

**PERBANDINGAN 3 METODE DALAM DATA MINING  
UNTUK PENJURUSAN SISWA DI SMA N 3 BOYOLALI**

**NASKAH PUBLIKASI  
PROGRAM STUDI INFORMATIKA  
FAKULTAS KOMUNIKASI DAN INFORMATIKA**



**Diajukan oleh :**

*Syarifah Nur Haryati*

*Yusuf Sulistyono Nugroho, S.T., M.Eng.*

**PROGRAM STUDI INFORMATIKA  
FAKULTAS KOMUNIKASI DAN INFORMATIKA  
UNIVERSITAS MUHAMMADIYAH SURAKARTA**

**MARET, 2015**

HALAMAN PENGESAHAN

Publikasi Ilmiah dengan Judul :

PERBANDINGAN 3 METODE DALAM DATA MINING  
UNTUK PENJURUSAN SISWA DI SMA N 3 BOYOLALI

Yang dipersiapkan dan disusun oleh :

Syarifah Nur Haryati

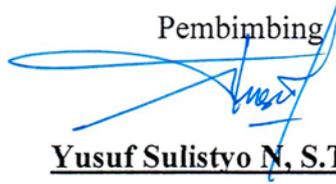
L200110067

Telah Disetujui Pada :

Hari ..<:<:'\|.....

Tanggal . 6 Mar-t\ 2-C\|f

Pembimbing



**Yusuf Sulistyono, S.T., M.Eng.**

NIK: 1197

Publikasi Ilmiah Ini Telah Diterima Sebagai Salah Satu Persyaratan

Untuk Memperoleh Gelar Sarjana

Tanggal .. .Y\ \..?:. \1

Mengetahui,

Ketua Program Studi

Informatika



**Dr. Heru Supriyono, M.Sc.**

NIK: 970



**UNIVERSITAS MUHAMMADIYAH SURAKARTA  
FAKULTAS KOMUNIKASI DAN INFORMATIKA  
PROGRAM STUDI INFORMATIKA**

Jl. A Yani Tromol Pos I Pabelan Kartasura Telp. (0271)717417, 719483 Fax (0271) 714448  
Surakarta 57102 Indonesia. Web : <http://informatika.ums.ac.id>. Email: [informatika@fki.ums.ac.id](mailto:informatika@fki.ums.ac.id)

---

**SURAT KETERANGAN LULUS PLAGIASI**

**/A.3-11.3/INF-FKI/11112015 ,**

Assalamu 'alaikum Wr. Wb

Biro Skripsi Program Studi Informatika menerangkan bahwa :

Nama	SYARIFAH NUR HARYATI
NIM	L200110067
Judul	PERBANDINGAN 3 METODE DALAM DATA MINING UNTUK PENJURUSAN SISWA DI SMA N 3 BOYOLALI
Program Studi	Informatika
Status	<b><u>Lulus</u></b>

Adalah benar-benar sudah lulus pengecekan plagiasi dari Naskah Publikasi Skripsi, dengan menggunakan aplikasi Turnitin.

Demikian surat keterangan ini dibuat agar dipergunakan sebagaimana mestinya.

Wassalamu 'alaikum Wr. Wb

*Surakarta, 11 Maret 2015*

Biro Skripsi  
Informatika

**Adjie Sapetra, S.Kom**

## Turnitin Originality Report

PERBANDINGAN 3 METODE DALAM  
DATA MINING UNTUK PENJURUSAN  
SISWA DI SMA N 3 BOYOLALI by Syarifah  
Nur Haryati

From publikasi (publikasi)

Processed on 10 Mar 2015 13:12 WIB  
ID: 514648058  
Word Count: 2329

Similarity Index

**29%**

Similarity by Source

Internet Sources:	1200
Publications:	1%
Student Papers:	27%

### sources:

1 5% match (student papers from 17-Jun-2014)  
Class: publikasi maret 2014  
Assignment:

Paper ID: [435295687](#)

2 5% match (student papers from 04-Feb-2014)  
Class: publikasi maret 2014  
Assignment:

PaperID: 393366361

3 5% match (student papers from 10-Mar-2015)  
Class: publikasi  
Assignment:

PaperID: 514646729

4 2% match (student papers from 08-Jul-2014)  
Class: publikasi maret 2014  
Assignment:

PaperID::i\_38 4757.60

5 2% match (student papers from 13-Jun-2014)  
Class: publikasi maret 2014  
Assignment:

Paper ID:

6 2% match (student papers from 04-Feb-2014)  
Class: publikasi maret 2014  
Assignment:

PaperID:393366374

7 2% match (student papers from 08-Jul-2014)  
Class: publikasi maret 2014  
Assignment:

Paper ID: [xjph134b](#)

8 1% match (Internet from 06-Nov-2012)

## **PERBANDINGAN 3 METODE DALAM DATA MINING UNTUK PENJURUSAN SISWA DI SMA N 3 BOYOLALI**

Syarifah Nur Haryati, Yusuf Sulisty Nugroho  
Program Studi Informatika, Fakultas Komunikasi dan Informatika  
Universitas Muhammadiyah Surakarta  
Email: [syarifah.1200110067@gmail.com](mailto:syarifah.1200110067@gmail.com)

### *Abstraksi*

SMA N 3 Boyolali merupakan salah satu Sekolah Menengah Atas di kota Boyolali yang terdapat 2 jurusan yaitu IPA dan IPS. Penjurusan siswa ini mengarahkan peserta didik agar lebih focus mengembangkan kemampuan dan minat yang dimiliki. Jurusan yang tidak tepat bisa sangat merugikan siswa dan masa depannya. Dengan penjurusan tersebut diharapkan dapat memaksimalkan potensi, bakat atau talenta individu, sehingga juga akan memaksimalkan nilai akademisnya. Untuk mengatasi permasalahan tersebut adalah dengan cara menerapkan proses *data mining*.

Teknik *data mining* yang digunakan dalam penentuan jurusan ini menggunakan metode *Decision Tree* Algoritma C4.5, *Naive Bayes* dan *Clustering* Algoritma *K-Means*. Sedangkan atribut yang digunakan terdiri dari Gender, Minat, Rata-rata IPA, Rata-rata IPS, Psikotest IPA, Psikotest IPS, Asal Sekolah dan Jurusan. Dalam melakukan analisa ini menggunakan bantuan software RapidMiner 5 untuk mengetahui metode apa yang paling baik.

Pengimplementasian *data mining* menggunakan perbandingan 3 metode dapat diketahui bahwa berdasarkan dari nilai *precision*, metode *naive bayes* lebih baik digunakan untuk penelitian ini dibandingkan dengan metode yang lain dengan nilai 77,51%. Sedangkan berdasarkan nilai *recall* dan *accuracy*, *decision tree* lebih baik digunakan dibanding metode yang lain dengan nilai *recall* sebesar 90,80% dan nilai *accuracy* sebesar 79,14%. Variabel yang paling berpengaruh dalam menentukan penjurusan yaitu rata-rata IPA sehingga perlu dijadikan pertimbangan bagi pihak sekolah.

**Kata kunci :** *Data Mining, Decision Tree Algoritma C4.5, Naive Bayes, Clustering Algoritma K-Means*

## PENDAHULUAN

SMA N 3 Boyolali yang berdiri sejak tahun 1989 merupakan salah satu Sekolah Menengah Atas di Kota Boyolali yang beralamat di Jalan Perintis Kemerdekaan, Pulisen, Boyolali. Akan tetapi di SMA tersebut hanya tersedia dua jurusan yaitu IPA dan IPS. Penjurusan siswa ini bertujuan mengarahkan peserta didik agar lebih fokus mengembangkan kemampuan dan minat yang dimiliki. Jurusan yang tidak tepat bisa sangat merugikan siswa dan masa depannya. Dengan penjurusan tersebut diharapkan dapat memaksimalkan potensi, bakat atau talenta individu, sehingga juga akan memaksimalkan nilai akademisnya.

Penentuan jurusan ini akan berdampak terhadap kegiatan akademik selanjutnya dan mempengaruhi pemilihan bidang ilmu atau studi bagi siswa-siswi yang ingin melanjutkan ke perguruan tinggi nantinya. Menentukan jurusan yang dilakukan secara manual mempunyai banyak kelemahan. Karena data yang digunakan cukup banyak sehingga menyita waktu dan

menguras tenaga, serta menuntut ketelitian ekstra.

Dengan melakukan *mining*, diharapkan dapat digali suatu potensi yang lebih dari sekedar informasi data sekolah saja tetapi juga dapat menganalisis penjurusan siswa. Sehingga dengan demikian dapat dianalisis penjurusan siswa yang sudah ada ataupun menemukan peluang-peluang yang baru serta menemukan rencana strategis dalam proses pengklasifikasian jurusan terhadap siswa.

Berdasarkan permasalahan tersebut, maka dalam penelitian ini akan dilakukan perbandingan terhadap 3 metode untuk penjurusan siswa menggunakan *data mining*, yaitu metode *decision tree* algoritma C4.5, *naive bayes* dan *clustering* dengan algoritma *k-means*. Dengan harapan setelah diolah dengan *data mining*, dapat memberikan rekomendasi bagi siswa yang mengalami kebingungan dalam mengambil pilihan jurusan serta memudahkan bagi pihak sekolah untuk menentukan penjurusan siswa.

## TINJAUAN PUSTAKA

### 1. Telaah Penelitian

Pada penelitian Nugroho (2014), data yang berlimpah membuka peluang diterapkannya *data mining* untuk pengelolaan pendidikan yang lebih baik dan *data mining* dalam pelaksanaan pembelajaran berbantuan komputer yang lebih efektif. Penelitian ini dilakukan untuk memanfaatkan data-data yang melimpah tersebut sebagai sumber informasi strategis untuk mengklasifikasi masa studi dan predikat kelulusan dengan menggunakan teknik *Decision Tree* algoritma C4.5 dan *Naive Bayes* digunakan untuk melakukan prediksi masa studi dan predikat kelulusan mahasiswa yang masih aktif. Interpretasi hasil penelitian mengindikasikan bahwa variabel yang perlu digunakan sebagai pertimbangan bagi fakultas untuk memperoleh tingkat masa studi yang efektif adalah renta SKS, sedangkan variabel yang perlu digunakan sebagai pertimbangan untuk memperoleh predikat kelulusan yang maksimal adalah peran serta mahasiswa untuk menjadi asisten

### 2. Landasan Teori

#### a. *Data Mining*

*Data Mining* adalah serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. (Munawaroh, 2013)

#### b. *Decision Tree*

*Decision Tree* adalah struktur pohon, dimana setiap *node* merepresentasikan atribut yang telah diuji, setiap cabang merupakan suatu pembagian hasil uji, dan *node* daun (*leaf*) merepresentasikan kelompok kelas tertentu. *Level node* teratas dari sebuah *Decision Tree* adalah *node* akar (*root*) yang biasanya berupa atribut yang paling memiliki pengaruh terbesar pada suatu kelas tertentu. (Hastuti, 2012)

#### c. *Naive Bayes*

*Naive Bayes* merupakan teknik prediksi berbasis *probabilistic* sederhana yang berdasar pada penerapan Teorema *Bayes* dengan asumsi independensi yang kuat. (Prasetyo, 2012)

#### d. Algoritma *K-Means*

Algoritma *K-Means* adalah algoritma klustering yang paling

sederhana dibandingkan dengan algoritma klustering lain. *K-Means* membagi data kemudian mengelompokkannya ke dalam beberapa klaster yang memiliki kemiripan dan memisahkan setiap klaster berdasarkan perbedaan antar masing-masing klaster. (Hariyadi, 2012)

## METODE PENELITIAN

### a. Penentuan Atribut

Setelah semua data dianalisa dan diseleksi melalui beberapa pertimbangan dari data yang diperoleh, ditetapkan atribut-atribut yang digunakan yaitu :

**Tabel 1.** Atribut yang digunakan

Atribut	Variabel
Jurusan	Y
Gender	X1
Minat	X2
Rata – rata IPA	X3
Rata – rata IPS	X4
Psikotest IPA	X5
Psikotest IPS	X6
Asal Sekolah	X7

### b. Penentuan sampel

Untuk mendapatkan sampel yang dapat menggambarkan dan mewakili jumlah populasi menggunakan bantuan metode *slovin* dengan nilai maksimal  $e = 5\%$ . (Umar, 2004) seperti pada **persamaan 1.**

$$\text{---} \dots\dots\dots (1)$$

Dimana :

$n$  = ukuran sampel

$N$  = ukuran populasi

$e$  = persen toleransi ketidakteelitian

### c. Pengelompokan Data

Setelah selesai menentukan atribut-atribut yang akan digunakan kemudian nilai dari atribut tersebut dikelompokkan untuk mempermudah dalam proses *data mining* menggunakan 3 metode tersebut.

### d. Implementasi Data Mining

#### 1. *Decision Tree* Algoritma C4.5

*Decision Tree* adalah salah satu metode untuk pengklasifikasian data. Hal yang harus dilakukan adalah menghitung *entropy* dan *information gain*. (Ranny dkk, 2012)

**Persamaan 2** Rumus *Entropy*

$$-p_1 \log_2 p_1 - p_2 \log_2 p_2 \dots - p_n \log_2 p_n \dots \dots \dots (2)$$

**Persamaan 3** Rumus *Information Gain* :

$$-\dots \dots \dots (3)$$

2. *Naive Bayes*

*Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan *statistic* untuk memprediksi peluang di masa depan. (Bustami, 2013)

**Persamaan 4** teorema *Bayes* :

$$\dots \dots \dots (4)$$

3. *Clustering* Algoritma *K-Means*

*K-Means Clustering* merupakan salah satu metode data *clustering* non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih *cluster*/kelompok. (Agusta, 2007)

**Persamaan 5** Rumus jarak *Euclidean* :

$$\dots \dots \dots (5)$$

**HASIL PEMBAHASAN**

**1. Pengambilan sampel**

Apabila diketahui siswa SMA N 3 Boyolali selama 5 tahun memiliki jumlah 1240 siswa dan memiliki toleransi ketidakteelitian 5%. Maka jumlah sampel yang diambil yaitu:

$$n = \dots$$

$$n = 1240 / 1 + 1240 \times (0,05)^2$$

$$n = 1240 / 1 + 1240 \times 0,0025$$

$$n = 1240 / 1 + 3,1$$

$$n = 12430 / 4,1$$

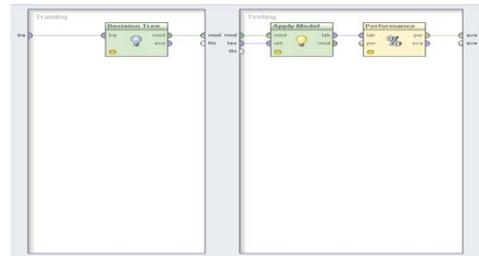
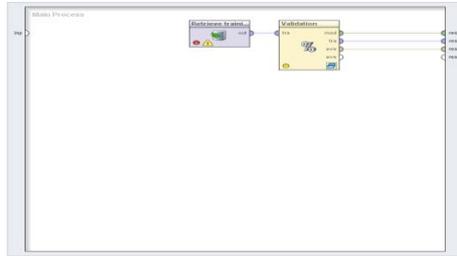
$$n = 302,43290 \text{ siswa}$$

Jadi dibulatkan menjadi 302 siswa yang digunakan bahan *sampling*.

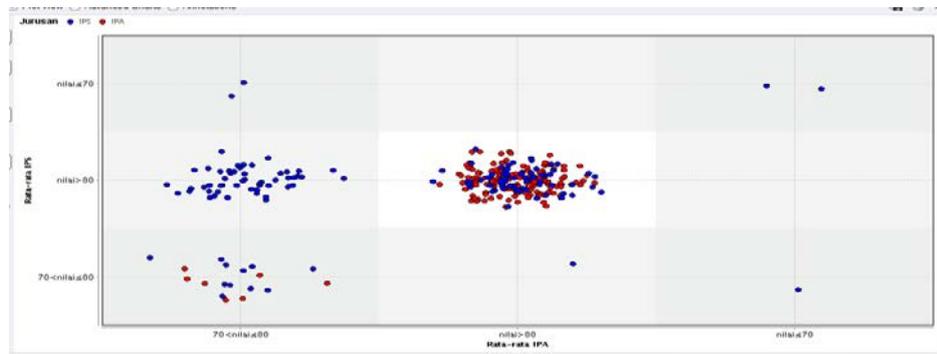
**2. Hasil Implementasi *Decision Tree* menggunakan *RapidMiner* 5**

Rancangan proses klasifikasi data penjurusan siswa menggunakan aplikasi

RapidMiner 5 ditunjukkan pada gambar 1.



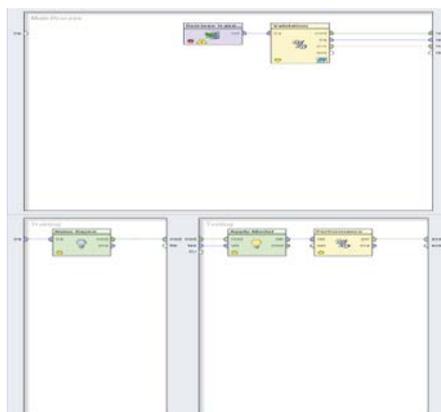
**Gambar 1.** Rancangan proses *Decision Tree*



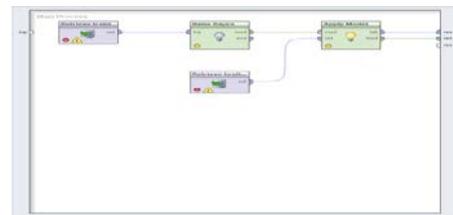
**Gambar 2.** Tampilan hasil *decision tree* pada *Scatter Plot*

### 3. Implementasi *Naive Bayes* menggunakan *RapidMiner 5*

Rancangan proses untuk prediksi data penjurusan siswa menggunakan *RapidMiner 5*.



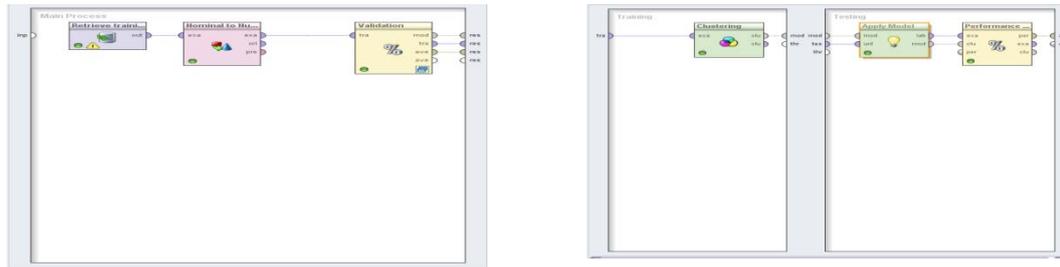
**Gambar 3.** Rancangan proses *Naive Bayes* menggunakan *data training*



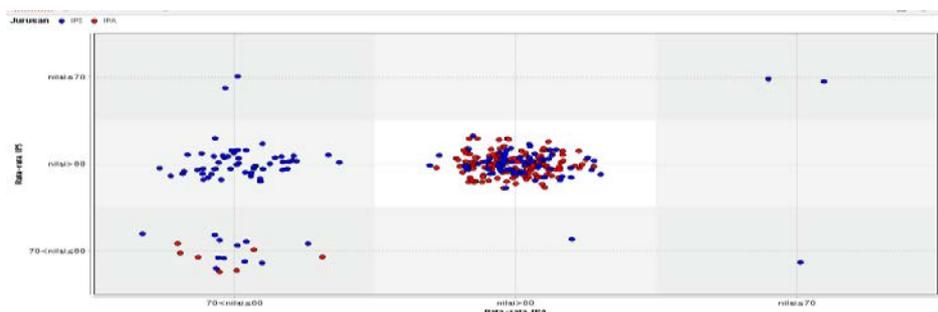
**Gambar 4.** Rancangan proses *Naive Bayes* menggunakan *data testing*

### 4. Implementasi *Clustering K-Means* menggunakan *RapidMiner 5*

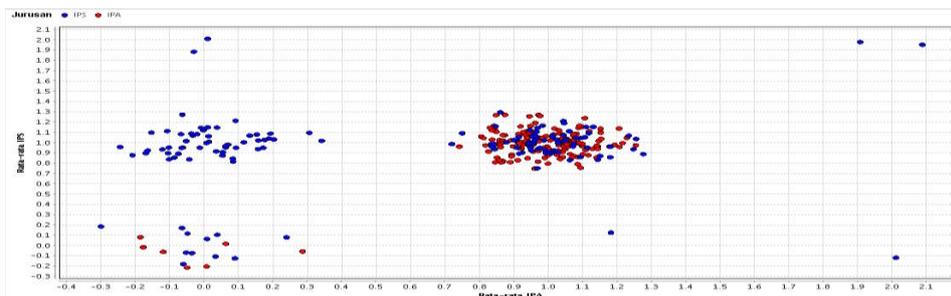
Rancangan proses untuk pengelompokan data penjurusan siswa menggunakan *RapidMiner5* ditunjukkan gambar 5.



**Gambar 5.** Rancangan proses *K-Means*



**Gambar 6.** Scatter Plot Naive Bayes menggunakan data training



**Gambar 7.** Tampilan hasil *Clustering K-Means* pada Scatter Plot

## 5. Implementasi Perhitungan *Decision Tree*

### a. Menentukan *Root Node*

Atribut yang memiliki nilai information gain paling tinggi di pilih sebagai *root node*.

**Tabel 2.** *Information gain* tertinggi

Atribut	Nilai <i>gain</i>
Gender	0,172
Minat	0,054
Rata-rata IPA	0,207
Rata-rata IPS	0,023
Psikotest IPA	0,003
Psikotest IPS	0,009
Asal Sekolah	0,011

b. Menentukan *Internal Node* pertama

Menentukan *internal node* pada rata-rata IPA  $70 < \text{nilai} \leq 80$  didapatkan nilai *information gain* seperti pada tabel 3.

Tabel 3. Nilai *Information gain*

Nilai Information Gain	Rata-rata IPA $70 < \text{nilai} \leq 80$
Gender	0,170
Minat	0,063
Rata-rata IPS	0,220
Psikotest IPA	0,029
Psikotest IPS	0,034
Asal Sekolah	0,040

Dari hasil tersebut dapat disimpulkan bahwa atribut Rata-rata IPS merupakan *internal node* pada rata-rata IPA  $70 < \text{nilai} \leq 80$ .

c. Menentukan *internal node* kedua

Menentukan *internal node* kedua pada rata-rata IPA  $70 < \text{nilai} \leq 80$  dan rata-rata IPS  $70 < \text{nilai} \leq 80$  didapatkan nilai *information gain* seperti pada tabel 4.

Tabel 4. Nilai *Information Gain*

Nilai Information Gain	Rata-rata IPA, IPS $70 < \text{nilai} \leq 80$
Gender	0,386
Minat	0,245
Psikotest IPA	0,065
Psikotest IPS	0,060
Asal Sekolah	0,114

d. Menentukan *leaf node*

Menentukan *leaf node* pada rata-rata IPA dan IPS  $70 < \text{nilai} \leq 80$  dengan *gender* laki-laki.

Tabel 5. Nilai *Information Gain*

Nilai Information Gain	Rata-rata IPA, IPS $70 < \text{nilai} \leq 80$ , Gender laki-laki
Minat	0,000
Psikotest IPA	0,000
Psikotest IPS	0,000
Asal Sekolah	0,000

Dari hasil tersebut dapat disimpulkan bahwa *gender* laki-laki menghasilkan *leaf node* IPS, dikarenakan hasil dari semua *information gain* bernilai 0 dan probabilitas siswa jurusan IPA tidak ada.

## 6. Implementasi Perhitungan *Naive Bayes*

Sebagai contoh diambil salah satu data uji yang memiliki

ciri sebagai berikut: Laki-laki, minat IPS, rata-rata IPA  $70 < \text{nilai} \leq 80$ , rata-rata IPS  $70 < \text{nilai} \leq 80$ , nilai psikotest IPA  $25 < \text{skor} \leq 45$ , psikotest IPS  $\text{skor} > 65$ , asal sekolah wilayah II. Apakah siswa tersebut masuk jurusan IPA atau IPS?

Perhitungan data test berdasarkan data *training* :

Fakta menunjukkan :

$$P(Y=IPA) = 163/302 = 0,539$$

$$P(Y=IPS) = 139/302 = 0,460$$

Fakta :

$$P(X_1=\text{Laki-laki} | Y=IPA)$$

$$= 28/163 = 0,171$$

$$P(X_1=\text{Laki-laki} | Y=IPS)$$

$$= 89/139 = 0,640$$

$$P(X_2=IPS | Y=IPA) = 17/163 = 0,104$$

$$P(X_2=IPS | Y=IPS) = 45/139 = 0,323$$

$$P(X_3=70 < \text{nilai} \leq 80 | Y=IPA)$$

$$= 7/163 = 0,042$$

$$P(X_3=70 < \text{nilai} \leq 80 | Y=IPS)$$

$$= 64/139 = 0,460$$

$$P(X_4=70 < \text{nilai} \leq 80 | Y=IPA)$$

$$= 7/163 = 0,042$$

$$P(X_4=70 < \text{nilai} \leq 80 | Y=IPS)$$

$$= 13/139 = 0,093$$

$$P(X_5=25 < \text{Skor} \leq 45 | Y=IPA)$$

$$= 37/163 = 0,226$$

$$P(X_5=25 < \text{Skor} \leq 45 | Y=IPS)$$

$$= 33/139 = 0,237$$

$$P(X_6=\text{Skor} > 65 | Y=IPA)$$

$$= 8/163 = 0,049$$

$$P(X_6=\text{Skor} > 65 | Y=IPS)$$

$$= 11/139 = 0,079$$

$$P(X_7=\text{Wilayah II} | Y=IPA)$$

$$= 102/163 = 0,625$$

$$P(X_7=\text{Wilayah II} | Y=IPS)$$

$$= 99/139 = 0,712$$

HMAP dari keadaan ini dapat dihitung :

$$P(X_1=\text{Laki-laki}, X_2=IPS, X_3=70 < \text{nilai} \leq 80, X_4=70 < \text{nilai} \leq 80, X_5=25 < \text{Skor} \leq 45, X_6=\text{Skor} > 65, X_7=\text{Wilayah II} | Y=IPA)$$

$$= 28/163 * 17/163 * 7/163 * 7/163 *$$

$$= \text{IPS}$$

$$\begin{aligned} & 37/163 * 8/163 * 102/163 * 163/302 \\ & = 0,000000117 \end{aligned}$$

$$\begin{aligned} & P(X_1=\text{Laki-laki}, X_2=\text{IPS}, X_3=70 < \text{nilai} \\ & \leq 80, X_4=70 < \text{nilai} \leq 80, X_5=25 < \text{Skor} \leq \\ & 45, X_6=\text{Skor} > 65, X_7=\text{Wilayah II} | Y \\ & = \text{IPS}) \end{aligned}$$

$$\begin{aligned} & = 89/139 * 45/139 * 64/139 * 13/139 * \\ & 33/139 * 11/139 * 99/139 * 139/302 \\ & = 0,0000542 \end{aligned}$$

❖ KEPUTUSAN PENJURUSAN

## 7. Implementasi Perhitungan Clustering K-Means

Pada metode ini untuk mengelompokan data penjurusan siswa dilakukan perhitungan dengan berbagai tahap, diantaranya:

a. Menentukan *cluster*

Pada tahap ini data penjurusan siswa dibagi menjadi 5 *cluster*.

b. Menentukan *Centroid*

Pada tahap ini mencari nilai *centroid* pada setiap pusat.

**Tabel 6.** Hasil *Centroid* per *cluster*

<i>Cluster</i>	<i>Centroid</i>						
	X1	X2	X3	X4	X5	X6	X7
1-60	7	9,2	10	11,6	12,4	18	7,2
61-120	7,2	10,2	10,6	11,6	11,8	19	12,2
121-180	6,8	9,2	9,2	11	14,6	21,2	13,2
181-240	7,2	9,4	9,8	12	12,6	17,4	10,4
241-302	8,8	10	7,2	11	13,6	18,6	8,8

Berdasarkan tabel diatas merupakan data *centroid* awal dari 5 *cluster* yang sudah ditentukan.

c. Menentukan *Euclidean distance*

Pada tahap ini mencari jarak antar data dalam melakukan pengelompokan data dengan menggunakan rumus *Euclidean distance*.

## 8. Hasil Perbandingan 3 Metode

Untuk mengetahui metode apa yang paling baik dalam penelitian ini dilakukan

perbandingan 3 metode dalam hitungan *Accuracy*, *Precision* dan *Recall*.

**Tabel 7.** Hasil Perbandingan 3 Metode

<b>Komponen</b>	<b><i>Decision Tree</i></b>	<b><i>Naive Bayes</i></b>	<b><i>K-Means</i></b>
<i>Accuracy</i>	79,14%	76,82%	36,40%
<i>Precision</i>	75,51%	77,51%	64,25%
<i>Recall</i>	90,80%	80,37%	25,40%

Berdasarkan hasil perbandingan diatas dapat disimpulkan bahwa metode *Decision Tree* lebih baik digunakan untuk penelitian ini dikarenakan tingkat *accuracy* dan *recall*-nya lebih tinggi dibandingkan dengan metode lainnya. Sedangkan metode *naive bayes* memiliki nilai *precision* yang lebih tinggi dibanding metode lainnya.

## 9. Interpretasi Hasil Penelitian

Interpretasi hasil penelitian mengindikasikan bahwa variabel yang perlu dipertimbangkan bagi pihak sekolah untuk penentuan jurusan di SMA N 3 Boyolali adalah variabel rata-rata IPA.

penelitian ini. Hal ini terbukti bahwa rata-rata IPA menempati posisi sebagai *root node*.

2. Berdasarkan dari nilai *precision*, metode *naive bayes* lebih baik digunakan untuk penelitian ini dibandingkan dengan metode yang lain dengan nilai 77,51%.
3. Berdasarkan nilai *recall* dan *accuracy*, *decision tree* lebih baik digunakan dibanding metode yang lain dengan nilai *recall* sebesar 90,80% dan nilai *accuracy* sebesar 79,14%.
4. Berdasarkan analisa dari 3 metode diatas, metode *decision tree* merupakan metode yang paling mudah dimengerti, dikarenakan metode ini memiliki pola yang jelas untuk mengetahui variabel apa yang paling berpengaruh.

## KESIMPULAN

1. Rata-rata IPA merupakan variabel yang paling mempengaruhi dalam

## DAFTAR PUSTAKA

- Agusta, Y. 2007. *K-means – Penerapan Permasalahan dan Metode Terkait*. Jurnal Sistem dan Informatika Vol.3 (Februari 2007):47-60.
- Bustami. 2013. '*Penerapan Algoritma Naive Bayes untuk Mengklasifikasi Data Nasabah Asuransi*', Jurnal Penelitian Teknik Informatika.
- Hariyadi, Teguh. 2012. '*Penerapan Algoritma K-Means untuk Pengelompokan Data Nilai Siswa*', Jurnal Jurusan D3 Manajemen Informatika Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang.
- Hastuti, Khafiizh. 2012. *Analisis Komparasi Algoritma Klasifikasi Data Mining untuk Prediksi Mahasiswa Non Aktif*. Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2012. ISBN 979 – 26 – 0255 – 0.
- Munawaroh, Holisatul. 2013. '*Perbandingan Algoritma ID3 dan C5.0 dalam Identifikasi Penjurusan Siswa SMA*', Jurnal Sarjana Teknik Informatika Vol.1, No.1.
- Nugroho, Yusuf Sulistyoyo. 2014. '*Klasifikasi dan Prediksi Masa Studi dan Prestasi Mahasiswa Fakultas Komunikasi dan Informatika Universitas Muhammadiyah Surakarta*', Jurnal KomuniTI, Vol VI, No 1, Maret 2014.
- Prasetyo, Eko. 2012. *Data Mining konsep dan aplikasi menggunakan matlab*. Penerbit Andi. Yogyakarta
- Ranny dan Budi. 2012. '*Pemilihan Diet Nutrien bagi Penderita Hipertensi Menggunakan Metode Klasifikasi Decision Tree*', Jurnal Teknik ITS, Vol. 1, No.1.
- Umar, Husein. 2004. *Metode Penelitian untuk Skripsi dan Tesis Bisnis*. Cetakan ke-6. PT Raja Grafindo Persada. Jakarta

## **BIODATA PENULIS**

Nama : Syarifah Nur Haryati

Tempat / Tanggal Lahir : Boyolali, 5 Januari 1994

Jenis Kelamin : Perempuan

Agama : Islam

Jurusan : Teknik Informatika

Peguruan Tinggi : Universitas Muhammadiyah Surakarta

Alamat : Jl. A. Yani Tromol Pos 1 Pabelan, Kartasura

Telp./Fax : (0271)717417, 719483 / (0271)714448

Alamat Rumah : Kemasan Rt.02/Rw.04, Sawit, Boyolali

No. HP : 085728822002

Alamat e-mail : [syarifah.l200110067@gmail.com](mailto:syarifah.l200110067@gmail.com)