

Implementasi Algoritma Naïve Bayes untuk Menentukan Kesiapan Siswa Dalam Menghadapi Ujian Nasional

Muchammad Ainurrofiq¹

^{1,3} Jurusan Teknik Informatika, FASILKOM UDINUS
Jln. Nakula 1 No 5-11 Semarang 50131 INDONESIA

¹rofiq@mhs.dinus.ac.id

Abstract— Try Out is the evaluation phase to face the National Examination indeed, readiness in the exam is essential to be done by students in grades 3, tryout a final exam practice for students in grade 3 will follow the national exam, ie. The school, teacher guidance and counseling has an important role in the delivery of services to students tryout. Services tryout before the national exam students require special handling professionally, because it involves the success of students in national exams. Mistakes in determining the readiness of students for national examinations can be a negative influence on the process and results of National Examination siswaitu own. So we need a method that can help minimize the impact of errors when determining the readiness of the student, by the technique of grouping data from the results of data mining. The need for data mining because of the large amount of data that can be used to produce useful information and knowledge. Naïve Bayes is a machine learning method that uses probability calculations. The use of this algorithm is assessed according as naive Bayesian classifier is a classification algorithm that is simple but has the capability and high accuracy. Function to predictions made using Matlab Tools, which produces an accuracy of 83.8798% with an error rate of 16.1202%

Keywords— Data Mining, Prediction, Majority, Algorithm Naïve Bayes, Matlab.

I. PENDAHULUAN

Apa yang kita amati di berbagai sekolah terutama berkaitan dengan akan dilaksanakannya Ujian Nasional (UN), di mana sekolah-sekolah mencari berbagai terobosan untuk dapat meningkatkan kualitas dan kuantitas lulusannya. Oleh karena itu strategi yang dilakukan adalah melakukan gerak cepat, terarah dan hasilnya memuaskan.

Banyak hal yang dilakukan oleh sekolah untuk mendongkrak agar nilai UN lebih baik. Tidak hanya sekolah-sekolah, pemerintah daerah pun baik tingkat provinsi maupun kabupaten/kota melakukan berbagai terobosan, sehingga daerah yang bersangkutan akan memperoleh peringkat terbaik.

Memperhatikan standar kelulusan yang ditentukan Badan Standar Nasional Pendidikan (BSNP) pada UN setiap tahunnya terus meningkat, hal ini dirasa cukup berat. Oleh karena itu, sekolah selaku pengemban tugas untuk mencerdaskan kehidupan bangsa melakukan berbagai upaya agar siswanya mampu mencapai kriteria kelulusan tersebut.

Upaya yang dilakukan sekolah dalam mempersiapkan siswa antara lain mengadakan bimbingan belajar, les, remedial, try out UN. Bahkan, membagikan standar kompetensi lulusan dan ruang lingkup materi yang diujikan kepada siswanya yang akan menghadapi UN. Salah satu upaya sekolah dalam mempersiapkan siswanya untuk mengikuti ujian nasional yaitu Tryout,

Tryout adalah tahap evaluasi untuk menghadapi Ujian Nasional sesungguhnya, Kesiapan dalam menghadapi ujian merupakan hal yang mutlak untuk dilakukan oleh siswa SMA

dan SMK kelas 3, oleh sebab itu perlu adanya sistem untuk membuat keputusan kesiapan siswa dan siswi dalam menghadapi Ujian Nasional. Try out UN setiap tahunnya telah dilakukan semua sekolah baik tingkat SMP/ sederajat maupun SMA/ sederajat. Try out yang dilakukan biasanya pada tingkat provinsi, kabupaten dan kota dilakukan satu kali hasilnya belum menggembirakan. Kemudian beberapa lembaga kursus dan bimbingan belajar juga tidak kalah geraknya melaksanakan try out bagi siswa-siswa yang mengejar pendidikan di lembaga yang bersangkutan,

Hasil try out UN, meskipun belum dapat dikatakan sebagai kemampuan maksimal siswa karena memang sebagian materi pelajaran yang belum diberikan, namun pada dasarnya sebagian besar soal yang di-try out-kan bersumber dari materi yang sudah disajikan. Penyusunan soal diambil dari materi pelajaran kelas I, II dan III, dan tidak menutup kemungkinan soal-soal diperoleh dari berbagai sumber media massa, dan materi tersebut berkait erat dengan kisi-kisi dari materi yang sudah diajarkan oleh guru di sekolah. Tentunya tidak terlepas dari kurikulum yang berlaku, akan tetapi mungkin ada hal-hal baru yang tidak pernah diajarkan oleh guru.

Paling tidak, hasil try out tahap pertama ini memberikan gambaran, sebagian besar siswa belum menyiapkan diri dengan belajar sungguh-sungguh. Tidak menutup kemungkinan, sebagian besar siswa terbiasa belajar santai sehingga tidak begitu siap menghadapi try out UN yang diselenggarakan dan hasilnya pun cukup mengawatirkan.

Maka dari itu, perlu dilaksanakan try out kali kedua, ketiga dan seterusnya. Karena semakin sering dilaksanakan try

out ini, maka akan memberikan dampak perubahan terhadap hasil-hasil selanjutnya, artinya setiap kali dilakukan try out nilai skor akan lebih baik dan meningkat.

Try out ini bisa dilakukan berkali-kali hanya dapat diselenggarakan di setiap sekolah, dan sekolah yang sering melaksanakan try out akan terlihat nyata nilai UN-nya meningkat. Dari data diatas penulis akan memfokuskan bagi para siswa yang mempunyai nilai terendah, supaya diwajibkan mengikuti kelas tambahan agar mereka lebih siap dalam menghadapi ujian nasional nanti, jika yang mempunyai nilai tertinggi tidak diwajibkan mengikuti kelas tambahan,

Disini penulis bekerja sama dengan SMK Negeri 1 Dukuturi untuk membuat sebuah sistem guna menentukan kesiapan Siswa dan Siswi dalam menghadapi ujian nasional, dalam hal ini atribut yang akan digunakan dalam sistem meliputi hasil dari tryout yang dilaksanakan oleh SMK Negeri 1 Dukuturi, dengan hasil tersebut nantinya akan melalui sebuah proses guna menentukan kesiapan siswa dan siswi dalam menghadapi ujian nasional

II. STUDI PUSTAKA

2.1. Penelitian Terkait

Penulis memulai penelitian ini dengan terlebih dahulu melakukan studi kepustakaan dari penelitian-penelitian dan sumber-sumber lain. Penelitian tersebut membahas tentang topik yang terkait dengan penelitian penulis, antara lain adalah penelitian mengenai algoritma yang akan digunakan penulis.

1. "Penerapan Algoritma Klasifikasi Data Mining ID3 untuk Menentukan Penjurusan Siswa SMA N 6 Semarang" yang ditulis oleh Obbie Kristanto.
2. "Perbandingan Algoritma ID3 dan C5.0 dalam Identifikasi Penjurusan Siswa SMA" yang ditulis oleh Holisatul Munawaroh, Bain Khusnul K, Yeni Kustiyahningsih.
3. "Penentuan Jurusan Sekolah Menengah Atas dengan Algoritma Fuzzy C-Means" yang ditulis oleh Bahar.
4. "Penerapan Algoritma Naive Bayes untuk Penentuan Status Turn-Over Pegawai" yang ditulis oleh Yeffriansjah Salim.
5. "Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa" yang ditulis oleh Arief Jananto.
6. "Penerapan Algoritma Naive Bayes untuk Mengklasifikasi Data Nasabah Asuransi" yang ditulis oleh Bustami.

2.2. Tinjauan Pustaka

A. Data Mining

Data mining (penambangan data) adalah suatu proses untuk menemukan suatu pengetahuan atau informasi yang berguna dari data berskala besar. Sering juga disebut sebagai bagian proses KDD (Knowledge Discovery in Databases). (Santosa, 2007) [19].

Karakteristik data mining sebagai berikut [21]:

- a. Data mining berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
- b. Data mining biasa menggunakan data yang sangat besar. Biasanya data yang besar digunakan untuk membuat hasil lebih percaya.
- c. Data mining berguna untuk membuat keputusan yang kritis, terutama dalam strategi..

B. Algoritma 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.

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

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[13].

Dimana :

X = data dengan class yang belum diketahui

H = hipotesis data X, merupakan suatu class spesifik

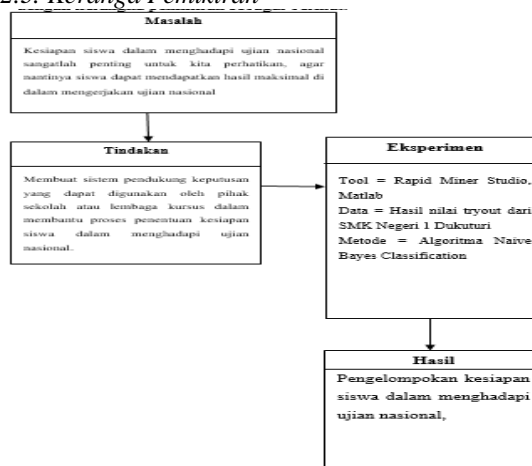
$P(H|X)$ = probabilitas hipotesis H berdasarkan kondisi X (posteriori probability)

$P(H)$ = probabilitas hipotesis H (prior probability)

$P(X|H)$ = probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$ =probabilitas dari X

2.3. Kerangka Pemikiran



III METODE PENELITIAN

3.1 Instrumen Penelitian

Berdasarkan permasalahan yang telah diuraikan pada bab sebelumnya, maka bahan dan peralatan yang diperlukan untuk penelitian ini meliputi :

A. Bahan

Dalam penelitian ini bahan yang dibutuhkan adalah data siswa kelas X SMA Kesatrian 1 Semarang tahun ajaran 2014/2015, yang beralamat di Jalan Pamularsih 116 Semarang.

B. Peralatan

Peralatan dalam penelitian ini meliputi kebutuhan perangkat lunak dan kebutuhan perangkat keras. Dibawah ini merupakan kebutuhan dari sistem, diantaranya:

Kebutuhan perangkat lunak :

- a. Microsoft Office Word Professional Plus 2010
Software ini digunakan untuk mengolah laporan hasil penelitian.
- b. Microsoft Office Excel Professional Plus 2010
Software ini digunakan sebagai media penulisan datasheet.
- c. Sistem operasi Microsoft Windows 8 Pro
Sistem Operasi yang digunakan dalam notebook penulis.
- d. Matlab version 7.10.0.499 (R2010a)
Framework yang akan digunakan untuk mengolah datasheet dalam klasifikasi penjurusan siswa menggunakan metode data mining.
- e. RapidMiner Studio 5.3.013
Framework yang akan digunakan untuk melihat hasil akurasi dari algoritma yang digunakan terhadap datasheet yang sedang diteliti.

Kebutuhan perangkat keras :

- a. Prosesor AMD E-450 APU with Radeon(tm) HD Graphics (2 CPUs), ~ 1.6GHz
- b. Layar monitor 12.1''
- c. Ram 2048MB
- d. Harddisk 500GB
- e. Satu buah mouse

3.2 Pengumpulan Data

Data yang diperoleh dalam penelitian ini merupakan data primer yaitu data yang didapatkan secara langsung dari sumber data, Selain itu dalam membantu penyusunan tugas akhir ini digunakan beberapa studi pustaka yang merupakan data

3.3 Teknik Analisis Data

Data siswa baru SMK negeri 1 dukuturi tahun ajaran 2014/2015 diperoleh dari pengumpulan data sebanyak 430 record, namun tidak semua data dapat digunakan dan tidak semua atribut digunakan karena harus melalui beberapa tahap pengolahan awal data (preparation data). Untuk mendapatkan data yang berkualitas, beberapa teknik yang dilakukan sebagai berikut:

- 1.Data cleaning.
- 2.Data integration and transformation
- 3.Data reduction

Tabel 3.1 Atribut

No	Nama Atribut	Penjelasan
1	Nama	Nama Siswa
2	Nilai Bahasa Indonesia	Nilai Bahasa Indonesia
3	Nilai Bahasa Inggris	Nilai Bahasa Inggris
4	Nilai Matematika	Nilai Matematika
5	Nilai IPA	Nilai IPA
6	Minat	Minat jurusan
7	Jurusan	Jurusan siswa

Tabel 3.2 Transformation data

Nilai	Predikat
3,33 – 4	1
2,33 – 3,32	2
1,33 – 2,32	3
0 – 1,32	4

A	B	C	D
NAMA	NEM BI	Skala 1-4	Predikat
ALDO BANIDA RAFID	5,6	2,24	3
BAIHAQY HADI T	6,6	2,64	2
MAHARNUM PRAMITYA LILIMADANI	7,0	2,8	2
DEA NABILA RAHMADIKA	7,8	3,12	2
SOFIE ANTONIA HANJANI	7,0	2,8	2
YULIANTINA BUNGA KINANTI	5,6	2,24	3
IRFAN RIFALDI	7,6	3,04	2
MOCH. GALUH RAGA PRAMESWARA	7,2	2,88	2
MUHAMMAD HUTOMO ADI N.	5,2	2,08	3
ZAHRA PUTSUMAZKI	6,8	2,72	2

Gambar 3.1 Konversi nilai ujian ke angka predikat

IV HASIL PENELITIAN DAN PEMBAHASAN

Didalam bab ini akan dibahas mengenai data yang akan digunakan dalam penelitian, data tersebut akan di hitung menggunakan algoritma naïve bayes dan kemudian diuji menggunakan cross validation.

4.1 Data yang digunakan

Pada penelitian ini, data yang digunakan adalah data siswa baru SMK Negeri 1 DukuTuri tahun ajaran 2014/2015 dengan jumlah 366 record data.

	A	B	C	D	E	F	G
	NAMA	N_IND	N_ING	N_MAT	N_IPA	MINAT	Status
2	ALDO BANIDA RAFID	3	2	3	3	1	Siap
3	BAlHAQY HADI T	2	3	3	3	1	TIDAK
4	MAHARNUM PRAMITYA LILIMADANI	2	2	2	2	1	TIDAK
5	DEA NABILA RAHMADIKA	2	1	2	3	1	TIDAK
6	SOFIE ANTANIA HANJANI	2	2	3	3	3	Siap
7	YULIANTINA BUNGA KINANTI	3	2	3	2	1	TIDAK
8	IRFAN RIFALDI	2	3	3	3	1	TIDAK
9	MOCH. GALUH RAGA PRAMESWARA	2	2	3	2	1	TIDAK
10	MUHAMMAD HUTOMO ADI N.	3	3	3	3	2	TIDAK
11	ZAHRA PUTSUMAZKI	2	2	2	2	1	TIDAK
12	WILDAN PRAKOSO	2	2	3	2	2	TIDAK
13	ANDHIKA ARYA PERDANA SUYANA PU	2	2	2	2	1	KHUSUS
14	AULIA KEMAL MUHAMMAD	2	2	3	2	1	TIDAK
15	IKA RISKI LESTARI	2	3	3	3	1	TIDAK
16	VERA FEBRIYANTI	2	3	3	2	1	TIDAK
17	SHEILA ZALFATIKA	1	2	3	2	1	TIDAK
18	DEBBY MAYLINDA VIRTRARIES	2	3	4	3	2	TIDAK
19	ALMIRA JOVANKOVA YUNAN	1	2	2	2	1	Siap
20	MOCHAMAD RAFLI RAMADHAN	2	2	1	1	1	KHUSUS
21	DIMASTA WARDHANA	2	2	3	2	1	TIDAK

Gambar 4. 1 Tabel Data Siswa yang telah dikonversi

4.2 Perhitungan Algoritma Naïve Bayes

Dibawah ini adalah contoh perhitungan manual penerapan algoritma naïve bayes untuk menentukan kesiapan siswa dalam menghadapi ujian nasional menggunakan data training pada gambar 4.1 dan data testing sebagai berikut:

Tabel 4.1 Perhitungan Algoritma Naïve Bayes

Nama	N_Ind	N_Ing	N_Mat	N_Ipa	Minat	status
A	2	2	1	2	1	?

Dalam perhitungan algoritma naïve bayes ada beberapa langkah yang harus dilakukan, berikut adalah langkah langkahnya

- a. Tahap pertama yang dilakukan adalah menghitung jumlah *class/label* :

$$P(Y=KHUSUS) = 2/20$$

$$P(Y=TIDAK) = 15/20$$

$$P(Y=SIAP) = 3/20$$

- a. Tahap kedua menghitung jumlah kasus yang sama dengan *class* yang sama :

$$P(N_IND=2|Y=KHUSUS) = 2/2$$

$$P(N_IND=2|Y=TIDAK) = 12/15$$

$$P(N_IND=2|Y=SIAP) = 1/3$$

$$P(N_ING=2|Y=KHUSUS) = 2/2$$

$$P(N_ING=2|Y=TIDAK) = 8/15$$

$$P(N_ING=2|Y=SIAP) = 3/3$$

$$P(N_MAT=1|Y=KHUSUS) = 1/2$$

$$P(N_MAT=1|Y=TIDAK) = 0/15$$

$$P(N_MAT=1|Y=SIAP) = 0/3$$

$$P(N_IPA=1|Y=KHUSUS) = 1/2$$

$$P(N_IPA=1|Y=TIDAK) = 0/15$$

$$Accuracy = \frac{Count11+Count22+Count33}{Count11+Count12+Count13+Count21+Count22+Count23+Count31+Count32+Count33}$$

$$P(N_IPA=1|Y=SIAP) = 0/3$$

$$P(MINAT=1|Y=KHUSUS) = 2/2$$

$$P(MINAT=1|Y=TIDAK) = 12/15$$

$$P(MINAT=1|Y=SIAP) = 2/3$$

- b. Tahap ketiga kalikan semua hasil variable KHUSUS, TIDAK dan Siap, KHUSUS

$$= \{P(P(N_IND=2|Y=KHUSUS)) \cdot$$

$$P(N_ING=2|Y=KHUSUS) \cdot$$

$$P(N_MAT=1|Y=KHUSUS) \cdot$$

$$P(N_IPA=1|Y=KHUSUS) \cdot$$

$$P(MINAT=1|Y=KHUSUS)$$

$$= 2/2 \cdot 2/2 \cdot 1/2 \cdot 1/2 \cdot 2/2$$

$$= 1 \cdot 1 \cdot 0,5 \cdot 0,5 \cdot 1$$

$$= 0,25$$

IPS

$$= \{P(P(N_IND=2|Y=TIDAK)) \cdot$$

$$P(N_ING=2|Y=TIDAK) \cdot P(N_MAT=1|Y=TIDAK) \cdot$$

$$P(N_IPA=1|Y=TIDAK) \cdot P(MINAT=1|Y=TIDAK)$$

$$= 12/15 \cdot 8/15 \cdot 0/15 \cdot 0/15 \cdot 12/15$$

$$= 0,8 \cdot 0,53 \cdot 0 \cdot 0 \cdot 0,8$$

$$= 0$$

BAHASA

$$= \{P(P(N_IND=2|Y=SIAP)) \cdot P(N_ING=2|Y=SIAP) \cdot P(N_MAT=1|Y=SIAP) \cdot P(N_IPA=1|Y=SIAP) \cdot P(MINAT=1|Y=SIAP)$$

$$= 1/3 \cdot 3/3 \cdot 0/3 \cdot 0/3 \cdot 2/3$$

$$= 0,3 \cdot 1 \cdot 0 \cdot 0 \cdot 0,6$$

$$= 0$$

- c. Tahap keempat bandingkan hasil *class* KHUSUS, TIDAK, dan Siap,

Karena Hasil (P|KHUSUS) lebih besar dari (P|TIDAK) dan (P|Siap) maka keputusannya adalah KHUSUS, 0,25 > 0 maka "KHUSUS"

A	2	2	1	2	1	KHUSUS
----------	----------	----------	----------	----------	----------	---------------

Multiclass Classification Performance Annotations

Table View Plot View

accuracy: 83.86% +/- 4.06% (mikro: 83.88%)

	true Bahasa	true IPS	true IPA	class precision
pred. Bahasa	24	0	0	100.00%
pred. IPS	16	263	33	84.29%
pred. IPA	0	10	20	66.67%
class recall	60.00%	96.34%	37.74%	

4.3 Pengujian Cross Validation

Gambar 4.2 Nilai akurasi dari RapidMiner

Dari *confusion matrix* di atas, terlihat bahwa akurasi dengan menggunakan algoritma naïve bayes adalah sebesar 83,87%. Berikut adalah perhitungan akurasi dan Error Rate.

$$\text{Accuracy} = \frac{24+263+20}{24+0+0+16+263+33+0+10+20}$$

$$\text{Accuracy} = \frac{307}{366}$$

$$\text{Accuracy} = 0,8387978$$

$$\text{Error Rate} = \frac{\text{Count12}+\text{Count13}+\text{Count21}+\text{Count23}+\text{Count31}+\text{Count32}}{\text{Count11}+\text{Count12}+\text{Count13}+\text{Count21}+\text{Count22}+\text{Count23}+\text{Count31}+\text{Count32}+\text{Count33}}$$

$$\text{Error Rate} = \frac{0+0+16+33+0+10}{24+0+0+16+263+33+0+10+20}$$

$$\text{Error Rate} = \frac{59}{366}$$

$$\text{Error Rate} = 0,1612021$$

4.4 Pengujian pada Command Window Matlab

```
>> tepat = ((a+e+i) / (a+b+c+d+e+f+g+h+i)) * 100

tepat =

    83.8798

>> error = ((b+c+d+f+g+h) / (a+b+c+d+e+f+g+h+i)) * 100

error =

    16.1202

fx >>
```

4.5 Hasil Akurasi dari pengujian Matlab dan Rapidminer

	Naive Bayes	
	Matlab	RapidMiner
Accuracy	83,8798 %	83,87 %
Error	16,1202 %	16,12 %

Tabel 4.2 : Hasil Pengujian Matlab dan RapidMiner

Pada pengujian menggunakan tools Matlab dan RapidMiner, keduanya memiliki tingkat akurasi dan kesalahan yang sama. Hal ini menunjukkan bahwa penggunaan metode algoritma *naive bayes classifier* bisa digunakan untuk memprediksi penjurusan siswa di sekolah menengah atas.

V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil penelitian yang telah dilakukan oleh peneliti, dapat disimpulkan bahwa klasifikasi data siswa baru di SMK Negeri 1 Dukuturi tahun ajaran 2014/2015 dapat diselesaikan menggunakan teknik data mining, yaitu dengan menggunakan algoritma *Naive Bayes Classifier* untuk membantu pihak sekolah menengah atas dalam proses penentuan jurusan siswa.

Dari hasil percobaan prediksi kesiapan siswa menggunakan matlab dengan metode *Naive Bayes*, diperoleh akurasi sebesar 83,8798 % dengan error rate sebesar 16,1202 %.

5.2 Saran

Agar penelitian ini terus berkembang, peneliti memberikan saran sebagai berikut :

1. Penelitian ini dapat dikembangkan dengan metode data mining klasifikasi lainnya untuk melakukan perbandingan.
2. Penelitian ini dapat dikembangkan dengan menambah algoritma optimasi atau algoritma fitur seleksi untuk memilih variable yang mempengaruhi hasil prediksi.

UCAPAN TERIMA KASIH

Penulis mengucapkan teima kasih kepada Universitas Dian Nuswantoro, Rektor UDINUS, Dekan Fakultas Ilmu Komputer, Kaprodi Teknik Informatika-S1, Dosen pembimbing, Dosen-dosen pengampu kuliah di Fakultas Ilmu Komputer, serta teman-teman dan sahabat yang selama ini telah mendampingi penulis selama kuliah di Universitas Dian Nuswantoro.

REFERENSI

- [1] Holisatul Munawaroh, Bain Khusnul, and Yeni Kustiyahningsih, "Perbandingan Algoritma ID3 dan C5.0 dalam Identifikasi Penjurusan Siswa SMA," *Jurnal Sarjana Teknik Informatika*, vol. 1, pp. 1-12, Juni 2013.
- [2] Departemen Pendidikan Nasional, *Panduan Penyusunan Laporan Hasil Belajar Peserta Didik Sekolah Menengah Atas (SMA)*. Jakarta, Indonesia, 2006.
- [3] Fauzan A. Maharani. (2014, Juli) <http://www.m-edukasi.web.id>. [Online]. <http://www.m-edukasi.web.id/2014/07/peningkatan-mutu-pendidikan-smasmk.html?m=1>
- [4] Tri Mucharam Tjandra, "Penerapan Kurikulum 2013 dan Penjurusan Siswa," SMA Kesatrian 1 Semarang, Semarang, Interview 2014.
- [5] Kementerian Pendidikan dan Kebudayaan, "Permendikbud Nomor 81A tahun 2013 tentang Implementasi Kurikulum," Kementerian Pendidikan dan Kebudayaan, 2013.
- [6] Witten Ian H., Eibe Frank, and Hall Mark A., *Data Mining : Practical Machine Learning Tools and Techniques 3rd Edition*. Elsevier, 2011.
- [7] Obbie Kristanto, "Penerapan Algoritma Klasifikasi Data Mining ID3 untuk Menentukan Penjurusan Siswa SMAN 6 Semarang," 2014.
- [8] Selvia Lorena Br Ginting and Reggy Pasya Trinanda, *Penggunaan Metode Naive Bayes Classification pada Aplikasi Perpustakaan*. Bandung, Indonesia, 2010.
- [9] I. Rish, "An Empirical Study of the Naive Bayes Classifier," in *International Joint Conference on Artificial Intelligence*, California, 2006.
- [10] Bahar, "Penentuan Jurusan Sekolah Menengah Atas dengan Algoritma Fuzzy C Means," 2011.
- [11] Yeffriansjah Salim, "Penerapan Algoritma Naive Bayes untuk Penentuan Status Turn-Over Pegawai," *Media Sains*, vol. 4 Nomor 2, no. ISSN 2085-3548, Oktober 2012.
- [12] Arif Jananto, "Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa," *Jurnal Teknologi Informasi DINAMIK*, vol. 18 No.1, no. ISSN : 0854-9524, pp. 09-16, Januari 2013.
- [13] Bustami, "Penerapan Algoritma Naive Bayes untuk Mengklasifikasi Data Nasabah Asuransi," *Jurnal Informatika*, vol. 8, Januari 2014.
- [14] Usama Fayyad, *Advances in Knowledge Discovery and Data Mining*.: MIT Press, 1996.
- [15] Kusri and E. T. Luthfi, *Algoritma Data Mining*. Yogyakarta, Indonesia: Andi Offset, 2009.
- [16] Daniel T Larose, *Discovering Knowledge in Data : An Introduction to Data Mining*.: Wiley Interscience, 2005.
- [17] C. Dennis dkk Aprilla, *Belajar Data Mining dengan RapidMiner*. Jakarta, Indonesia, 2013.
- [18] Andry Pujiriyanto, *Cepat Mahir Matlab*. Jakarta, Indonesia, 2004.
- [19] Budi Santoso, *Data Mining, Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta, Indonesia: Graha Ilmu, 2007.
- [20] J. Han and M. Kamber, *Data Mining Concept and Techniques Second Edition Morgan Kauffman*., 2006.
- [21] P. Chapman, J. Clinton, R. Kerber, and dkk, *CRIS-DM 1.0 : Step by Step Data Mining Guide*., 2000.
- [22] Eko Prasetyo, *Data Mining : Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta, Indonesia: C.V Andi Offset, 2012.
- [23] "Peraturan Menteri Pendidikan dan Kebudayaan No. 66 tahun 2014 tentang Standar Penilaian Pendidikan," Kemetrian Pendidikan dan Kebudayaan, 2013.
- [24] "Peraturan Menteri Pendidikan Nasional Nomor 19 Tahun 2007," Menteri Pendidikan Nasional, 2007.