

Perbandingan Metode Naive Bayes Classifier Dan Decision Tree C4.5 Dalam Mencari Pola Minat Pemilihan Jurusan Di Madrasah Aliyah (Studi Kasus:MA El-Bayan Majenang)

Dede Yusuf, Zulfikar Yusya Mubarak, Annisa Rahayu Pangesti, Nuni Wulansari, Rizki Zulqornain

Universitas Al-Irsyad Cilacao, Jawa Tengah, Negara

Email: yusuf.dede17@gmail.com, zulfikaryusya@gmail.com, r.nisapangesti@gmail.com
nuniwulansari@universitasalirsyad.ac.id, rizqizulqornain056@gmail.com

Abstract

Pemilihan jurusan siswa merupakan salah satu aturan pengambilan keputusan dalam menentukan jurusan berdasarkan minat dan bakat, ini bertujuan untuk memahami potensi diri dan peluang yang ada. Dalam menentukan jurusan siswa pada kurikulum 2013, dilakukan pada awal masuk semester pertama, yaitu pada kelas 10(X) Sekolah Menengah Atas. Penerapan kurikulum 2013 bertujuan untuk mendukung adaptasi program pendidikan dengan menangkap karakteristik dan potensi dari peserta didik. [1]. Dalam penerapan kurikulum 2013 berdampak pada salah satunya pihak sekolah terutama guru BK belum mengetahui bakat, minat, karakter serta kemampuan finansial keluarga siswa untuk memilih mata pelajaran tertentu, sehingga guru BK harus benar-benar bisa merekomendasikan jurusan yang sesuai dengan minat dan bakat siswa. Berdasarkan permasalahan tersebut maka dilakukan penerapan data mining menggunakan algoritma Naïve Bayes dan algoritma Decesion Tree C4.5 untuk klasifikasi penjurusan siswa pada MA El-Bayan Majenang. Algoritma Naïve Bayes [2] dan algoritma Decesion Tree C4.5 [2] merupakan algoritma dengan pola klasifikasi pohon keputusan yang digunakan karena memiliki beberapa kelebihan dibandingkan algoritma lainnya. Dalam penelitian ini penulis membandingkan 2 metode tersebut, diawali dari pengumpulan data selanjutnya cleaning data dan dilanjutkan memproses data testing dan data training. Setelah semua data diproses maka dilanjutkan pada proses klasifikasi untuk masing-masing algoritma. Dari penelitian ini bisa kita simpulkan bahwa: 1) Algoritma Naïve Bayes dan Algoritma Decesion TreeC4.5 bisa digunakan untuk mencari pola minat siswa dalam menentukan jurusan dengan berbagai atribut yang berbeda. 2) Dalam beberapa tahun terakhir nilai Ujian Nasional dan nilai Ujian Sekolah sangat menentukan ketika masuk ke jurusan tertentu karena pada pengujian decision tree selalu berada pada posisi root node. 3) Algoritma Naïve bayes mempunyai tingkat akurasi tertinggi dengan 73 % dibanding dengan algoritma decision treeC4.5 dengan tingka takurasi 55 %.

Keywords: Data Mining, Educational Data Mining, Algoritma Naïve Bayes, Algoritma Decesion Tree C4.5

1. PENGANTAR

Untuk mewujudkan tujuan dari pendidikan nasional secara menyeluruh dan maksimal maka setiap siswa diwajibkan menempuh jenjang pendidikan formal setidaknya sampai siswa menempuh setera Sekolah Menengah Atas (SMA)). Pendidikan level menengah terdiri dari pendidikan level menengah umum dan pendidikan level menengah kejuruan. SMA, MA, SMK, MAK atau bentuk pendidikan menengah lainnya yang sederajat. Teknologi informasi dan komunikasi serta sistem informasi diharapkan dapat membantu pihak sekolah dalam memperkuat dan membantu berjalannya pembelajaran yang sesuai dengan tujuan sekolah tersebut terutama dalam menghadapi persaingan antar sekolah satu daerah. Sejalan dengan hal tersebut maka pihak Sekolah MA-ELBAYAN Majenang yang beralamat di Jl. K. H. Najmuddin, Benda Kulon, Padangjaya, Majenang, Kabupaten Cilacap, Jawa Tengah 5325 dapat membantu keputusan siswa dalam menentukan pemilihan penjurusan bidang studi yaitu IPA, Bahasa Jepang, IPS, berdasarkan minat dan kemampuan siswa/i nya.

Pada penelitian kali ini penulis akan mencari pola yang terjadi pada pemilihan jurusan yang sudah dilakukan di MA El-Bayan Majenang yang sesuai dengan minat siswa dalam 5 tahun terakhir, ini dikarenakan di Sekolah tersebut tahun ke tahun siswa yang mendaftar semakin banyak. Bisa dilihat data itu pada penerimaan siswa baru tahun 2016/2017 siswa yang mendaftar sebanyak 320 siswa, pada tahun 2017/2018 sebanyak 342 siswa dan tahun 2018/2019 sebanyak 382 siswa. Data ini penulis dapatkan dari hasil wawancara dengan panitia penerimaan siswa baru, penulis melihat data tersebut semakin kompleks karena jumlah penerimaan siswa semakin banyak dari tahun ke tahun. Data yang cukup banyak yang dimiliki oleh pihak MA tidak dimanfaatkan dengan cara efektif dan efisien, maka penulis disini ingin mencoba mencari pola kesesuaian jurusan yang diambil oleh siswa di MA El-Bayan Majenang, hal-hal apa saja yang menjadi pendorong bagi para siswa SMP/MTS untuk memilih jurusan yang ada disana. Sehingga siswa benar-benar yakin dalam memilih jurusan tersebut, yang nantinya jurusan tersebut akan siswa jalankan selama 3 tahun ke depan, dan pastinya membawa dampak positif untuk siswa di kemudian hari. Jadi kedepannya siswa SMP/MTS yang akan masuk ke jurusan tertentu harus mulai berfikir apakah siswa cocok masuk ke jurusan tersebut atau masuk ke jurusan yang lain. Dari permasalahan yang ada penulis akan membandingkan dua metode yaitu metode Algoritma Naïve Bayes dengan Algoritma C4.5 untuk bisa mengetahui pola minat pada pemilihan jurusan. Mana dari 2 metode tersebut akurasi paling tinggi dalam menentukan pola pemilihan jurusan yang sesuai minat di MA El-Bayan Majenang selama 5 tahun terakhir.

Pencarian pola ini dilakukan untuk melihat sejauh mana atribut-atribut yang berpengaruh terhadap pemilihan jurusan, sehingga nanti kedepannya akan mendapatkan pola bahwa siswa dengan jurusan tertentu harus memiliki atribut dengan nilai-nilai tertentu. Ini akan menjadi acuan bagi siswa, guru dan pihak lainnya bahwa melakukan penjurusan memang harus memiliki berbagai aspek pendukung [3]. Sehingga nanti dalam pemilihan jurusan tidak semata-merta hanya sesuai minat siswa saja tapi harus juga diukur dengan kemampuan siswa atau bisa dilihat dari aspek lainnya.

Pengolah Data menggunakan pola data mining adalah salah satu dari berbagai cara pengolahan untuk melihat pencarian pola tersebut, karena data mining menurut (M. Ridwan dan H, Suyono 2013) [4] Penambangan data adalah proses yang menggabungkan ilmu statistik, kecerdasan buatan, matematika, dan teknik pembelajaran mesin untuk mengekstraksi & menganalisa informasi yang berguna dan pengetahuan yang relevan dari basis data besar. Istilah data mining adalah disiplin ilmu yang tujuan utamanya yaitu menemukan, mengeksplorasi, dan mengekstraksi pengetahuan dari data dan informasi. Penambangan data sering disebut sebagai Knowledge Discovery in Databases (KDD). KDD adalah kegiatan yang melibatkan pengumpulan penggunaan data, pola pelacakan, dan pola atau hubungan dalam kumpulan data yang besar..

2. METODE

2.1. Metode penelitian

Metode penelitian adalah metode analisis deskriptif berdasarkan pendekatan kuantitatif, yaitu penelitian yang berpusat pada analisis data numerik (bilangan), dengan tujuan untuk memahami keadaan berdasarkan data yang diperoleh melalui penyajian, pengumpulan, dan analisis. adalah. Menganalisis data menjadi informasi baru yang dapat digunakan untuk menganalisis masalah yang diteliti [6].

2.2. Metode Pengumpulan Data

Setelah masalah dirumuskan, langkah selanjutnya adalah mengumpulkan informasi yang diperlukan. Data yang dibutuhkan adalah data yang mempengaruhi subjek yang berisi yang sesuai dengan objek penelitian:

- a. Data hasil Wawancara
Hasil keluaran dari wawancara siswa adalah nilai yang akan bisa menjadi masukan untuk pihak BK mendapatkan nilai yang harus dimasukkan ke dalam atribut itu, berupa jumlah tanggungan orang tua, penghasilan orang tua per bulannya dan jarak tempat tinggal ke lokasi sekolah
- b. Data minat siswa terhadap jurusan
Data ini diperoleh dari angket yang disebar oleh guru BK (Bimbingan Konseling), sebelum proses siswa menentukan jurusan yang

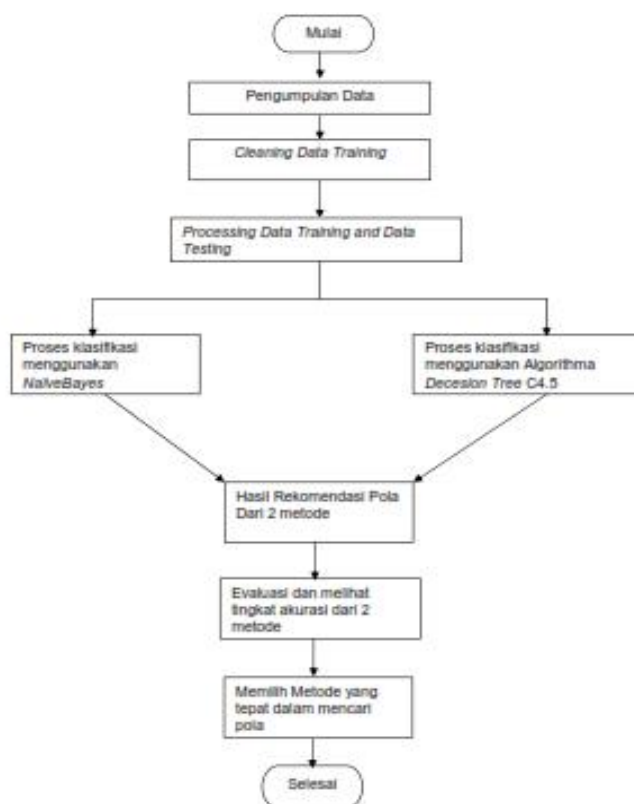
berlangsung di MA El-Bayan Majenang. Hasil keluarannya adalah siswa masuk pada jurusan yang diinginkan bisa jurusan IPA, IPS, atau Bahasa Jepang

2.3. Metode Analisis Data

Metode data mining pendidikan berasal dari sumber literatur yang berbeda termasuk data mining, pembelajaran mesin, psikometri, dan area lain dari model komputasi, statistik, dan visualisasi informasi. Sehingga untuk menganalisis data dalam penerapandata mining ini menggunakan tahapan Educational Data Mining yang terdiri dari dua kategori, yaitu 1) Web Mining dan 2) statistik dan visualisasi [7].

2.4. Gambaran Alur Penelitian

Ini adalah gambaran alur penelitian yang penulis lakukan yaitu dari mulai pengumpulan data sampel pada hasilnya yaitu mendapatkan pola atau atribut yang sesuai dalam pemilihan jurusan yang sesuai dengan minat dari siswa SMP/MTS selama 5 tahun terakhir ini ada pada gambar 3.5



Gambar 2.1 Alur Penelitian

3. HASIL DAN DISKUSI

a. Analisis Data

Algoritma Naive Bayes dan Algoritma Decision Tree C4.5 merupakan metode klasifikasi pada data mining, hanya perbedaan dari 2 algoritma ini adalah jika naive bayes berfokus pada asumsi yang berdasarkan teorema Bayes [8] sedangkan Algoritma Decision tree menggunakan pohon keputusan [9] untuk hasil akhirnya. Data yang digunakan adalah data pilihan pemilihan jurusan di MA El-Bayan Majenang ini, data PPDB di MA El-Bayan Majenang tahun 2014/2015 sampai 2018/2019, dengan jumlah siswa sebanyak 1617 Siswa, dengan 3 pilihan jurusan. Data yang akan digunakan oleh penulis yaitu 885 Siswa itu akumulasi dari data siswa 5 tahun terakhir.

b. Seleksi data

Pada tahap ini akan dilakukan proses pengumpulan data, dimana data yang akan digunakan adalah data pilihan jurusan yang dipilih oleh siswa SMP yang akan masuk ke MA El-Bayan Majenang, data PPDB di MA El-Bayan Majenang tahun 2014/2015, 2015/2016, 2016/2017, 2017/2018, dan 2018/2019.

Tabel 3.1 Data Mahasiswa Setiap Tahun

Tahun Ajar	Jumlah Siswa
2014/2015	180
2015/2016	187
2016/2017	181
2017/2018	166
2018/2019	171
Total	885

Pada table dibawah ini menunjukkan data mahasiswa yang sudah disesuaikan dengan atribut dan akan digunakan sebagai data *cleaning*, dari data ini akan dipilih data mahasiswa yang bisa digunakan, karena pada kenyataannya banyak data mahasiswa yang tidak lengkap. Sehingga kita harus melakukan proses *cleaning*.

Tabel 3.2 Contoh Atribut Data Mahasiswa

Gender	Penghasilan orang tua	Jumlah Tanggungan orang tua	Perkiraan jarak tempat tinggal ke sekolah	Asal Sekolah	Alamat	Pekerjaan Ayah	Pendidikan Terakhir Ibu	Rata-rata Nilai Us	Rata-rata Nilai Uts	Hasil Wawancara	Jurusan
P	<=1 juta	1	2 Km	MITS EL BAYAN MAJENANG	MAJENANG	PETANI	SD	80	71	Bahasa Jepang	Bahasa Jepang
P	1 - 2 juta	4	1 Km	MITS EL BAYAN MAJENANG	MAJENANG	PEDAGANG	SMP	85	71	Bahasa Jepang	Bahasa Jepang
L	1 - 2 juta	3	1 Km	MITS EL BAYAN MAJENANG	MAJENANG	PEDAGANG	SMP	75	70	Bahasa Jepang	Bahasa Jepang
P	1 - 2 juta	2	2 Km	MITS EL BAYAN MAJENANG	MAJENANG	PEDAGANG	SMP	74	69	Bahasa Jepang	Bahasa Jepang
P	1 - 2 juta	2	7 Km	SMP DIPONEGORO MAJENANG	MAJENANG	PEDAGANG	SMP	73	72	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	1	13 Km	SMP NEGERI 1 DAYEUHLUHUR	MAJENANG	PEDAGANG	SMK	72	71	Bahasa Jepang	Bahasa Jepang
P	1 - 2 juta	4	9 Km	SMP MANUHUNDYAN 1 DAYEUHLUHUR	MAJENANG	PEDAGANG	SMK	71	70	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	1	1 Km	MITS EL BAYAN MAJENANG	MAJENANG	PEDAGANG	SMK	70	69	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	1	2 Km	MITS EL BAYAN MAJENANG	MAJENANG	PEDAGANG	SMK	69	71	Bahasa Jepang	Bahasa Jepang
P	1 - 2 juta	1	1 Km	MITS EL BAYAN MAJENANG	MAJENANG	PEDAGANG	SMK	67	70	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	3	10 Km	MITS EL BAYAN MAJENANG	MAJENANG	PEDAGANG	SMK	66	65	Bahasa Jepang	Bahasa Jepang
L	1 - 2 juta	1	5 Km	MITS EL BAYAN MAJENANG	MAJENANG	PEDAGANG	SMK	65	71	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	2	2 Km	MITS EL BAYAN MAJENANG	MAJENANG	PETANI	SD	64	70	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	1	12 Km	MITS EL BAYAN MAJENANG	MAJENANG	PETANI	SD	75	70	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	3	17 Km	MITS EL BAYAN MAJENANG	MAJENANG	PETANI	SD	74	70	Bahasa Jepang	Bahasa Jepang
P	1 - 2 juta	2	3 Km	MITS EL BAYAN MAJENANG	MAJENANG	PETANI	SD	73	69	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	1	1 Km	MITS EL BAYAN MAJENANG	MAJENANG	PETANI	SD	72	71	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	4	1 Km	MITS EL BAYAN MAJENANG	MAJENANG	PETANI	SD	71	70	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	1	2 Km	SMP NEGERI 2 CEMINGGU	CEMINGGU	PETANI	SD	70	69	Bahasa Jepang	Bahasa Jepang
L	1 - 2 juta	1	1 Km	SMP NEGERI 1 SATI AJAR CEMINGGU	CEMINGGU	PEDAGANG	SD	69	67	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	1	1 Km	SMP AL ISLAM CIPARI	CIPARI	PEDAGANG	SD	71	65	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	2	2 Km	MITS EL BAYAN MAJENANG	CIPARI	PEDAGANG	SMP	70	65	Bahasa Jepang	Bahasa Jepang
P	1 - 2 juta	1	7 Km	MITS EL BAYAN MAJENANG	CIPARI	PEDAGANG	SMP	69	73	Bahasa Jepang	Bahasa Jepang
P	<=1 juta	4	13 Km	MITS EL BAYAN MAJENANG	CIPARI	PEDAGANG	SMP	67	74	Bahasa Jepang	Bahasa Jepang
L	<=1 juta	1	9 Km	MITS EL BAYAN MAJENANG	CIPARI	PEDAGANG	SMP	85	69	IPA	Bahasa Jepang
P	<=1 juta	1	1 Km	SMP AL HUSNA KAWUNGANTEN	KAWUNGANTEN	PEDAGANG	SMP	85	71	Bahasa Jepang	Bahasa Jepang
P	1 - 2 juta	1	2 Km	SMP NEGERI 1 KAWUNGANTEN	KAWUNGANTEN	PEDAGANG	SMP	75	70	Bahasa Jepang	Bahasa Jepang

Berdasarkan tabel 3.2, penulis akan melakukan analisa terhadap data tersebut untuk atribut mana yang akan digunakan pada tahapan-tahapan selanjutnya. Sehingga muncul beberapa atribut yang bisa digunakan sebagai variable untuk menentukan jurusan yang tepat, ada beberapa atribut yang sudah diseleksi :

Table 3.3 Atribut

NO	ATRIBUT
1	Gender
2	Jumlah Tanggungan Orang Tua
3	Penghasilan Orang Tua
4	Perkiraan Jarak Tempat Tinggal Ke Sekolah
5	Asal Sekolah
6	Alamat
7	Pekerjaan Ayah
8	Pendidikan Terakhir Ibu
9	Rata-Rata Nilai Us
10	Rata-Rata Nilai Un
11	Hasil Wawancara Dengan Pihak Sekolah

c. Data *Cleaning* (Pembersihan Data)

Selanjutnya tahap pembersihan data, pada tahap ini data yang tidak sesuai atau tidak relevan maka akan dihapus. Data yang dihapus yaitu

data yang tidak lengkap, itu dikarenakan siswa ketika mengisi formulir tidak lengkap jadi ada beberapa data siswa yang tidak dimasukkan ke dalam data ini. Contoh data siswa yang tidak dimasukkan ke dalam data ini dikarenakan data yang tidak lengkap dan tidak relevan.

Tabel 3.4 Hasil Data Cleaning

No	NAMA	Gender	Penghasilan orang tua	Jumlah Tanggungan orang tua	Perkiraan Jarak(sekolah ke tempat tinggal)	Asal Sekolah	Alamat	Pekerjaan Ayah	Pendidikan Terakhir Ibu	Rata-rata Nilai US	Rata-rata Nilai UN
1	Imas Arifah	P	>=4 juta		3 Km	MTS EL-BAYAN MAJENANG	CIMANGGU	WIRASWASTA	SMA	73	84
2	HADIS FADILAH	L	>=4 juta	1		MTS EL-BAYAN MAJENANG	WANAREJA	PENJAHIT		69	72
3	HALIMATUS SADIYYAH	L		2	1 Km		WANAREJA		SMK	72	71

d. Data Transformation (Transformasi Data)

Tidak semua format data dapat diolah oleh aplikasi Rapid Miner, data masih harus melalui beberapa tahapan lagi. Setelah melalui tahapan Data Selection dan Cleaning, data masih harus melalui tahapan Transformation. Beberapa format data yang dapat diolah dalam aplikasi Rapid Miner antara lain, data berformat *.xlsx, *.csv. dalam penelitian ini peneliti menggunakan data berformat *.xlsx untuk di olah kedalam aplikasi Rapid Miner.

Tabel 3.5 Hasil data transformation

Gender	Jumlah Saudara Kandung	Penghasilan orang tua	Perkiraan Jarak(sekolah ke tempat tinggal)	Hasil Wawancara	Rata2 Nilai US	Rata Nilai UN	Jurusan
F	3	<=1 juta	1 Km	Bahasa Jepang	59,22	53	Bahasa Jepang
F	2	1-2 juta	1 Km	Bahasa Jepang	69,57	46,5	Bahasa Jepang
L	1	1-2 juta	1 Km	Bahasa Jepang	65,56	52,75	Bahasa Jepang
L	4	1-2 juta	2 Km	Bahasa Jepang	65,57	43	Bahasa Jepang
L	3	<=1 juta	1 Km	Bahasa Jepang	71,22	55,75	Bahasa Jepang
F	3	1-2 juta	1 Km	Bahasa Jepang	64,22	58,25	Bahasa Jepang
F	2	1-2 juta	1 Km	Bahasa Jepang	65,11	62,5	Bahasa Jepang
F	2	1-2 juta	3 Km	Bahasa Jepang	73,00	63	Bahasa Jepang
F	1	<=1 juta	2 Km	IPS	68,48	70	IPS
F	2	1-2 juta	1 Km	IPS	69,70	44,5	IPS

e. Data Mining

Tahap ini dilakukan pada aplikasi RapidMiner dengan menggunakan Klasifikasi Naïve Bayes dan algoritma Decesion TreeC4.5, pengolahan data dilakukan ketika file data sudah siap diolah, file yang digunakan berektensi *.xlsx setelah file siap, maka pada bahasan ini akan dilakukan

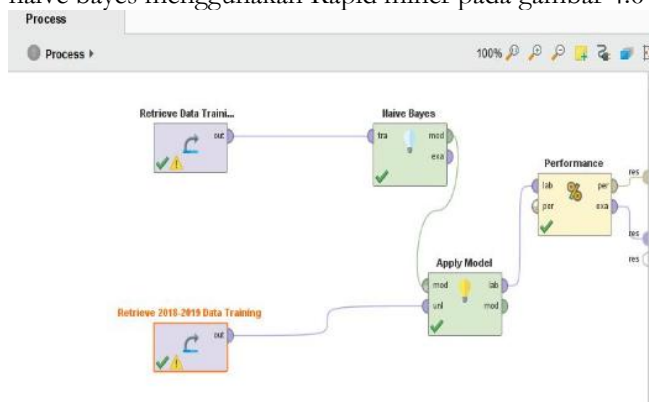
proses pengolahan data yang disebut proses Data Mining, Proses Data Mining dengan menggunakan aplikasi RapidMiner.

f. Tahap Klasifikasi

Pada tahap klasifikasi, dataset yang telah ditentukan kemudian di-import untuk selanjutnya akan dilakukan proses *training* dan *testing* dengan algoritma yang diusulkan, yaitu proses pertama menggunakan Naïve Bayes dan yang kedua menggunakan algoritma Decesion TreeC4.5.

g. Hasil Rekomendasi Naïve Bayes menggunakan Rapid Miner9.2

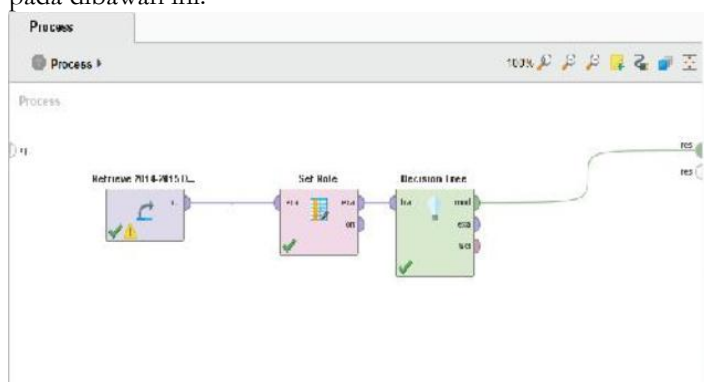
Rancangan proses untuk prediksi data penjurusan siswa dengan metode naïve bayes menggunakan Rapid miner pada gambar 4.6



Gambar 3.1 Hasil Naïve Bayes

h. Hasil Rekomendasi Decesion TreeC4.5 menggunakan Rapid Miner9.2

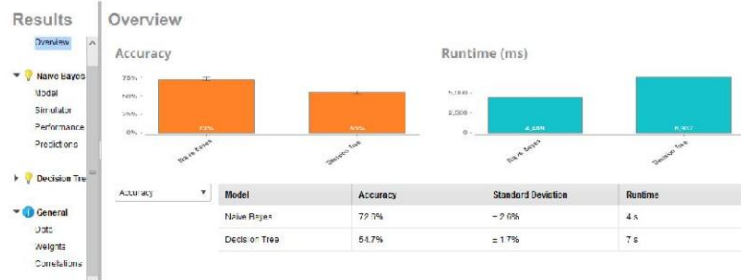
Rancangan proses klasifikasi data penjurusan siswa dengan metode decision tree C4.5 menggunakan aplikasi Rapid Miner 9.2 ditunjukan pada dibawah ini.



Gambar 3.2 Decesion Tree C45

i. Evaluasi

Pada tahap ini bisa mengecek akurasi dan runtime dari masing-masing metode, untuk naïve bayes tingkat akurasinya yaitu 72.6 % dengan runtime 4 s dan Decesion Tree C4.5 akurasinya yaitu 54,7 % dengan runtime 7 s.



Gambar 3.3 Hasil Perbandingan 2 Metode

4. KESIMPULAN

Sebagai hasil dari penelitian yang sudah dilakukan, penulis mendapatkan beberapa kesimpulan dari hasil penelitian yang bisa jadi rujukan khususnya bagi MA El-Bayan Majenang dan umumnya untuk sesama peneliti. Kesimpulan yang didapatkan yaitu ::

- Algoritma Naïve Bayes dan Algoritma Decesion Tree C4.5 bisa digunakan untuk mencari pola minat siswa dalam menentukan jurusan dengan berbagai atribut yang berbeda.
- Dalam beberapa tahun terakhir Rata-rata nilai Ujian Nasioanal dan Rata-rata nilai Ujian Sekolah sangat menentukan ketika masuk ke jurusan tertentu karena pada pengujian Decision tree C4.5 selalu berada pada posisi root node.
- Algoritma Naïve bayes mempunyai tingkat akurasi tertinggi dengan 73 % dibanding dengan Algoritma Decision tree C4.5 dengan tingkat akurasi 55 %.

References

- [1] P. PRAMUJO, "IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA C4.5 UNTUK KLASIFIKASI PENJURUSAN SISWA PADA SMA NEGERI 2 PEMALANG," *FIK (Universitas Dian Nuswantoro)*,

pp. 10-11, 2016.

- [2] T. Emha and Kusrini, *Algoritma data mining*, Yogyakarta: Andi Offset, 2009.
- [3] Z. Y. Mubarak, U. Ema and E. T. Luthfi, "Sistem Pendukung Keputusan Penyeleksian Kelayakan Proposal Penelitian Dosen Menggunakan Metode Simple Additive Weighting dan Weight Product," *SEMNAS TEKNOLOGI ONLINE*, pp. 2-8-109, 2018.
- [4] M. Ridwan, S. Hadi and S. M, "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier,," *EECCIS*, vol. Vol 7 No 1, p. 204, 2013.
- [5] B. Santosa, *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*, Yogyakarta: Graha Ilmu, 2007.
- [6] Sugiyono, *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*, Bandung: Alfabeta, 2014.
- [7] A. Algarni, "Data Mining in Education," (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, Vols. Vol. 7, No. 6, pp. 457 - 461, 2016.
- [8] A. F. Cahyanti, R. Saptono and S. W. Sihwi, "Penentuan Model Terbaik pada Metode Naive Bayes Classifier dalam Menentukan Status Gizi Balita dengan Mempertimbangkan Independensi Parameter," *JURNAL ITSMART*, vol. Vol 4. No 1., pp. 28 - 35, 2015.
- [9] Zulfian Azmi and M. Dahria, "DECISION TREE BERBASIS ALGORITMA UNTUK PENGAMBILAN KEPUTUSAN," *Jurnal SAINTIKOM*, Vols. Vol. 12, No.3, pp. 157 - 164, 2013.