

KOMBINASI ALGORITMA *FUZZY NAIVE BAYES* PADA PEMILIHAN BIDANG KEAHLIAN MAHASISWA TEKNIK INFORMATIKA UNISLA

Nurul Fuad

Program Studi Teknik Informatika, Universitas Islam Lamongan
E-mail: nurulfuad2@gmail.com

Abstract

The success of college can be seen in the smooth process of students in taking lectures on time, it means that if students determine the choice of majors in accordance with their field of expertise, then certainly there will be no difficulty to develop themselves in the lecture process. Therefore, in this study by using two algorithms, Fuzzy and Naive Bayes to determine the areas of expertise of students. Informatics Engineering students have the opportunity to choose a field of expertise, namely: the network expertise, programming, and design. Not all students will master the area of expertise, so they must choose according to their skills, hobbies, and favorite subjects, and grades. This research will create a system for classifying the expertise of UNISLA informatics engineering students. The algorithm implemented is Fuzzy to categorize the values or input numbers, while Naive Bayes is used for classification. The trial process results accuracy of 100%.

Keywords: Naive Bayes, Fuzzy, Science Expertise.

Abstrak

Keberhasilan perkuliahan di kampus bisa dilihat pada kelancaran mahasiswa dalam menempuh perkuliahan tepat waktu, artinya jika mahasiswa dalam menentukan pilihan jurusan sesuai dengan bidang keahliannya, maka dipastikan tidak akan ada kesulitan untuk mengembangkan diri dalam proses perkuliahan. Oleh karena itu dalam penelitian ini memanfaatkan dua algoritma, yaitu Fuzzy dan Naive Bayes untuk menentukan bidang keahlian mahasiswa. Mahasiswa Teknik Informatika mempunyai kesempatan dalam memilih bidang keahlian, yaitu: bidang keahlian jaringan, pemrograman, dan desain. Mahasiswa tidak semua akan menguasai bidang keahlian tersebut, sehingga harus memilih sesuai skill, hobby dan mata kuliah kesukaan, dan nilai. Penelitian ini akan membuat sistem untuk mengklasifikasi bidang keahlian mahasiswa Teknik Informatika UNISLA. Algoritma yang diimplimentasikan, yaitu Fuzzy untuk mengkategorikan dari nilai atau inputan angka, sedangkan Naive Bayes digunakan untuk klasifikasi. Proses uji coba menghasilkan akurasi yang 100%.

Kata kunci: Naive Bayes, Fuzzy, Bidang Keahlian.

1. PENDAHULUAN

Untuk pengelompokkan mahasiswa berdasarkan minat bidang keahlian sangat

dibutuhkan di fakultas guna mendukung dan mempermudah kebijakan yang dikeluarkan oleh fakultas untuk mendapatkan *output*

sumber daya manusia yang sesuai kebutuhan masyarakat umum. Bidang keahlian merupakan bagian dari kurikulum berbasis kompetensi. Bidang keahlian adalah kumpulan dari beberapa mata kuliah pendukung yang akan mengantarkan mahasiswa menuju proses penyelesaian skripsi. Bidang peminatan yang ada pada Program Studi Teknik Informatika Unisla adalah Jaringan, Pemrograman, dan Desain. Mahasiswa harus mampu menjadi tenaga yang profesional, sehingga perlu disesuaikan dengan bidang keahlian. Untuk menentukan bidang keahlian tersebut dapat diambil dari beberapa variabel diantaranya *skill*, kesukaan mata kuliah, hobi dan IPK agar dapat memilih bidang sesuai dengan keahlian karena tidak jarang juga mahasiswa dalam menentukan bidang minat secara asal - asalan yang mereka minati tidak berdasarkan kriteria yang telah dianjurkan oleh jurusan Universitas Islam Lamongan. seiring dengan perkembangan teknologi permasalahan tersebut dapat diantisipasi dengan cara klasifikasi atau pengelompokan bidang minat sesuai dengan kriteria yang telah ditentukan.

Pada implementasi system ini digunakan untuk mengklasifikasikan atau menenukan bidang keahlian mahasiswa berdasarkan hobi, skill, mata kuliah favorit, IPK dengan Algoritma Fuzzy Naïve Bayes. Data *input* angka (IPK) dilakukan konversi ke bentuk kategori menggunakan Algoritma *Fuzzy* sedangkan untuk menentukan bidang keahlian mahasiswa menggunakan Algoritma *Naïve Bayes*.

1.1 Himpunan Fuzzy

Berikut gambaran himpunan fuzzy. Pada himpunan tegas (*crisp*), nilai keanggotaan suatu item x dalam suatu himpunan A , yang sering ditulis dengan $\mu_A[x]$, memiliki 2 kemungkinan, yaitu:

1. satu (1), yang berarti bahwa suatu item menjadi anggota dalam suatu himpunan, atau
2. nol (0), yang berarti bahwa suatu item tidak menjadi anggota dalam suatu himpunan.

Himpunan *fuzzy* memiliki 2 (dua) atribut, yaitu:

1. Linguistik, yaitu penamaan suatu grup yang mewakili suatu keadaan atau kondisi tertentu dengan menggunakan bahasa alami.
2. Numeris, yaitu suatu nilai (angka) yang menunjukkan ukuran dari suatu variabel.

Ada beberapa hal yang perlu diketahui dalam memahami Sistem *Fuzzy*, yaitu:

1. Variabel *Fuzzy*
Variabel *fuzzy* merupakan variabel yang hendak dibahas dalam suatu sistem *fuzzy*.
2. Himpunan *Fuzzy*
Himpunan *fuzzy* merupakan suatu grup yang mewakili suatu kondisi atau keadaan tertentu dalam suatu variabel *Fuzzy*.
Contoh: Variabel temperatur, terbagi menjadi 5 himpunan *fuzzy*, yaitu: Dingin, Sejuk, Normal, Hangat, dan Panas.

Fungsi Keanggotaan

Fungsi Keanggotaan (*membership function*) adalah suatu kurva yang menunjukkan pemetaan titik-titik *input* data ke dalam nilai keanggotaannya (sering juga disebut dengan derajat keanggotaan) yang memiliki interval antara 0 sampai 1. Salah satu cara yang dapat digunakan untuk mendapatkan nilai keanggotaan adalah dengan melalui pendekatan fungsi. Ada beberapa fungsi yang bisa digunakan.

Penalaran Monoton

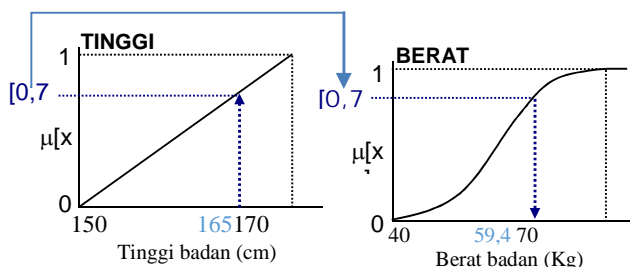
Metode penalaran secara monoton digunakan sebagai dasar untuk teknik implikasi *fuzzy*. Meskipun penalaran ini

sudah jarang sekali digunakan, namun terkadang masih digunakan untuk penskalaan *fuzzy*. Jika 2 (dua) daerah *fuzzy* direlasikan dengan implikasi sederhana sebagai berikut:

$$\begin{aligned} &\text{IF } x \text{ is } A \text{ THEN } y \text{ is } B \\ &\text{transfer fungsi:} \\ &y = f(x, A, B) \end{aligned}$$

maka sistem *fuzzy* dapat berjalan tanpa harus melalui komposisi dan dekomposisi *fuzzy*. Nilai *output* dapat diestimasi secara langsung dari nilai keanggotaan yang berhubungan dengan antesedennya.

Misalkan ada 2 himpunan *fuzzy* yaitu TINGGI (menunjukkan tinggi badan orang Indonesia) dan BERAT (menunjukkan berat badan orang Indonesia) seperti terlihat pada Gambar 2.1



Gambar 1. Himpunan *Fuzzy* TINGGI dan BERAT.

Hubungan antara kedua himpunan diekspresikan dengan aturan tunggal sebagai berikut:

IF TinggiBadan is TINGGI THEN BeratBadan is BERAT

Implikasi secara monoton akan menyeleksi daerah *fuzzy* A dan B dengan algoritma sebagai berikut:

- Untuk suatu elemen x pada domain A, tentukan nilai keanggotannya dalam daerah *fuzzy* A, yaitu: $\mu_A[x]$;
- Pada daerah *fuzzy* B, nilai keanggotaan yang berhubungan dengan tentukan permukaan *fuzzy*-nya. Tarik garis lurus ke arah domain. Nilai pada sumbu domain, y ,

merupakan solusi dari fungsi implikasi tersebut. Dapat dituliskan:

$$y_B = f(\mu_A[x], D_B)$$

Gambar 1 menunjukkan kerja algoritma tersebut. Seseorang yang memiliki tinggi badan 165 cm, memiliki derajat keanggotaan 0,75 pada daerah *fuzzy* TINGGI diperoleh dari:

$$\begin{aligned} \mu_{\text{TINGGI}}[165] &= (165 - 150) / (170 - 150) \\ &= 15 / 20 \\ &= 0,75 \end{aligned}$$

Nilai ini dipetakan ke daerah *fuzzy* BERAT yang akan memberikan solusi berat badan orang tersebut yaitu 59,4 kg diperoleh dari:

$$\mu_{\text{BERAT}}[y] = S(y; 40, 55, 70) = 0,75$$

Karena $0,75 > 0,5$, maka letak y adalah antara 52,5 sampai 70, sehingga:

$$\begin{aligned} \Leftrightarrow 1 - 2[(70 - y) / (70 - 40)]^2 &= 0,75 \\ \Leftrightarrow 1 - 2(70 - y)^2 / 900 &= 0,75 \\ \Leftrightarrow 2(70 - y)^2 / 900 &= 0,25 \\ \Leftrightarrow (70 - y)^2 &= 112,5 \\ \Leftrightarrow (70 - y) &= \pm \sqrt{112,5} \\ \Leftrightarrow y = 70 \pm 10,6 &\text{---> ambil (-) nya, karena nilainya harus } < 70 \\ \Leftrightarrow y = 59,4 \end{aligned}$$

Tiap-tiap aturan (proposisi) pada basis pengetahuan *fuzzy* akan berhubungan dengan suatu relasi *fuzzy*. Bentuk umum dari aturan yang digunakan dalam fungsi implikasi adalah:

IF x is A THEN y is B

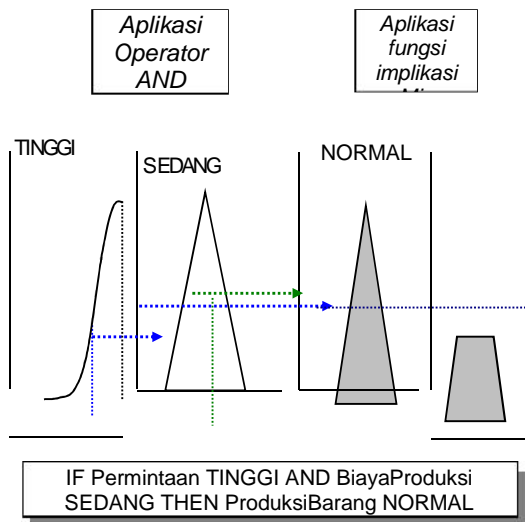
dengan x dan y adalah skalar, A dan B adalah himpunan *fuzzy*. Proposisi yang mengikuti IF disebut sebagai anteseden, sedangkan proposisi yang mengikuti THEN disebut sebagai konsekuen. Proposisi ini dapat diperluas dengan menggunakan *operator fuzzy*, seperti:

IF (x_1 is A_1) • (x_2 is A_2) • (x_3 is A_3) • • (x_N is A_N) THEN y is B

dengan adalah *operator* (misal: OR atau AND).

Secara umum, ada 2 (dua) fungsi implikasi yang dapat digunakan, yaitu:

- a. Min (*minimum*). Fungsi ini akan memotong *output* himpunan *fuzzy*. Gambar 2.9 menunjukkan salah satu contoh penggunaan fungsi min.



Gambar 2. Fungsi Implikasi

1.2 Naive Bayes

Algoritma *Naive Bayes* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naive dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naive Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. Persamaan dari Teorema Bayes adalah:

$$P(H | E) = \frac{P(E | H) * P(H)}{P(E)} \quad (1)$$

Keterangan:

- X : Data dengan *class* yang belum diketahui
- H : Hipotesis data X merupakan suatu *class* spesifik
- P(H|X) : Probabilitas hipotesis H berdasar kondisi X (posteriori probability)
- P(H) : Probabilitas hipotesis H (prior probability)
- P(X|H) : Probabilitas X berdasarkan kondisi pada hipotesis H
- P(X) : Probabilitas X

Dalam penelitian Nafi'iyah dan Mujilawati (2018) menyebutkan bahwa algoritma *Naive Bayes* dapat digunakan dalam proses identifikasi jenis kelamin manusia dengan nilai akurasi 80%, dan inputan dataset dari *Naive Bayes* lebih baik dalam bentuk kategori.

Penelitian lainya dalam klasifikasi, yaitu penelitian Nafi'iyah dan Faticah (2018) algoritma ID3 dataset inputan lebih baik dibuat dalam bentuk kategori, sehingga algoritma *fuzzy* digunakan sebagai cara mengkonversi inputan angka ke kategori.

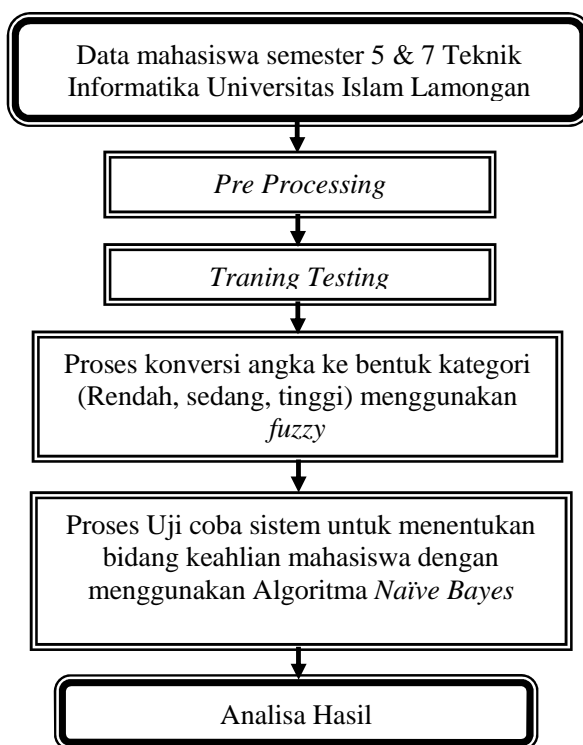
Penelitian Darmawan (2017) proses penentuan dosen pembimbing seminar proposal menggunakan data 100 mahasiswa yang berbeda judul beserta dosen pembimbing. Dari data lama tersebut digunakan untuk menentukan mahasiswa baru yang akan mengajukan proposal skripsi menggunakan algoritma *Naive Bayes*.

Algoritma *Naive Bayes* harus menggunakan dataset terlebih dahulu sebagai sumber belajar dalam menentukan keputusan. Dataset dari Algoritma *Naive Bayes* lebih baik dalam bentuk kategori, agar dapat mengubah bentuk kategori diterapkan Algoritma *Fuzzy*. Di mana *fuzzy* terlebih dahulu diset himpunan

keanggotaannya, selanjutnya dilakukan tahap inferensi menggunakan *rule*, dan dihasilkannya *value* kategori.

2. METODE PENELITIAN

Penelitian ini merancang dan mengembangkan sebuah sistem untuk menentukan bidang keahlian mahasiswa pada Universitas Islam Lamongan Jurusan Teknik Informatika. Alur penelitian dapat dilihat pada Gambar 3.



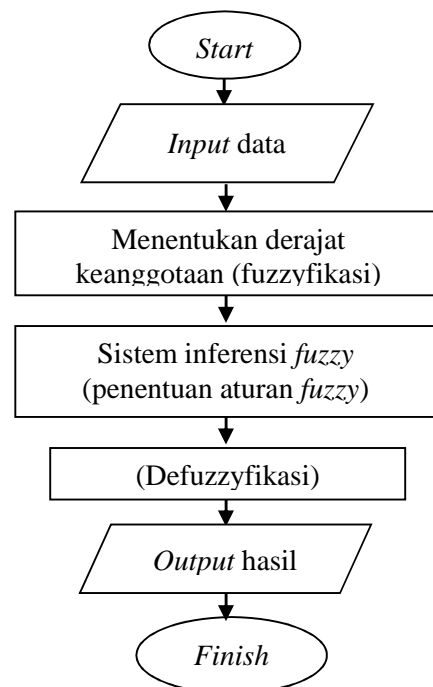
Gambar 3. Alur Penelitian

Langkah pertama yang dilakukan adalah mengumpulkan data informasi mahasiswa semester 5 dan 7 melalui kuisioner *online* kemudian melakukan *preprocessing* dengan cara melakukan penapisan data untuk selanjutnya disimpan pada *database* yang nantinya dijadikan data *training* dan *testing* pada uji coba bidang minat. Sistem pemilihan bidang keahlian ini berdasarkan *skill*, hobi, dan mata kuliah kesukaan, dan nilai. Agar dapat membantu mahasiswa dalam memilih bidang keahlian, maka peneliti akan

membuat sistem untuk mengklasifikasi bidang keahlian mahasiswa Teknik Informatika Universitas Islam Lamongan. Algoritma yang digunakan, yaitu *fuzzy* untuk mengkategorikan dari nilai atau inputan angka, sedangkan *Naïve Bayes* digunakan untuk klasifikasi.

2.1 Proses Konversi IPK Menggunakan Fuzzy

1. Inputan sistem ada data masasiswa semester 5 & 7
2. *Pre processing* ini adalah menyiapkan data dan melakukan *filtering* pada data yang kurang layak
3. Pada bagian ini ada proses konversi angka ke bentuk kategori (Rendah, sedang, tinggi) menggunakan *fuzzy*.



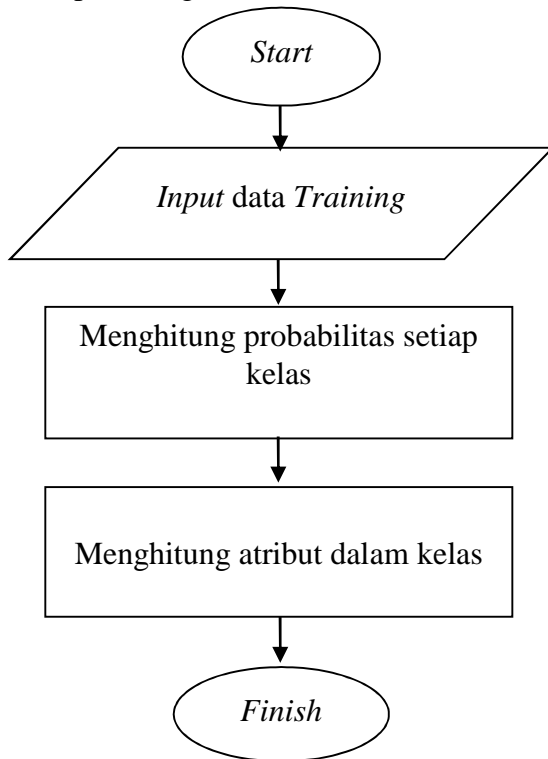
Gambar 4. Alur Konversi IPK dengan Fuzzy

2.2 Proses Pemilihan Jurusan Menggunakan Naive Bayes

a. Proses Training

Pada proses *training* ini data mahasiswa yang telah melalui *preprocessing* dan pengolahan *fuzzy* akan dimasukkan *database*

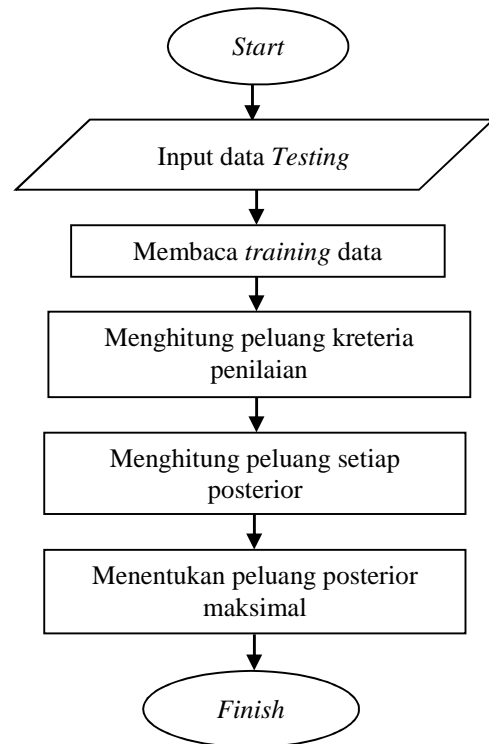
kemudian dilakukan penghitungan probabilitas menggunakan *Naive Bayes* selanjutnya menentukan kelas berdasarkan hasil perhitungan.



Gambar 5. Pemilihan Jurusan dengan *Naive Bayes*

b. Proses Testing

Pada proses data testing (Gambar 6) data diinputkan atau *diload* dari *database* kemudian dilakukan pembelajaran dari data *training* dan selanjutnya dilakukan perhitungan peluang berdasarkan kriteria secara sistem dan terakhir menghitung peluang posterior dan hasil



Gambar 6. Proses *Testing*

3. HASIL DAN DISKUSI

Konversi IP Menggunakan Fuzzy

Proses pengolahan IPK menggunakan *fuzzy* adalah dari data awal yang telah *diimport* kedalam tabel kemudian dilakukan perhitungan menggunakan program dan menghasilkan nilai baru seperti ditampilkan pada form data IP hasil *Fuzzy*, fitur baru ini yang akan digunakan untuk proses data *training* menggantikan IPK yang lama, selanjutnya untuk proses *training* dengan melakukan pilih tombol latihan dengan *Naive Bayes*. Hasilnya dapat dilihat pada Gambar 7.



Gambar 7. Konversi Nilai IP Menggunakan *Fuzzy*

Tabel 1. Contoh Data *Training*

Nama	Hobby	Mata Kuliah (Favorit)	IPK	Bidang Minat
MIRZA	Olahraga	Alpro	3,68	Pemrograman
HASAN	Membaca	Islamologi	3,86	Desain
HIMAM	<i>Editing</i>	B. Inggris	3,23	Desain
ALVIN	<i>Editing</i>	Alpro	3,14	Desain
YAFI	Membaca	Pancasila	3,09	Jaringan
FILA	<i>Editing</i>	Alpro	3,86	Desain
QORNI	Olahraga	B. Inggris	2,77	Design
RIZKY	<i>Editing</i>	Islamologi	3,36	Jaringan
ROSI	Olahraga	Islamologi	3,36	Pemrograman
ROSIDA	Olahraga	Pancasila	3,36	Pemrograman

Pada Tabel 2 berikut ini adalah contoh perhitungan *testing*.

Tabel 2. Contoh Data *Testing*

Nama	Hobby	Mata Kuliah (Favorit)	IPK	Bidang Minat
xxxx	<i>editing</i>	alpro	33,14	?

Implementasi Algoritma *Naive Bayes* untuk menentukan bidang minat menggunakan data *training* Tabel 1.

Dalam proses perhitungan menggunakan Algoritma *Naive Bayes*, ada beberapa langkah yang dilakukan, berikut ini adalah langkah-langkahnya:

1. Tahap 1 menghitung jumlah *class/label*
 $P(Y= \text{DESAIN}) = 5/10$ ‘jumlah data “DESAIN” pada kolom ‘BIDANG MINAT’ dibagi jumlah data

$P(Y= \text{PEMOGRAMAN}) = 3/10$ ‘ jumlah data “PEMOGRAMAN” pada kolom ‘BIDANG MINAT’ dibagi jumlah data

$P(Y= \text{JARINGAN}) = 2/10$ ‘ jumlah data “JARINGAN” pada kolom ‘BIDANG MINAT’ dibagi jumlah data

2. Tahap 2 menghitung jumlah kasus yang sama dengan class yang sama

$$P(\text{HOBBY} = \text{EDITING} \mid Y= \text{DESAIN}) = 3/5$$

‘jumlah data *Hobby* “Editing” dengan keterangan “DESAIN” dibagi jumlah data DESAIN

$$P(\text{HOBBY} = \text{EDITING} \mid Y= \text{PEMOGRAMAN}) = 2/3$$

‘jumlah data *Hobby* “Editing” dengan keterangan “PEMOGRAMAN” dibagi jumlah data PEMOGRAMAN

$$P(\text{HOBBY} = \text{EDITING} \mid Y= \text{JARINGAN}) = 1/2$$

jumlah data *Hobby* “Editing” dengan keterangan “JARINGAN” dibagi jumlah data JARINGAN

$$P(\text{MATA KULIAH} = \text{ALPRO} \mid Y= \text{DESAIN}) = 2/5$$

'jumlah data MK "ALPRO" dengan keterangan "DESAIN" dibagi jumlah data DESAIN

$P(\text{KULIAH} = \text{ALPRO} \mid Y = \text{PEMOGRAMAN}) = 1/3$
 jumlah data MK "ALPRO" dengan keterangan "PEMOGRAMAN" dibagi jumlah data PEMOGRAMAN

$P(\text{KULIAH} = \text{ALPRO} \mid Y = \text{JARINGAN}) = 0/2$

jumlah data MK "ALPRO" dengan keterangan "JARINGAN" dibagi jumlah data JARINGAN

$P(\text{IPK} = 3.23 \mid Y = \text{DESAIN}) = 1/5$

'jumlah data IPK "3.23" dengan keterangan "DESAIN" dibagi jumlah data DESAIN

$P(\text{IPK} = 3.23 \mid Y = \text{PEMOGRAMAN}) = 0/3$

'jumlah data IPK "3.23" dengan keterangan "PEMOGRAMAN" dibagi jumlah data PEMOGRAMAN

$P(\text{IPK} = 3.23 \mid Y = \text{JARINGAN}) = 0/2$

'jumlah data IPK "3.23" dengan keterangan "JARINGAN" dibagi jumlah data JARINGAN

3. Tahap 3 kalikan semua hasil variabel DESAIN, PEMROGRAMAN, & JARINGAN

$P(\text{HOBBY} = \text{EDITING}), (\text{MATA KULIAH} = \text{ALPRO}), (\text{IPK} = 3.23) \text{DESAIN}$

$= \{P(\text{HOBBY} = \text{EDITING} \mid Y = \text{DESAIN}), P(\text{MATA KULIAH} = \text{ALPRO} \mid Y = \text{DESAIN}), P(\text{IPK} = 3.23 \mid Y = \text{DESAIN})\}$
 $= 3/5.2/5.1/5 = 0.048$

$P(\text{HOBBY} = \text{EDITING}), (\text{MATA KULIAH} = \text{ALPRO}), (\text{IPK} = 3.23) \text{PEMOGRAMAN}$

$= \{P(\text{HOBBY} = \text{EDITING} \mid Y = \text{DESAIN}), P(\text{KULIAH} = \text{ALPRO} \mid Y = \text{PEMOGRAMAN}), P(\text{IPK} = 3.23 \mid Y = \text{DESAIN})\}$
 $= 2/3.1/3.0/3 = 0$

$P(\text{HOBBY} = \text{EDITING}), (\text{MATA KULIAH} = \text{ALPRO}), (\text{IPK} = 3.23) \text{JARINGAN}$

$= \{P(\text{HOBBY} = \text{EDITING} \mid Y = \text{DESAIN}), P(\text{KULIAH} = \text{ALPRO} \mid Y = \text{JARINGAN}), P(\text{IPK} = 3.23 \mid Y = \text{DESAIN})\}$
 $= 1/2.0/2.0/2 = 0$

4. Tahap 4 Bandingkan hasil *class* DESAIN, PEMOGRAMAN & JARINGAN

Karena hasil $(P|\text{DESAIN})$ lebih besar dari $(P|\text{PEMOGRAMAN})$ lebih besar dari $(P|\text{JARINGAN})$, maka keputusannya adalah "DESAIN"

Tabel 3. Contoh Hasil Data *Testing*

No	Nama	Hobby	Mata Kuliah (Favorit)	IPK	Bidang Minat
1	xxxx	<i>editing</i>	alpro	3,14	DESAIN

4. KESIMPULAN

Hasil dari analisa dan uji coba dalam penelitian ini dapat disimpulkan bahwa:

1. Aplikasi yang dibuat pada penelitian ini adalah membuat program aplikasi klasifikasi bidang minat dengan menggunakan Program Matlab dan Algoritma *Fuzzy Naive Bayes* uji coba yang pertama menghasilkan akurasi 78.2% dari 81 data selanjutnya dilakukan perbaikan data, kemudian dilakukan uji coba yang ke dua akurasinya 100% dari 243 data yang diolah pada proses *training* dengan menggunakan kfold (10).
2. Algoritma *fuzzy* digunakan untuk konversi variabel ke dalam bentuk kategori dan Algoritma *Naive Bayes* digunakan untuk menentukan pemilihan bidang minat.

Saran

Pada penelitian selanjutnya disarankan untuk menggunakan kombinasi algoritma yang lainnya seperti SVM, KNN, bagpro, dll,

untuk mendapatkan hasil yang lebih maksimal dari data *training* bervariasi.

DAFTAR PUSTAKA

- Darmawan, R. 2017. Perancangan Implementasi Sistem Pendukung Keputusan Penentuan Dosen Pembimbing Seminar Proposal Menggunakan Metode Naive Bayes.
- Hartatik. 2017. Klasifikasi Konsentrasi Jurusan Mahasiswa Universitas Amikom Yogyakarta.
- Nafi'iyah N. dan Faticah, C. 2018. Metode Fuzzy ID3 untuk Klasifikasi Bentuk Wajah Manusia Menggunakan Dental Panoramic.
- Nafi'iyah N. dan Mujilahwati, S. 2018, Agustus. Analisis Algoritma Backpropagation dan Naive Bayes dalam Identifikasi Jenis Kelamin Manusia Berdasarkan Foto Panoramik Gigi.

HALAMAN INI SENGAJA DIKOSONGKAN