

PENERAPAN DATA MINING UNTUK KLASIFIKASI PENJURUSAN SEKOLAH MENENGAH ATAS PADA SMA 1 KAJEN DENGAN MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER

Maghriza Fakri Adilla¹, T. Sutojo²

^{1,2} Jurusan Teknik Informatika-S1, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

Jln. Nakula 1 No 5-11 Semarang 501131, (024) 3517261

E-mail : 111201207169@mhs.dinus.ac.id¹, tsutojo@dosen.dinus.ac.id²

Abstrak

Penjurusan siswa adalah suatu proses pengambilan keputusan dalam memilih bidang keahlian studi berdasarkan kemampuan potensi diri dan peluang yang ada. Secara formal pemilihan jurusan merupakan ketentuan yang sudah ditetapkan oleh pemerintah melalui kurikulum yang berlaku, yaitu kurikulum 2013. Penjurusan dilakukan dikelas X, dampaknya adalah pihak sekolah khususnya guru bimbingan karir belum mengetahui bakat dan karakter siswa dalam program studi tertentu. Sehingga dikhawatirkan siswa akan mengalami kesulitan dalam mengikuti pembelajaran yang menyebabkan rendahnya prestasi belajar siswa. Dalam menentukan penjurusan jika dilakukan dengan cara manual maka akan menyita banyak waktu dan memerlukan ketelitian ekstra karena data yang cukup banyak sehingga memungkinkan terjadi kesalahan dalam melakukan proses penjurusan. Oleh karena itu diperlukan teknik Data Mining dengan menggunakan metode klasifikasi algoritma Naive Bayes untuk mengklasifikasi penjurusan siswa di SMA 1 Kajen. Algoritma Naive Bayes merupakan metode machine learning yang menggunakan perhitungan probabilitas yang digunakan untuk menghitung peluang dan dapat memprediksi peluang dimasa depan berdasarkan pengetahuan dimasa sebelumnya. Dengan menerapkan algoritma ini diharapkan dapat membantu untuk mengklasifikasi penjurusan siswa secara tepat dan akurat. Dari hasil penelitian, disimpulkan bahwa klasifikasi data siswa baru SMA 1 Kajen tahun ajaran 2015/2016 dapat diselesaikan dengan menggunakan teknik data mining, yaitu dengan metode algoritma Naive Bayes. Akurasi yang dihasilkan dari metode algoritma Naive Bayes Classifier adalah sebesar 86,1842 % dengan error rate sebesar 13.8158 %.

Kata Kunci : Penjurusan, Kurikulum 2013, Data Mining, Klasifikasi, *Naive Bayes Classifier*

I. PENDAHULUAN

Penjurusan siswa adalah suatu proses pengambilan keputusan dalam memilih bidang keahlian studi berdasarkan kemampuan potensi diri dan peluang yang ada. Secara formal pemilihan jurusan merupakan ketentuan yang sudah ditetapkan oleh pemerintah melalui kurikulum 2013 yang berlaku saat ini. Untuk menentukan penjurusan siswa dilakukan pada awal masuk, yaitu pada kelas X Sekolah Menengah Atas. Penerapan kurikulum 2013 ini bertujuan untuk penyesuaian program pendidikan pada satuan pendidikan dengan kondisi dan kekhasan potensi yang ada didaerah siswa [1]. Dampak diberlakukannya kurikulum 2013 adalah pihak sekolah khususnya guru BK belum mengetahui bakat dan karakter siswa dalam program studi tertentu. Sehingga dikhawatirkan siswa akan mengalami kesulitan dalam mengikuti pembelajaran yang menyebabkan rendahnya prestasi belajar siswa.

SMA 1 Kajen merupakan Sekolah Menengah Atas yang memiliki dua program penjurusan, yaitu jurusan Ilmu Pengetahuan Alam (IPA) dan Ilmu Pengetahuan Sosial (IPS). Penjurusan di SMA 1 Kajen dilakukan agar para siswa bisa menyalurkan bakat dan minat serta kemampuan di jurusan yang ada. Sehingga siswa dapat mengembangkan potensi yang dimilikinya. Pemilihan jurusan yang tepat akan menentukan keberhasilan siswa. Begitupun sebaliknya, jika tidak tepat

dalam pemilihan jurusan kesempatan yang sangat baik bagi siswa akan hilang. Maka sekolah harus mampu memfasilitasi siswa agar sesuai dengan kemampuan yang dimiliki sehingga siswa dijuruskan pada jurusan yang tepat. Penjurusan ini sangat membantu siswa dalam mempersiapkan diri memasuki perguruan tinggi, dimana pada saat itu siswa harus memiliki spesialisasi program studi tertentu. Sering kali siswa asal dalam memilih jurusan sehingga nantinya setelah lulus SMA siswa akan kebingungan memilih jurusan apa diperguruan tinggi sehingga dapat menghambat cita-citanya. Penjurusan di SMA 1 Kajen, baik IPA dan IPS dipertimbangkan oleh beberapa faktor, yaitu nilai Ujian Nasional IPA dan Matematika Sekolah Menengah Pertama (SMP), nilai raport IPA dan Matematika Sekolah Menengah Pertama (SMP) selama 5 semester, nilai kualitas, nilai IQ, dan minat jurusan [2]. Dalam menentukan penjurusan dengan memperhatikan banyak faktor yang kompleks, jika dilakukan dengan cara manual maka akan menyita banyak waktu dan memerlukan ketelitian ekstra karena data yang cukup banyak dan mempunyai banyak kelemahan sehingga memungkinkan terjadi kesalahan dalam melakukan proses penjurusan [3]. Oleh karena itu diperlukan klasifikasi penjurusan yang tepat dan akurat, salah satunya menggunakan teknologi dibidang data mining.

Data mining merupakan disiplin ilmu yang mempelajari metode untuk mengekstrak pengetahuan atau pola dari suatu

data, sering disebut juga knowledge discovery in database [4]. Data mining dapat digunakan untuk mengelompokkan, memprediksi, mengestimasi, dan menentukan kaidah asosiasi dalam suatu data yang ada. Perlunya data mining karena adanya sejumlah data besar yang dapat digunakan sehingga menghasilkan informasi dan knowledge yang berguna [5]. Salah satu metode data mining adalah algoritma *Naive Bayes*. *Naive Bayes* merupakan metode machine learning yang menggunakan perhitungan probabilitas yang digunakan untuk menghitung peluang dan dapat memprediksi peluang dimasa depan berdasarkan pengetahuan dimasa sebelumnya. *Naive Bayes* adalah salah satu algoritma yang terdapat pada teknik klasifikasi. *Bayes Optimal Classifier* dapat menghitung peluang dari satu kelas dari masing-masing kelompok atribut dan dapat menentukan kelas mana yang paling optimal [6].

Berdasarkan uraian masalah diatas, maka akan dibuat sistem untuk mengklasifikasi penjurusan siswa pada SMA 1 Kajen. Sistem akan dibuat menggunakan teknologi data mining dengan menggunakan algoritma *Naive Bayes*. Diharapkan algoritma *Naive Bayes* tersebut mampu menjadi sistem pendukung keputusan dalam menentukan penjurusan siswa di SMA 1 Kajen.

II. STUDI PUSTAKA

2.1 Penelitian Terkait

Sebelum melakukan penelitian penulis terlebih dahulu melakukan tinjauan pustaka dari penelitian lain yang membahas permasalahan yang hampir sama, antara lain :

1. Algoritma Klasifikasi *Naive Bayes* Untuk Menilai Kelayakan Kredit yang ditulis oleh Claudia Clarentia Ciptohartono.
2. Algoritma *Naive Bayes* untuk Mencari Perkiraan Waktu Studi Mahasiswa yang ditulis oleh Algoritma *Naive Bayes*.
3. Data Mining Menggunakan Algoritma *Naive bayes* Untuk Klasifikasi Kelulusan Mahasiswa Universitas Dian Nuswantoro yang ditulis oleh Yuda Septian Nugroho.
4. Penerapan Algoritma C4.5 Untuk Prediksi Jurusan Siswa SMAN 3 Rembang yang ditulis oleh Faid Ari Prastya.

2.2 Tinjauan Pustaka

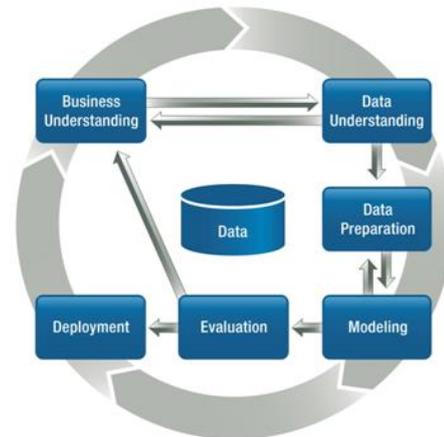
A. Data Mining

Data mining merupakan proses penggunaan teknik statistik, matematika, *artificial intelligence* (kecerdasan buatan) dan *machine learning* yang digunakan untuk mengekstrak serta mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [12].

B. CRISP-DM

CRISP-DM (*Cross Industry Standard Process for Data Mining*) adalah suatu konsorsium perusahaan yang didirikan oleh Komisi Eropa pada tahun 1996 dan sudah ditetapkan sebagai proses standar dalam data mining yang bisa diaplikasikan diberbagai sektor industri. Berikut

merupakan gambar dari proses siklus hidup pengembangan dari CRISP-DM [12] :



Gambar 2.1 : Siklus Hidup CRISP-DM

Berikut merupakan tahap dari siklus pengembangan data mining CRISP-DM :

- a. Pemahaman Bisnis (*Business Understanding*)
- b. Pemahaman Data (*Data Understanding*)
- c. Pengolahan Data (*Data Preparation*)
- d. Pemodelan (*Modelling*)
- e. Evaluasi (*Evaluation*)
- f. Penyebaran (*Deployment*)

C. Algoritma Naive Bayes

Algoritma *Naive Bayes* adalah salah satu algoritma yang terdapat pada teknik data mining klasifikasi. *Naive bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris yaitu Thomas bayes, *Naive Bayes* memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya, sehingga dikenal dengan 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 [15].

Persamaan dari teorema Bayes adalah :

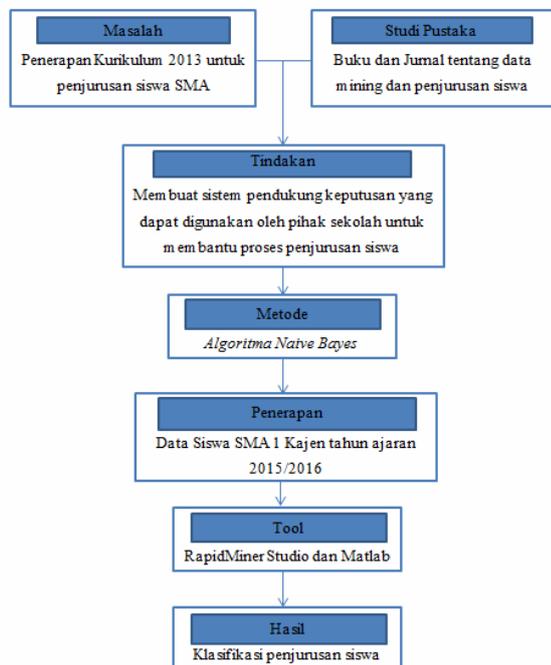
$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

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

2.3 Kerangka Pemikiran



Gambar 2.2 : Kerangka Pemikiran.

III METODE PENELITIAN

3.1 Instrumen Penelitian

Pada penelitian ini bahan dan peralatan yang diperlukan sebagai berikut :

A. Bahan

Dalam penelitian ini bahan yang dibutuhkan adalah data siswa kelas X SMA 1 Kajen tahun ajaran 2015/2016, yang beralamat di Jalan Mandurorejo Kajen, Kabupaten Pekalongan.

B. Peralatan

Peralatan yang digunakan dalam melakukan penelitian berupa kebutuhan perangkat keras dan perangkat lunak. Berikut adalah kebutuhan sistem :

Kebutuhan perangkat keras :

- Prosesor AMD A6-4455 APU with Radeon(tm) HD Graphics 2.10 GHz
- Layar monitor 14 inch
- Ram 4,00 GB
- Harddisk 500 GB

Kebutuhan perangkat lunak :

1. Sistem

- a. Windows Edition : Microsoft Windows 7 Ultimate
- b. System Type : 64-bit Operating System

2. Data Mining

- a. Matlab version R2012a (7.14.0.739)
Tools yang digunakan untuk mengolah data dalam klasifikasi penjurusan siswa SMA 1 Kajen dengan menggunakan metode data mining.
- b. RapidMiner 5.3.013
Tools yang digunakan untuk melihat tingkat keakurasian data yang diteliti dengan menggunakan algoritma *Naive Bayes*.

3.2 Metode Penelitian

Metode yang digunakan pada penelitian ini adalah model standarisasi data mining, yaitu CRISP-DM (*Cross Industry Standart Process for Data Mining*). Berikut adalah langkah-langkah metode CRISP-DM :

a. Pemahaman Bisnis

Penelitian ini mengutamakan penentuan tujuan proyek dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan, serta menerjemahkan tujuan dan batasan dari data yang diperoleh dari SMA 1 Kajen sehingga menjadi formula dari permasalahan data mining mulai dari menyiapkan strategi awal hingga metode yang dibutuhkan untuk menapai tujuan.

b. Pemahaman Data

Pada penelitian ini data yang digunakan adalah data primer. Data diperoleh melalui *softcopy* secara langsung dari SMA 1 Kajen. Data yang didapatkan adalah data siswa kelas X SMA 1 Kajen tahun ajaran 2015/2016.

c. Pengolahan Data

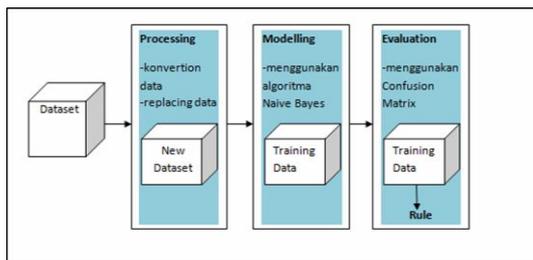
- Tahap Pertama yang dilakukan adalah penentuan data yang akan diolah. Dari data yang telah diperoleh, tidak semua data akan diolah karena tidak semua masuk dalam kriteria penelitian, penelitian yang akan dilakukan memiliki batasan-batasan data yang akan digunakan. Data awal terdiri dari 332 *record* data.
- Tahap Kedua yang dilakukan adalah menentukan atribut yang akan digunakan dari tahap pertama. Terdapat 15 variabel pada data awal, variabel yang akan digunakan sebanyak 9 variabel adalah nama siswa, rata-rata nilai raport ipa, rata-rata nilai raport matematika, nilai ujian nasional ipa, nilai ujian nasional matematika, nilai kualitas, tes IQ, minat jurusan dan jurusan.
- Tahap Ketiga yang dilakukan adalah penanganan data *missing value*. *Missing*

value adalah data yang tidak lengkap dikarenakan atribut tidak tercatat maupun atribut memang tidak dimiliki. Penanganan *missing value* dilakukan dengan penghapusan *record* yang kosong. Jumlah data awal adalah 332 data, terdapat data yang missing value sehingga menjadi 304 data *record* yang dapat digunakan.

- Tahap Keempat atau yang terakhir adalah melakukan konversi data. Guna memudahkan dalam analisa proses data mining maka data dengan atribut yang telah ditentukan atau dipilih akan dilakukan konversi. Konversi dilakukan pada atribut rata-rata nilai raport ipa, rata-rata nilai raport matematika, nilai ujian nasional ipa, nilai ujian nasional matematika, nilai kualitas, nilai IQ dan minat jurusan.

d. Pemodelan

Dalam penelitian ini metode yang digunakan adalah algoritma *Naive Bayes Classification*. Untuk pengukuran tingkat akurasi klasifikasi penjurusan siswa ini menggunakan *Tools RapidMiner* dan *Matlab*. Berikut adalah gambaran pemodelan penelitian :



Gambar 3.1 : Model Penelitian

e. Validasi dan Evaluasi

Pada tahap ini dilakukan validasi dan pengukuran tingkat keakuratan dari hasil yang dicapai oleh model dengan menggunakan *tools Matlab* yaitu *Confusion Matrix* untuk menguji tingkat keakuratan model yang digunakan.

f. Penyebaran

Hasil dari penelitian ini berupa analisis yang mengarah ke *Decision Support System (DSS)* yang diharapkan dapat digunakan oleh pihak sekolah SMA 1 Kajen khususnya guru BK sebagai bahan pertimbangan untuk proses penjurusan siswa serta menentukan langkah penjurusan siswa dengan atribut yang telah dijabarkan.

Pada penelitian ini, data yang digunakan adalah data siswa baru SMA 1 Kajen tahun ajaran 2015/2016 dengan jumlah 304 record.

	A	B	C	D	E	F	G	H	I
1	NAMA	RAPORTIP	RAPORTMT	N_UNIPA	N_UNMTK	ILAI KUALITA	IQ	MINAT	JURUSAN
2	ABDUL MUHAM	2	2	2	2	2	4	1	IPS
3	ADHE ASSHIDIQIE	2	2	3	3	3	4	2	IPS
4	ADHITYA SETYONUGROHO	2	2	1	2	3	4	1	IPA
5	ADITYO NUGROHO	2	1	2	2	3	4	1	IPA
6	AFIF ROSYIDI	2	2	2	2	3	4	2	IPS
7	AGNIS WIDIA RARASATI	2	2	2	3	3	4	2	IPS
8	AGUNG PRIYO UTOMO	1	1	2	2	3	4	2	IPS
9	AHMAD AZRIEL DAFFARISANDI	1	1	1	2	2	4	1	IPA
10	AIN HERLANGBANG	2	2	3	2	3	4	1	IPS
11	AINUN KHOLIZA	1	1	1	2	3	4	1	IPA
12	AKHMAD SOFYAN KHALIM	2	2	2	2	3	4	1	IPA
13	AKMAL WISNU YUWANA	2	2	2	3	3	4	2	IPS
14	ALDHA NAVA PRATIWI	2	2	3	3	3	4	2	IPS
15	ALDI RIZAL ZAENI	2	2	2	2	3	4	1	IPA
16	ALIEF PRIAMBODO	2	2	3	2	3	4	1	IPS
17	ALIFIA FADHILATUL ARDHIAN	1	1	2	2	2	4	1	IPS
18	AMALIA FADHILLA KARESYA PUTRI	2	2	2	3	3	4	2	IPS
19	AMALIATUL KHANIFA	1	2	2	2	2	4	1	IPA
20	AMLATUL KHOIRIYAH	1	1	2	2	2	4	1	IPA
21	AMRUL IKROM	1	1	1	1	2	3	1	IPA

Gambar 4. 1 Tabel Data Siswa

4.2 Perhitungan Algoritma Naive Bayes

Berikut merupakan perhitungan manual dari penerapan algoritma *Naive Bayes* untuk mengklasifikasi penjurusan siswa dengan menggunakan data *training* pada gambar 4.1 dan data testing sebagai berikut

Tabel 4.1 Data Testing

Nama	R_ipa	R_mtk	Un_ipa	Un_mtk	Qty	Iq	Mnt	Jrsn
A	1	2	2	1	2	3	1	?

Berikut adalah langkah-langkah perhitungan *Naive Bayes* :

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

$$P(Y=IPA) = 9/20$$

$$P(Y=IPS) = 11/20$$

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

$$P(R_RAPORTIPA=1|Y=IPA)=5/9$$

$$P(R_RAPORTIPA=1|Y=IPS)=2/11$$

$$P(R_RAPORTMTK=2|Y=IPA)=4/9$$

$$P(R_RAPORTMTK=2|Y=IPS)=9/11$$

$$P(N_UNIPA=2|Y=IPA)=5/9$$

$$P(N_UNIPA=2|Y=IPS)=7/11$$

$$P(N_UNMTK=1|Y=IPA)=1/9$$

$$P(N_UNMTK=1|Y=IPS)=0/11$$

$$P(N_KUALITAS=2|Y=IPA)=4/9$$

$$P(N_KUALITAS=2|Y=IPS)=2/11$$

IV HASIL PENELITIAN DAN PEMBAHASAN

4.1 Persiapan Data

$$P(IQ=3|Y=IPA)=1/9$$

$$P(IQ=3|Y=IPS)=0/11$$

$$P(MINAT=1|Y=IPA)=9/9$$

$$P(MINAT=1|Y=IPS)=4/11$$

- c. Tahap ketiga kalikan semua hasil variable IPA dan IPS

$$IPA = P(R_RAPORTIPA=1|Y=IPA).$$

$$P(R_RAPORTMTK=2|Y=IPA).$$

$$P(N_UNIPA=2|Y=IPA). P(N_UNMTK=1|Y=IPA).$$

$$P(N_KUALITAS=2|Y=IPA). P(IQ=3|Y=IPA).$$

$$P(MINAT=1|Y=IPA). P(Y=IPA)$$

$$= 5/9 \cdot 4/9 \cdot 5/9 \cdot 1/9 \cdot 4/9 \cdot 1/9 \cdot 9/9 \cdot 9/20$$

$$= 0,56 \cdot 0,44 \cdot 0,56 \cdot 0,11 \cdot 0,44 \cdot 0,56 \cdot 1 \cdot 0,45$$

$$= 0,00168$$

$$IPS = P(R_RAPORTIPS=1|Y=IPS).$$

$$P(R_RAPORTMTK=2|Y=IPS).$$

$$P(N_UNIPA=2|Y=IPS). P(N_UNMTK=1|Y=IPS).$$

$$P(N_KUALITAS=2|Y=IPS). P(IQ=3|Y=IPS).$$

$$P(MINAT=1|Y=IPS). P(Y=IPS)$$

$$= 2/11 \cdot 9/11 \cdot 7/11 \cdot 0/11 \cdot 2/11 \cdot 0/11 \cdot 4/11 \cdot 11/20$$

$$= 0,18 \cdot 0,81 \cdot 0,63 \cdot 0 \cdot 0,18 \cdot 0 \cdot 0,36 \cdot 0,55$$

$$= 0$$

- d. Tahap keempat bandingkan hasil *class* IPA dan IPS. Karena hasil (P|IPA) lebih besar dari (P|IPS) maka keputusannya adalah IPA.
 $0,00168 > 0$ maka “IPA”

4.3 Pengujian Menggunakan Matlab

Berdasarkan data training sebanyak 304 data dan 8 variabel diantaranya adalah rata-rata nilai raport ipa, rata-rata nilai raport matematika, nilai ujian nasional ipa, nilai ujian nasional matematika, nilai kualitas, tes IQ, minat jurusan dan jurusan, yang dimodelkan dengan algoritma *Naive Bayes* diperoleh hasil sebagai berikut :

```

Command Window
>> conmat=confusionmat(var_target,tar)

conmat =

    110    23
     19   152

>> TP=conmat(1,1)
TP =
    110

>> FN=conmat(1,2)
FN =
     23

>> FP=conmat(2,1)
FP =
     19

>> TN=conmat(2,2)
TN =
   152

```

Gambar 4.2 Confusion Matrix.

Berdasarkan data asli jumlah siswa yang diklasifikasikan IPS adalah sebanyak 129 *record*, akan tetapi untuk kelas IPS diklasifikasikan IPS (*True Positive*) adalah 110 *record* dan dari kelas IPS yang diklasifikasikan IPA (*False Positive*)

sebanyak 19 *record*. Dari data asli jumlah siswa yang diklasifikasikan IPA adalah sebanyak 175 *record*, untuk jumlah (*True Negative*) adalah 152 *record* diklasifikasikan IPA dan jumlah (*False Negative*) untuk kelas IPA yang diklasifikasikan IPS sebanyak 23 *record*.

Hasil dari *Confusion Matrix* diatas, *accuracy* yang dihasilkan dengan menggunakan algoritma *Naive Bayes* adalah sebesar 86.1842 dan *error rate* sebesar 13.8158 %. Perhitungan dari *accuracy* dan *error rate* dapat dilihat sebagai berikut :

```

>> accuracy=((TP+TN)/(TP+TN+FN+FP))*100

accuracy =

    86.1842

>> error=((FN+FP)/(TP+TN+FN+FP))*100

error =

    13.8158

```

Gambar 4.3 : Perhitungan Akurasi Confusion Matrix

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Accuracy = \frac{110+152}{110+152+19+23}$$

$$Accuracy = \frac{262}{304}$$

$$Accuracy = 0.8618$$

$$Error Rate = \frac{FP + FN}{TP + TN + FP + FN}$$

$$Error Rate = \frac{19+23}{110+152+19+23}$$

$$Error Rate = \frac{42}{304}$$

$$Error Rate = 0.1381$$

4.4 Implementasi Aplikasi

Berikut adalah tampilan aplikasi klasifikasi penjurusan siswa SMA 1 Kajen :

Gambar 4.4 : Tampilan Aplikasi

Pada klasifikasi penjurusan siswa terdapat form yang harus diisi oleh pihak SMA 1 Kajen, yaitu form rata nilai IPA, rata nilai Matematika, ujian nasional IPA dan ujian nasional Matematika di isi nilai siswa, form nilai kualitas di isi predikat A - C, form IQ di isi nilai IQ siswa dan untuk form minat jurusan di jurusan IPA atau IPS. Setelah di isi semua kemudian klik tombol proses maka sistem akan mengklasifikasi data yang diinputkan dan akan menampilkan informasi apakah siswa tersebut akan dijuruskan IPA atau IPS dan akan menampilkan hasil akurasi, eror dan jumlah data *training*.

V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil penelitian yang telah dilakukan oleh peneliti, maka dapat disimpulkan bahwa klasifikasi data siswa baru SMA 1 Kajen tahun ajaran 2015/2016 dapat diselesaikan dengan menggunakan teknik data mining, yaitu dengan metode algoritma *Naive Bayes Classifier*. Akurasi yang dihasilkan dari metode algoritma *Naive Bayes Classifier* menggunakan matlab adalah sebesar 86,1842 % dengan error rate sebesar 13,8158 %.

Jadi klasifikasi data siswa baru SMA 1 Kajen tahun ajaran 2015/2016 dengan menggunakan algoritma *Naive Bayes* dapat menjadi sistem pendukung keputusan untuk membantu pihak SMA 1 Kajen khususnya guru Bimbingan Karir (BK) dalam proses menentukan jurusan siswa.

5.2 Saran

Agar penelitian ini terus berkembang, berikut saran yang diusulkan peneliti :

1. Penelitian ini dapat dikembangkan dengan metode data mining klasifikasi lainnya untuk melakukan perbandingan.
2. Penelitian ini diharapkan dapat digunakan pihak SMA 1 Kajen sebagai pedoman pengambilan keputusan dalam melakukan penjurusan siswa.

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.

DAFTAR PUSTAKA

- [1] Kementerian Pendidikan dan Kebudayaan, Salinan – *Permendikbud Nomor 81A Tahun 2013 Tentang Implementasi Kurikulum*. Jakarta, Indonesia, 2013.
- [2] Abi Sukrisna, "Penerapan Kurikulum 2013 dan Penjurusan Siswa," SMA 1 Kajen, Pekalongan, Interview 2015.
- [3] Hikmatul Hidayah, "Klasifikasi Penjurusan Program Studi Sekolah Menengah Atas Dengan Algoritma Naive Bayes Classifier Pada SMA N 1 Subah," 2014.
- [4] Witten Ian H., Eibe Frank, and Hall Mark A., *Data Mining : Practical Machine Learning Tools and Techniques 3rd Edition*. Elsevier, 2011.
- [5] Obbie Kristanto, "Penerapan Algoritma Klasifikasi Data Mining ID3 untuk Menentukan Penjurusan Siswa SMAN 6 Semarang," 2014.
- [6] Selvia Lorena Br Ginting and Reggy Pasya Trinanda, *Penggunaan Metode Naive Bayes Classification pada Aplikasi Perpustakaan*. Bandung, Indonesia, 2010.
- [7] Claudia Clarentia Ciptohartono, "Algoritma Klasifikasi Naive Bayes Untuk Menilai Kelayakan Kredit," 2014.
- [8] Arief Jananto, "Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa," *Jurnal Teknologi Informasi DINAMIK*, vol. 18, no. 1, halaman 09-16, Januari 2013.
- [9] Yuda Septian Nugroho, "Data Mining Menggunakan Algoritma Naive Bayes untuk Klasifikasi Kelulusan Mahasiswa Universitas Dian Nuswantoro," 2014.

- [10] Faid Ari Prastya, "Penerapan Algoritma C4.5 untuk Prediksi Jurusan siswa SMAN 3 Rembang," 2015.
- [11] Kementerian Pendidikan dan Kebudayaan, *Pedoman Peminatan Peserta Didik*. Jakarta, Indonesia. Kementerian Pendidikan dan Kebudayaan, 2013.
- [12] Santosa, B. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Graha Ilmu. Yogyakarta. 2007.
- [13] Kusriani, dan Emha Taufiq Luthfi. *Alogaritma Data Mining*. Yogyakarta: Andi Publisher. 2009.
- [14] Sedy Winanta, Yetli Oslan, Gunawan Santoso, "Implementasi Metode Bayesian Dalam Penjurusan di SMA Brunderan Purworejo," *Jurnal EKSIS*, vol. 06, no. 02, halaman 21-28, November 2013.
- [15] Bustami, "Penerapan Algoritma Naive Bayes untuk Mengklasifikasi Data Nasabah Asuransi," *Jurnal Informatika*, vol. 8, Januari 2014.
- [16] Bahar, "Penentuan Jurusan Sekolah Menengah Atas dengan Algoritma Fuzzy C Means," 2011.
- [17] "Peraturan Menteri Pendidikan dan Kebudayaan No. 66 tahun 2014 tentang Standar Penilaian Pendidikan," Kementerian Pendidikan dan Kebudayaan, 2013.