menggunakan naive bayes untuk mengklasifikasi mahasiswa yang ada kedalam label potensi diri atau bakat, mencari tahu ada tidaknya pengaruh dari aspek bakat atau potensi diri terhadap IPK mahasiswa Fakultas Ilmu Komputer Universitas Amikom Yogyakarta serta mencari tahu potensi terbaik dari masing-masing program studi. Penelitian ini diharapkan dapat melengkapi penelitian-penelitian sebelumnya terkait mengetahui faktor yang memiliki pengaruh terhadap IPK sehingga nantinya dapat dijadikan parameter dalam perancangan model pengambilan keputusan bagi mahasiswa.

2. METODE PENELITIAN

2.1. Metodology Penelitian

Methodology yang digunakan pada penelitian ini terdiri dari 8 tahapan yang selanjutnya dapat dilihat pada Gambar 1 di bawah ini.



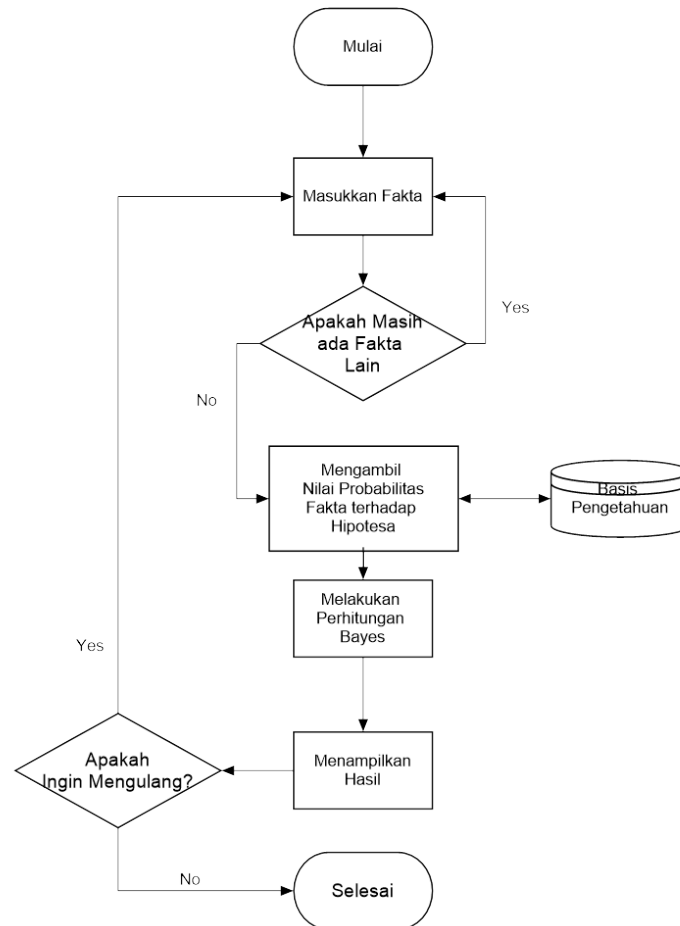
Gambar 1. Methodology penelitian

Pada Gambar 1 dapat dilihat alur penelitian yang dilakukan, mulai dari tahapan perumusan masalah, menentukan tujuan, melakukan studi pustaka, yakni mencari referensi yang berkaitan dengan penelitian yang dilakukan berdasarkan buku, jurnal, maupun penelitian-penelitian terdahulu. Metode pengumpulan data dengan cara melakukan wawancara terhadap psikolog dan mahasiswa semester 4 pada masing-masing program studi Fakultas Ilmu Komputer Universitas Amikom Yogyakarta, serta membagikan koesioner kepada mahasiswa Amikom Yogyakarta. Koesioner yang dibuat disesuaikan dengan hasil wawancara dengan psikolog, dimana didapatlah 8 label utama potensi diri, yakni kecerdasan visual, kecerdasan interpersonal, kecerdasan intrapersonal, kecerdasan linguistik, kecerdasan logika, kecerdasan musical, kecerdasan kinestetik dan kecerdasan naturalis. Pada penelitian ini peneliti membatasi potensi diri yang digunakan hanya 4 label, yakni kemampuan interpersonal, kemampuan logika, kemampuan visual dan kemampuan linguistik.

Berdasarkan data yang didapat dari koesioner, selanjutnya dilakukan konsultasi dengan pakar untuk membuat dataset dengan mengambil 30 dari 50 data, dan selanjutnya diberi label potensi diri paling dominan. Dataset yang ada dijadikan dasar dalam pembobotan pemodelan naive bayes, sedangkan 20 dari 50 data yang tersisa selanjutnya dijadikan data testing untuk menguji metode naive bayes. Penelitian ini menggunakan excell sebagai instrumen untuk membuat pemodelan naive bayes. Model naive bayes selanjutnya diuji dengan menggunakan confusion matrix untuk mengetahui nilai recall, precision, dan accuracy, jika pengujian dirasa belum sesuai maka melakukan pengumpulan data ulang, dan merubah dataset yang ada hingga dirasa sesuai. Langkah selanjutnya adalah melakukan analisis korelasi, metode yang digunakan adalah korelasi pearson product moment, sedangkan instrumen yang digunakan adalah IBM SPSS statistic versi 22. Pada tahap ini, nilai bayes dari masing-masing label pada data testing diuji korelasinya dengan variabel IPK dari masing-masing mahasiswa tersebut. Hasil dari analisis ini selanjutnya dijadikan kesimpulan dalam hal menjawab ada tidaknya pengaruh potensi diri dengan IPK mahasiswa tersebut, serta mencari tahu potensi diri mana yang terbaik pada masing-masing program studi.

2.2. Rancangan Flowchart Pemodelan Bayes

Dari Gambar 2 dapat dilihat langkah-langkah yang akan dilakukan dalam algoritma bayes, mulai dari memasukkan fakta, selanjutnya mengambil nilai probabilitas kemunculan fakta terhadap masing-masing hipotesa serta mengambil nilai probabilitas kemunculan masing-masing hipotesa tanpa dipengaruhi fakta dari basis pengetahuan. Melakukan perhitungan dengan menggunakan metode bayes. Mendapatkan hipotesa dengan nilai tertinggi yang akan dijadikan hasil akhir, selanjutnya adalah mengulangi proses mulai dari memasukkan fakta atau selesai dan memberhentikan proses. Daftar fakta atau pernyataan dapat dilihat pada Tabel 1.



Gambar 2. Rancangan Flowchart Pemodelan Bayes

Tabel 1. Daftar Fakta

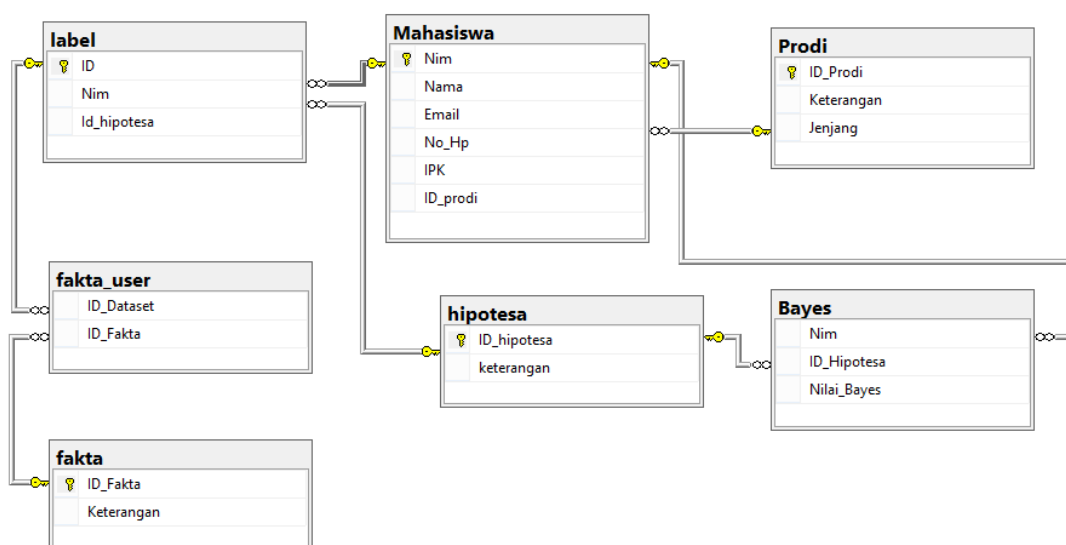
ID	Fakta
F1	Menjadi perwakilan dari suatu organisasi/institusi. Baik resmi maupun tidak resmi
F2	Mampu dengan mudah mengorganisir berbagai hal. Atau berbagai sumber daya yang dimilikinya
F3	Menyukai aktivitas memberikan perhatian atau perawatan dan dukungan kepada orang lain
F4	Senang atau mampu menciptakan sesuatu yang baru. seperti penulis, ilmuan dll.
F5	Memberikan atau mengirimkan sesuatu kepada orang-orang tertentu. Dan dalam jumlah tertentu
F6	Mampu membuat jurnal. Buku harian atau catatan kejadian sehari-hari. Atau menulis laporan dan berita untuk disiarkan
F7	Mampu merumuskan strategi promosi. Untuk mendorong orang agar mereka membeli lebih banyak produk/jasanya.
F8	Mampu menjadi penengah. Mengatasi dan menyelesaikan konflik antar dua pihak yang bertikai
F9	Membuat orang lain menjadi bersemangat. Tertarik. Dan berkomitmen untuk melakukan sesuatu sebaik mungkin
F10	Mengoperasikan dan menjaga mesin-mesin. Instrument atau peralatan lain.
F11	Memastikan kualitas yang memadai dan baik. Terutama dalam produk yang dibuat
F12	Mengembalikan sesuatu menjadi seperti kondisi/keadaan semula atau normal. Atau menjadi lebih baik

F13	Melindungi keselamatan atau keamanan sesuatu atau seseorang dari bahaya kerusakan, kehilangan dan pencurian
F14	Senang melayani dan mendahulukan orang lain
F15	Mampu dan pandai memilih seseorang untuk ditempatkan pada suatu posisi atau tugas tertentu
F16	Mampu melakukan perencanaan jangka panjang
F17	Gemar mengkombinasikan berbagai pandangan, ide, obyek, dll. Menjadi sesuatu hal yang baru
F18	Gemar berpikir jauh ke depan melampaui cakrawala

Tabel 1 menerangkan fakta apa saja yang dijadikan dasar dalam pembobotan model naive bayes, dimana dari total 18 fakta, mahasiswa diminta memilih 5-7 fakta atau pernyataan yang paling sesuai dengan dirinya.

2.3. Rancangan Basis Data

Basis data yang digunakan terdiri dari 7 tabel, tabel mahasiswa untuk menyimpan data responden, tabel prodi untuk menyimpan data program studi, tabel fakta untuk menyimpan daftar fakta, tabel label untuk menyimpan data responden yang dijadikan sebagai testing, tabel fakta user untuk menyimpan data fakta yang dipilih responden yang dijadikan testing, tabel hipotesa untuk menyimpan daftar label potensi diri yang dijadikan hipotesa dan tabel bayes digunakan untuk menyimpan data responden yang dijadikan sebagai data testing. Adapun rancangan basis datanya dapat dilihat pada Gambar 3.



Gambar 3. Rancangan Relasi Tabel

Dari gambar 3 diketahui bahwa tabel program studi berelasi dengan tabel mahasiswa, tabel label berelasi dengan mahasiswa dan hipotesa, tabel fakta user berelasi dengan tabel label dan tabel fakta, serta tabel bayes berelasi dengan tabel mahasiswa dan hipotesa. Untuk mendapatkan nilai probabilitas fakta terhadap hipotesa, dan probabilitas hipotesa secara dinamis maka dari rancangan database diatas dibuat 2 buah view seperti pada query berikut.

```

create view [Probabilitas_Hipotesa] AS
select Id_Hipotesa, count (Id_Hipotesa) as probabilitas_hipotesa from label
group by (Id_hipotesa)

create view [Probabilitas_Hipotesa_Fakta] as
select f.ID_fakta, count(f.ID_Fakta) as probabilitas_fakta, i.Id_hipotesa from label i
join fakta_user f on i.ID=f.ID_Dataset group by i.Id_hipotesa , f.ID_Fakta
    
```

Query yang digunakan untuk membuat view untuk menampilkan nilai probabilitas kemunculan hipotesa, dan nilai probabilitas kemunculan hipotesa berdasarkan fakta. Kedua nilai dari view tersebut yang nanti akan dijadikan dasar dalam perhitungan dengan algoritma naive bayes.

3. HASIL DAN PEMBAHASAN

Berdasarkan hasil studi pustaka dan pengumpulan data, langkah selanjutnya adalah melakukan pengolahan dari data yang sudah ada dengan algoritma *naive bayes*, untuk menentukan bakat atau potensi diri mahasiswa, berdasarkan fakta yang diinputkan.

3.1. Perhitungan Manual

Perhitungan manual dilakukan dengan cara mengklasifikasikan data testing kedalam beberapa label, berdasarkan *dataset* yang sudah dibuat. Diketahui nilai probabilitas kemunculan fakta terhadap hipotesa dapat dilihat pada tabel 2 dibawah ini.

Tabel 2. Nilai *Probabilitas* terhadap Hipotesa

Faktor	Nilai Probabilitas Kemunculan Fakta terhadap Hipotesa (%)			
	H1(15)	H2(6)	H3(3)	H4(6)
F1	80	70	70	10
F2	20	70	35	35
F3	80	35	70	70
F4	40	70	70	10
F5	60	10	35	70
F6	80	70	70	10
F7	20	35	100	70
F8	40	70	100	70
F9	60	10	70	70
F10	80	70	70	70
F11	10	35	35	35
F12	60	35	70	70
F13	10	10	0.1	70
F14	40	35	70	10
F15	40	70	70	70
F16	80	10	35	35
F17	80	35	10	70
F18	80	35	35	70

Keterangan :

F1 , F2, F3, F4, dan seterusnya merupakan id dari fakta yang ada dan dapat dilihat keterangannya pada tabel 1. H1 merupakan kemampuan Logika, H2 Kemampuan interpersonal, H3 kemampuan linguistik dan H4 merupakan kemampuan visual. Sedangkan nilai(15), (6), (3), (6) merupakan kemunculan hipotesa dari keseluruhan *dataset*, atau dari 30 *dataset* 16 diantaranya memiliki potensi diri dominan kemampuan logika, 6 kemampuan interpersonal, 3 kemampuan linguistik, dan 6 kemampuan visual.

Dari Tabel 2 diatas diketahui detail nilai kemunculan suatu fakta terhadap hipotesa tertentu, yang dapat juga diartikan bahwa saat fakta satu muncul (F1) maka kemungkinan akan menghasilkan hipotesa kemampuan logika sebesar 80%, kemampuan interpersonal 70% dan seterusnya. Angka tersebut didapat dari [2] menghitung kemunculan masing-masing fakta terhadap hipotesa yang ada pada *dataset*. Nilai probabilitas kemunculan hipotesa dari dataset yang ada dapat dilihat pada Tabel 3.

Tabel 3. Nilai Probabilitas Kemunculan Hipotesa

ID	Hipotesa	Probabilitas
H1	Kemampuan Logika	50
H2	Kemampuan Interpersonal	20
H3	Kemampuan Linguistik	10
H4	Kemampuan Visual	20

Dari tabel 3 diketahui bahwa Nilai probabilitas kemunculan hipotesa kemampuan logika pada *dataset* sebesar 50% atau 15 dari 30 data yang menjadi *dataset* memiliki kemampuan paling dominan yakni kemampuan logika. Nilai pada tabel 2 dan tabel 3 tersebut selanjutnya dijadikan dasar dalam perhitungan model *naive bayes* menggunakan persamaan 1 dibawah ini.

$$P(H_i | F) = \frac{P(F|H_i) \times P(H_i)}{\sum_{k=1}^n P(F|H_k) \times P(H_k)} \quad (1)$$

Keterangan:

$P(H_i | E)$ = Nilai probabilitas H_i dikatakan benar berdasarkan fakta F

$P(F | H_i)$ = Nilai probabilitas fakta F terhadap hipotesa H_i

$P(H_i)$ = Nilai probabilitas munculnya hipotesa H_i , tanpa memandang fakta apapun

n = Jumlah hipotesa.

Melakukan klasifikasi dengan menggunakan metode *naive bayes* apabila diketahui fakta pada data *testing* pertama sebagai berikut F_1, F_2, F_4, F_7, F_9 , langkah pertama adalah mencari nilai yang ada pada persamaan (2):

$$\sum_{k=1}^n P(F|H_i) \times P(H_i) \quad (2)$$

Keterangan :

$P(H_i)$ = Nilai probabilitas kemunculan H_i di dapat berdasarkan keseluruhan data pada *dataset* tanpa memperhatikan fakta.

- $(P(F_1 | H_1) \times P(F_2 | H_1) \times P(F_4 | H_1) \times P(F_7 | H_1) \times P(F_9 | H_1) \times P(H_1)) + (P(F_1 | H_2) \times P(F_2 | H_2) \times P(F_4 | H_2) \times P(F_7 | H_2) \times P(F_9 | H_2) \times P(H_2)) + (P(F_1 | H_3) \times P(F_2 | H_3) \times P(F_4 | H_3) \times P(F_7 | H_3) \times P(F_9 | H_3) \times P(H_3)) + (P(F_1 | H_4) \times P(F_2 | H_4) \times P(F_4 | H_4) \times P(F_7 | H_4) \times P(F_9 | H_4) \times P(H_4))$
- $(0.80 \times 0.20 \times 0.40 \times 0.20 \times 0.60 \times 0.50) + (0.70 \times 0.70 \times 0.70 \times 0.35 \times 0.10 \times 0.20) + (0.70 \times 0.35 \times 0.70 \times 1 \times 0.70 \times 0.1) + (0.10 \times 0.35 \times 0.10 \times 0.70 \times 0.70 \times 0.20)$
- $0.00384 + 0.002401 + 0.012005 + 0.000343$
- **0.018589**

Berdasarkan persamaan diatas didapatkanlah nilai jumlah probabilitas pada setiap hipotesa sebesar **0.018589**. Langkah selanjutnya adalah melakukan perhitungan pada setiap hipotesa menggunakan persamaan (1).

- $p(H_1 | F_1 F_2 F_4 F_7 F_9) = \frac{(0.8 \times 0.2 \times 0.4 \times 0.2 \times 0.6 \times 0.5)}{0.018589} = \mathbf{0.20657378}$
- $p(H_2 | F_1 F_2 F_4 F_7 F_9) = \frac{(0.7 \times 0.7 \times 0.7 \times 0.35 \times 0.1 \times 0.20)}{0.018589} = \mathbf{0.129162408}$

- $p(H_3 | F_1 F_2 F_4 F_7 F_9) = \frac{(0.7 \times 0.35 \times 0.70 \times 1 \times 0.7 \times 0.1)}{0.018589} = \mathbf{0.645812039}$
- $p(H_4 | F_1 F_2 F_4 F_7 F_9) = \frac{(0.1 \times 0.35 \times 0.1 \times 0.7 \times 0.7 \times 0.2)}{0.018589} = \mathbf{0.0184517726}$

Dari perhitungan diatas diketahui berdasarkan fakta yang didapat dari data testing pertama maka hipotesa yang diambil adalah H3 dengan nilai bayes sebesar 0.645812039, yang berarti mahasiswa tersebut cenderung memiliki potensi diri paling dominan yakni kemampuan linguistik sebesar 64%, diikuti dengan kemampuan logika sebesar 0.20657378 atau 20%. Dan memiliki kemampuan visual paling rendah bernilai 0.0184517726 atau hanya sebesar 1,84%. Langkah selanjutnya mengulangi proses perhitungan untuk melakukan klasifikasi pada data testing yang lain.

3.2. Pengujian Bayes

Pengujian yang dilakukan menggunakan *confusion matrix*, dimana hasil dari pemodelan bayes selanjutnya dibandingkan dengan hasil yang diberikan pakar, menyesuaikan dengan fakta yang ada pada data testing.

Tabel 4. Pengujian

No	Fakta	Hasil Bayes	Hasil Pakar	Sesuai (Y/T)
1	F1, F2, F4, F7, F9	H3	H3	Y
2	F1, F2, F3, F7, F16	H2	H2	Y
3	F4, F8, F11, F16, F18	H1	H1	Y
4	F5, F8, F10, F12, F16	H1	H1	Y
5	F6, F14, F15, F17, F18	H1	H2	T

Tabel 4 menampilkan sebagian proses uji coba yang telah dilakukan dengan kolom fakta merupakan fakta yang diberikan pada masing-masing data testing, kolom hasil bayes merupakan hasil keluaran dari model bayes, yakni hipotesa dengan nilai bayes tertinggi. Kolom hasil pakar merupakan klasifikasi yang dilakukan pakar terhadap fakta pada data testing, dan kolom sesuai merupakan perbandingan hasil sistem dengan hasil pakar. Adapun hasil pengujian pada seluruh data testing dapat dilihat pada tabel 5 dibawah ini.

Tabel 5. Confusion Matrix

Label		Klasifikasi Naive Bayes			
		Logika	Interpersonal	Linguistik	Visual
Klasifikasi Pakar	Logika	7	1	0	0
	Interpersonal	0	3	0	0
	Linguistik	0	1	3	0
	Visual	0	0	0	5
Total Testing		20			

Untuk mencari nilai akurasi, presisi dan recall pada data *multi-class* perlu mencari nilai rata-rata dari nilai akurasi, presisi dan recall terlebih dahulu[10]. Persamaan yang digunakan untuk mencari nilai akurasi, presisi dan recall dapat dilihat pada persamaan (3)(4)(5) berikut.

$$Akurasi = \sum_{i=1}^l \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} * 100\% \quad (3)$$

$$Presisi = \frac{\sum_{i=1}^l TP_i}{\sum_{i=1}^l (FP_i + TP_i)} * 100\% \quad (4)$$

$$Recall = \frac{\sum_{i=1}^l TP_i}{\sum_{i=1}^l (TP_i + FN_i)} * 100\% \quad (5)$$

Keterangan persamaan (3)(4)(5) :

- TPi adalah True Positive, yaitu jumlah data positif yang terklasifikasi dengan benar oleh sistem untuk label ke-i.
- TNi adalah True Negative, yaitu jumlah data negatif yang terklasifikasi dengan benar oleh sistem untuk label ke-i.
- FNi adalah False Negative, yaitu jumlah data negatif namun terklasifikasi salah oleh sistem untuk label ke-i.
- FPi adalah False Positive, yaitu jumlah data positif namun terklasifikasi salah oleh sistem untuk label ke-i
- l adalah jumlah label.

Mencari akurasi menggunakan persamaan (3)

- $((TP_1+TN_1)/TP_1+TN_1+FP_1+FN_1) + ((TP_2+TN_2)/TP_2+TN_2+FP_2+FN_2) +$
- $((TP_3+TN_3)/TP_3+TN_3+FP_3+FN_3) + ((TP_4+TN_4)/TP_4+TN_4+FP_4+FN_4)/4 * 100\%$
- $((7+0)/7+0+0+1) + ((3+0)/3+0+0+0) + ((3+0)/3+0+0+1) + ((5+0)/5+0+0+0)/4$
- $((0.875+1+0.75+1)/4) * 100\% = (3.625/4) * 100\% = 0.90625 * 100\% = \mathbf{90,625\%}$

Mencari Presisi dengan persamaan (4)

- $((TP_1+TP_2+TP_3+TP_4)/((FP_1+TP_1)+(FP_2+TP_2)+(FP_3+TP_3)+(FP_4+TP_4))) * 100\%$
- $((7+3+3+5)/((0+7)+(2+3)+(0+3)+(0+5))) * 100\% = (18/20) * 100\% = \mathbf{90\%}$

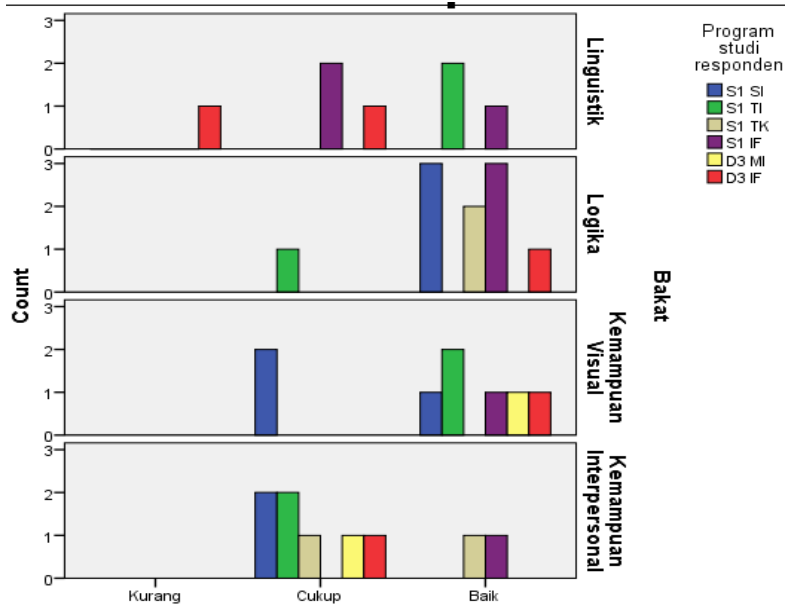
Mencari Recall dengan persamaan (5)

- $((TP_1+TP_2+TP_3+TP_4)/((TP_1+FN_1)+(TP_2+FN_2)+(TP_3+FN_3)+(TP_4+FN_4))) * 100\%$
- $((7+3+3+5)/((7+1)+(3+0)+(3+1)+(5+0))) * 100\% = (18/20) * 100\% = \mathbf{90\%}$

Berdasarkan hasil pengujian yang dilakukan, dari 20 percobaan terdapat 18 hasil yang sesuai dengan pakar, dan 2 hasil yang kurang sesuai, maka diketahui tingkat akurasi sebesar 90,625%, nilai presisi sebesar 90% dan nilai recall sebesar 90%. Akurasi yang diberikan mengalami penurunan dibandingkan dengan penelitian yang dilakukan M. Syukri [9] dengan metode yang sama bisa dipengaruhi oleh jumlah *dataset* yang digunakan.

3.3. Analisis Korelasi

Setelah data diklasifikasikan selanjutnya adalah mencari tahu ada tidaknya korelasi antara potensi diri dengan IPK mahasiswa. Adapun persebaran data potensi diri, IPK beserta program studi dapat dilihat pada Gambar 4.



Gambar 4. Persebaran Potensi diri dengan IPK Mahasiswa

Keterangan

Kurang : IPK <3

Cukup : IPK >=3 & IPK <=3.5

Baik : IPK >3.5

Count : Jumlah mahasiswa, berdasarkan bakat dan kategori IPK

Pada Gambar 3 dapat disimpulkan berdasarkan nilai persebaran IPK dengan potensi diri mahasiswa, bahwa mahasiswa yang memiliki potensi diri dominan pada kemampuan logika cenderung akan memiliki nilai IPK yang tinggi pada program studi S1 sistem informasi, S1 Informatika , serta S1 teknik komputer. Mahasiswa yang memiliki potensi diri dominan pada kemampuan visual cenderung akan memiliki IPK yang tinggi jika mengambil program studi S1 teknologi informasi, D3 manajemen informasi dan D3 informatika.

Selanjutnya melakukan analisis korelasi menggunakan SPSS untuk mencari tahu hubungan antara IPK dengan masing-masing label potensi menggunakan analisis korelasi *pearson*. Seluruh nilai *bayes* yang telah terklasifikasi pada data testing selanjutnya dibulatkan, lalu diuji korelasinya dengan IPK dari masing-masing data *testing*. Sebagian hasil klasifikasi dari data yang diuji dapat dilihat pada Tabel 5.

Tabel 5. Korelasi

No	IPK	Logika	Interpersonal	Linguistik	Visual
1	3.00	20	13	65	2
2	3.24	8	78	11	1
3	3.40	32	52	5	8
4	3.78	71	16	8	4
5	3.58	54	6	2	36

Dari tabel 5 diketahui bahwa data testing pertama memiliki IPK 3.00 dengan potensi diri paling dominan yakni kemampuan linguistik dengan 65%, diikuti kemampuan logika sebesar 20%, kemampuan interpersonal 13% dan kemampuan visual 2%. Hasil uji korelasi antara masing-masing label potensi diri dengan IPK mahasiswa dapat dilihat pada Tabel 6.

Tabel 6. Korelasi IPK dengan Potensi Diri

		V1	V2	V3	V4
IPK	Pearson Correlation	0.456*	-0.488	-0.391	0.535*
	Sig (2-tailed)	0.043	0.029	0.094	0.015
	N	20	20	20	20

Keterangan :

V1 = Kemampuan Logika

V2 = Kemampuan Interpersonal

V3 = Kemampuan Linguistik

V4 = Kemampuan Visual

*= 0.05

Berdasarkan Tabel 6 diatas diketahui nilai r hitung dari variabel kemampuan logika dan kemampuan visual dapat dilihat pada baris *pearson correlation* bernilai lebih dari nilai r tabel untuk data dengan 20 sampel yang bernilai 0.444, serta karena nilai r hitung variabel kemampuan logika dan visual bernilai positif maka berdasarkan nilai r hitung dapat diartikan bahwa variabel kemampuan logika dan kemampuan visual memiliki hubungan atau korelasi dengan IPK, dimana semakin tinggi kemampuan logika dan kemampuan visual mahasiswa, maka akan memiliki kecenderungan semakin tinggi juga IPK mahasiswa tersebut pada program studi fakultas ilmu komputer di Universitas Amikom Yogyakarta. Dilihat dari nilai signifikansi dari keempat variabel tersebut 3 variabel memiliki nilai signifikansi lebih kecil dari 0.05, yakni variabel kemampuan logika, kemampuan interpersonal, dan kemampuan visual maka dapat diartikan ketiga variabel tersebut memiliki korelasi yang signifikan terhadap meningkatnya IPK mahasiswa.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, berikut beberapa kesimpulan yang dihasilkan:

1. Potensi diri memiliki korelasi atau pengaruh terhadap IPK mahasiswa. Faktor kemampuan logika, interpersonal dan visual memiliki korelasi yang signifikan terhadap IPK mahasiswa FIK Universitas Amikom Yogyakarta, semakin tinggi kemampuan logika, interpersonal dan visual mahasiswa maka akan cenderung memiliki IPK yang semakin tinggi pula.
2. Potensi diri terbaik untuk program studi S1 sistem informasi, S1 Informatika, serta S1 teknik komputer adalah kemampuan logika. Potensi diri terbaik program studi S1 teknologi informasi, D3 manajemen informasi dan D3 informatika adalah kemampuan visual.
3. Metode bayes memiliki nilai akurasi sebesar 90% dalam hal mengklasifikasi potensi diri mahasiswa.

5. SARAN

Saran untuk peneliti yang ingin mengangkat tema sejenis yakni, menggunakan metode yang berbeda untuk mengetahui perbandingan nilai akurasi yang dihasilkan dalam klasifikasi bakat atau potensi diri. Menambahkan label yang digunakan dalam memetakan potensi diri serta menambah jumlah *dataset*, mengimplementasikan rancangan flowchart sistem dan rancangan basis data kedalam aplikasi baik berbasis desktop, web maupun mobile, serta menerapkan hasil dari penelitian ini untuk dijadikan salah satu faktor dalam memodelkan sistem pembantu keputusan untuk menentukan program studi berdasarkan bakat atau potensi diri mahasiswa.

DAFTAR PUSTAKA

- [1] Kementerian Riset, Teknologi, dan Pendidikan Tinggi, 2017, *Statistik Pendidikan Tinggi Tahun 2017, Pangkalan Data Pendidikan Tinggi*, Pusdatin Iptek Dikti, Setjen, Kemenristekdikti, Jakarta.
 - [2] Swastina, L, 2013, Penerapan Algoritma C4.5 untuk Penentuan Jurusan Mahasiswa. *Jurnal Gema Aktualita*, No. 1, Vol. 2, Hal. 93-98
 - [3] Febryanti, A. C., Darmawan, I., Andreswari, R., 2016, Pembobotan Kriteria Sistem Pendukung Keputusan Pemilihan Bidang Peminatan Menggunakan Metode Analytic Hierarchy Process (Studi Kasus: Program Studi Sistem Informasi Universitas Telkom), *Jurnal Rekayasa Sistem & Industri*, No. 4, Vol. 3, Hal. 7-15.
 - [4] Rozi, A. F., Purnomo, A. S., 2017, Rekomendasi Pemilihan Minat Studi Menggunakan Metode Mamdani Studi Kasus: Program Studi Sistem Informasi FTI UMBY, *Informatics Journal*, No. 3, Vol. 2, Hal. 138-147
 - [5] Beharu, W. T., 2018. Psychological Factors Affecting Students Academic Performance Among Freshman Psychology Students in Dire Dawa University, *Journal of Education and Practice*, No. 4, Vol. 9, Hal. 59-65
 - [6] Budiharto, W., 2014, *Artificial Intelligent*, Andi Offset, Yogyakarta.
 - [7] Sugiyono, 2011, *Statistika untuk Penelitian*, Alfabeta, Bandung.
 - [8] Mustafa, S. M., Ramadhan, M. R., Thenata, A. P., 2017, Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier, *Citec Journal*, No. 2, Vol 4, Hal. 151-162
 - [9] Bahar., Pratama, D. W., 2017, Penerapan Teorema Bayes Dalam Sistem Pakar Untuk Konsultasi Siswa Bermasalah, *JUITISI*, No. 2, Vol. 6, Hal. 1529-1540
 - [10] Sokolova, M., Lapalme, G., 2019, A Systematic analysis of performance measures for classification task, *Information Processing and Management*, No. 4, Vol. 45, Hal. 427-437.
-