

# METODE ANALYTICAL HIERARCHY PROCESS BERBASIS DECISION TREE C4.5 DALAM PEMILIHAN JURUSAN

<sup>1</sup>Siti Masamah, <sup>2</sup>Sudin Saepudin, <sup>3</sup>Dudih Gustian, <sup>4</sup>Mupaat

<sup>1,2,3,4</sup>Program Studi Sistem Informasi

<sup>1,2,3,4</sup>Sekolah Tinggi Teknologi Nusa Putra

<sup>1,2,3,4</sup>Jl.Raya Cibolang Kaler No.21 Kab, Sukabumi

e-mail: <sup>1</sup>siti.masamah@nusaputra.ac.id, <sup>2</sup>Sudin.Saepudin@nusaputra.ac.id, <sup>3</sup>dudih@nusaputra.ac.id,

<sup>4</sup>mupaat@nusaputra.ac.id

Korespondensi: <sup>2</sup>Sudin.Saepudin@nusaputra.ac.id

## ABSTRAK

Pendidikan kejuruan merupakan pendidikan menengah yang mempersiapkan peserta didik terutama untuk bekerja dalam bidang tertentu”. Maka dari itu diperlukan pengambilan keputusan jurusan yang tepat dan sesuai sasaran, melihat dari pengambilan keputusan yang perlu mempertimbangkan banyak hal seperti kurikulum, fasilitas, keinginan, dan faktor lain. Dan seperti artikel yang dikutip dari *news.okezone.com* yang berjudul “Duh, 87% mahasiswa di Indonesia salah jurusan”. Oleh karena itu penelitian ini bertujuan untuk menganalisis dari beberapa indikator yang dapat menyebabkan salah mengambil jurusan melalui cara mengurutkannya dengan metode Analytical Hierarchy Process, di mana dengan metode ini dapat mengurutkan masalah. Selain itu digunakan juga proses data mining menggunakan algoritma Decision Tree C4.5 yang bertujuan untuk proses klasifikasi dataset yang dibagi menjadi data training dan testing. Di mana menghasilkan model tree dari data training dan tingkat akurasi serta validasi dari data testing dengan nilai 96,88% dan 0.500% yang dikategorikan good classification.

**Kata Kunci :** *Analytical Hierarchy Process, Decision Tree C4.5, good classification*

## ABSTRACT

Vocational education is secondary education that prepares students especially to work in certain fields". Therefore, it is necessary to make appropriate and targeted department decisions, judging from the decision making that needs to consider many things such as curriculum, facilities, desires, and other factors. And as an article quoted from *news.okezone.com* entitled "Duh, 87% of students in Indonesia take the wrong major". Therefore, this study aims to analyze several indicators that can lead to wrong majors by sorting them using the Analytical Hierarchy Process method, where this method can sort problems. In addition, the data mining process using the Decision Tree C4.5 algorithm is also used which aims to classify the dataset which is divided into training and testing data. Where to generate a tree model from training data and the level of accuracy and validation of the test data testing with a value of 96.88% and 0.500% which are categorized as good classification.

**Kata Kunci :** *Analytical Hierarchy Process, Decision Tree C4.5, good classification*

## I. PENDAHULUAN

Sekolah Menengah Kejuruan (SMK) merupakan salah satu lembaga pendidikan formal yang secara khusus mempersiapkan peserta didiknya untuk siap bekerja di dunia industri, berwirausaha secara mandiri, atau melanjutkan pendidikan ke jenjang yang lebih tinggi sesuai kejuruannya [1].

Menurut UU Sisdiknas pasal 15 Depdiknas (2004:8) disebutkan bahwa “Pendidikan kejuruan merupakan pendidikan menengah yang mempersiapkan peserta didik terutama untuk bekerja dalam bidang tertentu”. Artikel yang berjudul “Duh, 87% mahasiswa di Indonesia salah jurusan” Ketika memilih jurusan pada saat masuk kuliah, tentu saja ada banyak poin yang perlu

dipertimbangkan. Jangan sampai salah jurusan yang berakibat fatal pada dunia kerja. Menurut Educational Psychologist dari Integrity Development Flexibility (IDF) Irene Guntur, M.Psi., Psi., CGA, sebanyak 87 persen mahasiswa di Indonesia salah jurusan. “Salah jurusan bisa memicu pada pengangguran. Supaya tidak ada pengangguran lagi ya jangan sampai ketika kuliah salah jurusan [2].

Sementara itu data dari Kementerian Pendidikan dan Kebudayaan Sekretariat Jenderal Pusat Data dan Statistik Pendidikan dan Kebudayaan 2016 mengenai Statistik Sekolah Menengah Kejuruan (SMK) tahun 2015/2016, menjelaskan bahwa gambaran umum indikator pendidikan SMK dengan jumlah siswa sebanyak 4.211.245 sementara angka mengulang sebanyak 8.461, angka putus sekolah sebanyak 77.899, dan angka lulusan dengan jumlah tingkat XII sebanyak 1.469.289 dengan lulusan sebanyak 1.429.870 [3].

No.	Jenis Indikator Type of Indicators	Nilai Indikator Value of Indicators
<i>Angka/Persentase / Rates/Percentage</i>		
1.	Angka Mengulang/ Repetition Rate	Siswa / Pupils 2014/2015 4.211.245 Mengulang Repeaters 8.461
2.	Angka Putus Sekolah / Drop-out Rate	Siswa / Pupils 2014/2015 4.211.245 Putus Sekolah Drop-outs 77.899
3.	Angka Lulusan / Completion Rate	Siswa Tk.XII 14/15 Pupils Gr.XII 14/15 1.469.289 Lulusan Graduates 1.429.870

Gambar 1. Gambaran umum indikator pendidikan SMK di tahun 2015/2016

SMK Harapan Bangsa adalah salah satu sekolah menengah kejuruan di Sukabumi yang mempunyai pilihan jurusan Otomotif dan Teknik Jaringan Komputer. Oleh karena itu melihat dari permasalahan yang ada peneliti merasa perlu untuk melakukan penelitian agar dengan sedini mungkin kita dapat meminimalisir kesalahan pengambilan keputusan guna pemilihan jurusan. Karena pada dasarnya siswa SMK disiapkan untuk siap diserap oleh dunia kerja. Dengan hal itu maka kita harus mengetahui urutan kriteria apa yang paling berpengaruh dalam pemilihan jurusan selain itu kita dapat memberikan alternatif keputusan

jurusan, sehingga diharapkan agar siswa tidak salah mengambil keputusan jurusan dan meminimalisir angka mengulang dan drop out.

Dari beberapa peneliti sebelumnya yang dilakukan oleh Obbie Kristanto dalam menentukan penjurusan siswa SMAN 6 Semarang dengan metode ID3 di peroleh kesimpulan mengetahui tingkat akurasi dengan menggunakan data dari guru BP sebagai perbandingan. Dengan menggunakan 371 dataset dan 20 data uji yang diinputkan terdapat 4 kasus yang meleset dan 16 kasus berhasil, sehingga didapat akurasi sebesar 80% [4]. Sementara penelitian yang dilakukan oleh Liliana Swatina (2013) didapat kesimpulan bahwa algoritma Decision Tree C4.5 akurat diterapkan untuk penentuan kesesuaian jurusan mahasiswa dengan keakuratan 93.31 % dan akurasi rekomendasi jurusan sebesar 82.64% [5]. Selain itu penelitian yang dilakukan oleh Yeni Kustiyahningsih dan Nikmatu Syafa'ah dalam penentuan jurusan dengan menggunakan metode KNN dan smart dapat diperoleh hasil yang cukup mendekati keakuratan dari data yang telah ada sebesar 62,5%, dihitung dari total hasil jumlah jurusan dengan menggunakan metode KNN dan SMART sebanyak 20 siswa dibagi dengan jumlah banyaknya siswa dalam kelas 1 tersebut sebanyak 32 siswa [6]. Dengan menggunakan metode C4.5 akan lebih mudah untuk membuat klasifikasi tingkat kemampuan bahasa Inggris dikampus XYZ berdasarkan faktor yang diteliti [7].

Sehingga dengan adanya penelitian ini diharapkan pihak manajemen dapat memberikan alternatif keputusan jurusan yang tepat bagi para siswa dengan harapan para siswa memiliki pengetahuan ilmu yang mantap dan sesuai dengan *passion* mereka selain itu manajemen dapat melihat urutan kriteria apa yang menjadi pengaruh siswa untuk menentukan jurusan, dengan hal ini pihak manajemen dapat mengubah ke hal yang lebih baik dan melakukan pendekatan, pemberian informasi mengenai jurusan yang ada.

## II. TINJAUAN PUSTAKA

### 2.1 Penelitian Terkait

Ari Puspita et.al (2015). Melakukan penelitian dengan judul “Algoritma C4.5 Berbasis Decision Tree Untuk Memprediksi Kelahiran Bayi Prematur”. Penelitian ini menggunakan metode a

C4.5 serta indikator yang diteliti yaitu usia, sistol, diastol, riwayat darah tinggi, riwayat keguguran, riwayat premature, trauma, konsumsi rokok, konsumsi ganda, dan keputihan. Dengan hasil dari kesimpulan yaitu dari hasil prediksi model algoritma C4.5 berbasis *decession tree* C4.5 memberikan nilai akurasi yaitu 93,60% dengan nilai akurasi yang sudah mencapai excellent classification hal ini menunjukkan bahwa *decession tree* merupakan model yang cukup baik [8].

Linda Atika (2010). Melakukan penelitian dengan judul “Sistem Penunjang Keputusan Penilaian Kinerja Pemilihan Dosen Berprestasi Menggunakan Metode AHP”. Indikator yang digunakan dalam penelitian ini diantaranya pengajaran, penelitian, pengabdian, aktivitas, dan quisioner. Dengan kesimpulan aplikasi dengan metode AHP ini dapat membantu dan memberikan alternative keputusan [9].

Yeni Kustiyarningsih et.al (2015). Melakukan penelitian dengan judul “Sistem Pendukung Keputusan Untuk Menentukan Jurusan Pada Siswa SMA Menggunakan Metode KNN dan Smart”. Kesimpulan dari penelitian ini ialah cukup baik mendekati keakuratan dari data yang telah ada sebesar 62,5%, dihitung dari total hasil jumlah jurusan dengan menggunakan metode KNN dan Smart sebanyak 20 siswa dibagi dengan jumlah banyaknya siswa dalam 1 kelas tersebut sebanyak 32 siswa. Sistem pendukung keputusan menggunakan metode KNN dan SMART ini digunakan untuk menentukan jurusan bagi siswa SMA berdasarkan bobot dan kreteria yang sudah ditentukan [10].

### III. METODOLOGI PENELITIAN

#### 3.1 Studi Literatur

Metode ini membantu peneliti dalam mencari teori-teori dasar yang diperlukan dalam penelitian, seperti teori tentang data mining, algoritma C4.5 dan AHP. Selain itu, dilakukan juga pengumpulan data training dan data testing untuk proses pemilihan jurusan.

#### 3.2 Studi Pustaka

Tahap ini bertujuan untuk mencari dan mengumpulkan jurnal ilmiah, buku, artikel, dan literatur lainnya yang berkaitan dengan penelitian ini sebagai referensi.

### 3.3 Rancangan Penelitian

Langkah-langkah perancangan pada penelitian ini yaitu :

1. Mempelajari dan memahami algoritma C4.5 pada *decision tree* dan AHP untuk pembobotan urutan masalah.
2. Mempelajari cara pengukuran akurasi sistem dan mengidentifikasi parameter-parameter yang digunakan untuk pengukuran akurasi.
3. Perancangan dan pembangunan aplikasi, yang dimana peneliti menggunakan Microsoft Visual studio 2010 berbasis desktop.

### 3.4 Teknik Pengumpulan dan Analisis Data

#### 1. Teknik Pengumpulan Data

Penelitian ini menggunakan teknik angket berupa kuisisioner yaitu daftar pertanyaan tertulis yang ditujukan kepada siswa siswi di SMK Harapan Bangsa.

#### 2. Analisis Data.

Analisa data yang digunakan yaitu analisa data

## IV. HASIL DAN PEMBAHASAN

### 4.1 Analisa dengan metode AHP

Proses analisis dengan AHP melalui beberapa tahapapan, maka setelah final diperoleh analisis hirarchy sebagai berikut ini :

#### 1. Kreteria Keinginan

Tabel 1. Kriteria keinginan

No	Perusahaan	Faktor Evaluasi	Faktor Bobot	Bobot Evaluasi
1	Otomotif	0.100	0.385	0.0385
2	TKJ	0.200	0.269	0.0538
3	Multimedia	0.100	0.192	0.0192
4	Bangunan	0.100	0.115	0.0115
5	RPL	0.143	0.038	0.0055
Jumlah			1.00	0.13

#### 2. Kriteria faktor lain

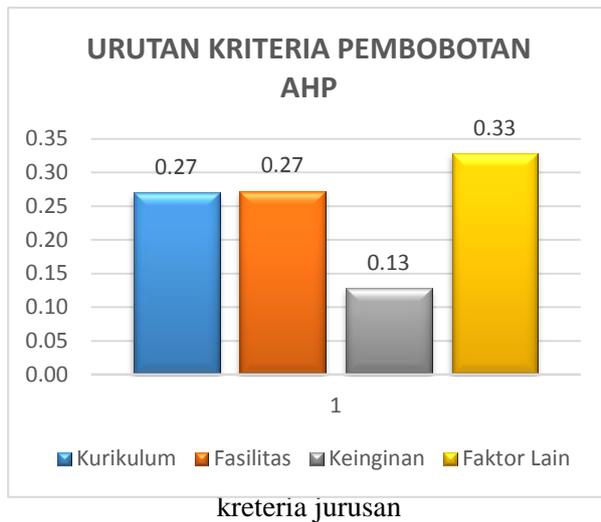
Tabel 2. Kriteria faktor lain

No	Perusahaan	Faktor Evaluasi	Faktor Bobot	Bobot Evaluasi
1	Otomotif	0.400	0.385	0.1538
2	TKJ	0.400	0.269	0.1077
3	Multimedia	0.200	0.192	0.0385

4	Bangunan	0.200	0.115	0.0231
5	RPL	0.143	0.038	0.0055
Jumlah			1.00	0.33

3. Melakukan perangkingan pada setiap bobot

Setelah melakukan ke lima tahapan di atas maka selanjutnya yaitu melakukan perangkingan pada setiap bobot. Adapun hasilnya diperlihatkan pada gambar 2 dibawah ini.



Pada gambar 2 diatas nampak bahwa setelah melakukan quisioner di SMK Harapan Bangsa dengan dataset 100 siswa, maka diperoleh data untuk parameter urutan dalam pemilihan jurusan oleh para siswa. Dengan urutan dari faktor lain menempati posisi pertama. Setelah melakukan pembobotan pada setiap urutan kreteria, dilanjutkan dengan proses perangkingan dari setiap kreteria yang ada untuk mengurutkan masalah.

Tabel 3. Pembobotan dan ranking kriteria

Kode	Kreteria	Nilai Bobot	Ranking
a	Kurikulum	0.27	3
b	Fasilitas	0.27	2
c	Keinginan	0.13	2
d	Faktor Lain	0.33	1

4.2 Analisa Klasifikasi Pemilihan Jurusan Menggunakan Metode C4.5

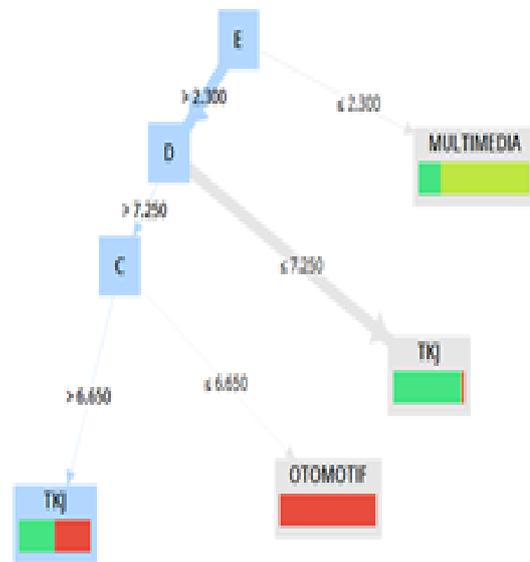
1. Penyusunan Data Penelitian

Dalam penelitian ini untuk membuat klasifikasi pemilihan jurusan menggunakan metode C4.5 dengan 100 dataset quisioner yang terdiri dari 70 data training dan 30 data testing yang diambil dari siswa kelas X di sekolah SMK Harapan Bangsa yang beralamat di jalan Neglasari no 7 RT 15 RW 4 Kec. Lengkong

Kab. Sukabumi. Yang nantinya 70 data training akan digunakan untuk pembuatan model data tree, pengujian data menggunakan 30 data testing, sementara untuk validasi data diambil dari data training yang telah diuji. Dalam data training maupun data testing sudah dikelompokan yang nantinya digunakan freemwork pada Rapid Miner untuk menghasilkan model tree nya. Adapun data training maupun data testing dapat dilihat pada lampiran.

2. Pembuatan model tree dari Data Training

Data Training ini di input ke Rapid Miner untuk membuat sebuah model tree dengan algoritma C45.



Gambar 3. Proses pembelajaran model Tree C45

Pada gambar 3 diatas nampak bahwa proses pembentukan model tree C45 menggunakan operator filter example. Hal ini karena terdapat data-data yang tidak dikenal oleh Rapid Miner, dalam kasus ini pada variable hasil terdapat missing sehingga model tree tidak dapat ditampilkan dengan benar, maka dengan ini digunakan operator filter example. Dari gambar tersebut secara berurutan di mana E adalah sebagai kreteria faktor lain, D sebagai keinginan, dan C sebagai fasilitas. Dan untuk variabel faktor lain menempati puncak tree, hal ini dikarenakan variabel tersebut mempunyai urutan pertama yang mempunyai pengaruh pada pemilihan jurusan oleh siswa.

### 3. Pengujian Data

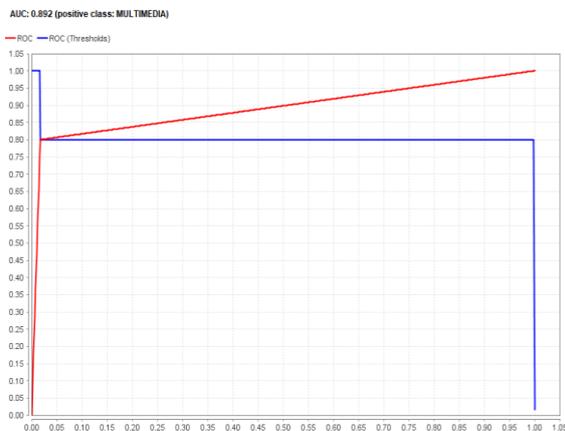
Setelah terbentuk model tree pada metode C4.5 oleh *Rapid Miner* seperti gambar di atas maka di lakukan pengujian data testing yang berjumlah 30 dari hasil quisioner di SMK Harapan Bangsa, dilakukan penyeleksian data tersebut apakah data tersebut sesuai dengan prediksi atau tidak.

accuracy: 96,88%

	true TKJ	true MULTIMEDIA	class precision
pred. TKJ	58	1	98,31%
pred. MULTIMEDIA	1	4	80,00%
class recall	98,31%	80,00%	

Dilihat dari gambar 4 diatas tercatat bahwa hasil akurasi dari data testing yakni 96,88 %, dengan presentasi indikasi kebenaran jurusan TKJ sekitar 98,31 %, indikasi kebenaran jurusan Multimedia 80 % dengan nilai masing-masing presisi klas 98,31% untuk prediksi jurusan TKJ dan 80% untuk prediksi jurusan multimedia.

### 4. Kurva ROC

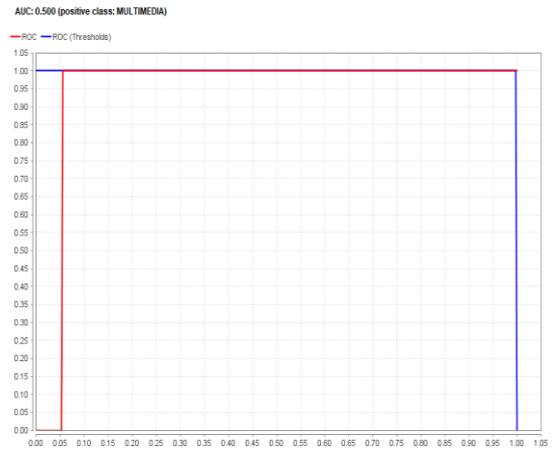


Pada gambar 5 diatas menunjukkan bahwa kurva ROC untuk data testing sekitar 0,892, hal ini dapat dikategorikan dalam *good classification*.

### 5. Validasi dari Data Testing dan Data Training

Langkah seanjutnya adalah melakukan proses validasi, di mana langkah ini bertujuan untuk menguji aplikasi data mining yang telah kita buat agar didapatkan presentase keakuratan sistem yang dibuat. Sementara tingkat akurasi dari confusion matriks yang dihasilkan sekitar 94,74%. Precision sekitar 50% dan recall sekitar 100 % Adapun tingkat akurasi AUC

yang dihasilkan dari validasi sekitar 0.500 ini berarti termasuk *good classification*. Untuk lebih jelasnya dapat dilihat pada gambar 6 di bawah ini.



Gambar 6. Tingkat akurasi yang dihasilkan dari data validasi

### 4.3 Implementasi Sistem

Pada implementasi sistem ini peneliti menggunakan *Software PHP* sebagai interface GUI (Graphic User interface). Dalam pengkodean ini peneliti menggunakan metode C4.5 dalam menganalisa pemilihan jurusan, dari inputan-inputan yang telah ditentukan user dapat mengetahui hasil analisa yang direkomendasikan oleh sistem, berikut adalah gambar 7 dibawah ini..

Gambar 7. Interface yang dibuat

#### 4.4 Pembahasan

Pengaruh faktor lain menempati urutan pertama dari hasil perankingan dengan metode AHP. Hal ini dapat dipahami karena “mungkin” pada dasarnya mereka dalam memilih pemilihan jurusan mereka mengikuti faktor lain diantaranya mengikuti ajakan teman. Sehingga dengan hal itu ada kemungkinan mereka salah mengambil jurusan sehingga pengambilan yang tidak tepat sesuai dengan sasaran. Dengan melihat kondisi tersebut maka ada kemungkinan pada pematapan ilmu pengetahuan yang kurang. Kemudian pada klasifikasi dengan metode C4.5 dilakukan proses pembentukan model tree dari data training, tingkat akurasi dari data testing, melakukan proses validasi dari ke dua data tersebut. Kurva ROC dan AUC dapat ditampilkan dengan tabel di bawah ini.

### V. PENUTUP

#### 5.1 Kesimpulan

Proses data mining menggunakan algoritma Decision Tree C4.5 yang bertujuan untuk proses klasifikasi dataset yang dibagi menjadi data training dan testing. Di mana menghasilkan model tree dari data training dan tingkat akurasi serta validasi dari data testing dengan nilai 96,88% dan 0.500% yang dikategorikan good classification.

#### 5.2 Saran

Kedepan sebaiknya digunakan metode optimasi guna meningkatkan tingkat akurasi dari pohon keputusannya.

### DAFTAR PUSTAKA

- [1] A. Muharam., ”Pengaruh Pelaksanaan Praktek Kerja Industri (PRAKERIN) Terhadap Kesiapan Siswa Bekerja Di Dunia Industri”, Bandung: 2013.
- [2] R. F. Harahap, “okezone.com,” MNC, 25 February 2014. [Online]. Available: <http://news.okezone.com/read/2014/02/24/373/945961/duh-87-mahasiswa-indonesia-salah-jurusan>. [Diakses 2017].
- [3] Anonim. Statistik Persekolahan SMK 2015/2016”, Sekertariat Jenderal Pusat Data dan Statistik Pendidikan dan Kebudayaan, Jakarta: 2016.
- [4] O.Kristanto, “Penerapan Algoritma Klasifikasi Data Mining ID3 Untuk Menentukan Penjurusan Siswa SMAN 6 Semarang”, Semarang, 2014.
- [5] L. Swatina, “Penerapan Algoritma C4.5 Untuk Penentuan Jurusan Mahasiswa” Banjarmasin: 2013.
- [6] Y. Kustiyarningsih, N. Syafa’ah, “Sistem Pendukung Keputusan Untuk Menentukan Jurusan Pada Siswa SMA Menggunakan Metode KNN dan Smart”, Madura: 2015.
- [7] R. D. H.Sugara, Agustian, I. M. Visiam, M. Farhan, D. Gustian. “Penerapan Metode C4.5 dalam Tingkat Kemampuan Bahasa Inggris Mahasiswa”, Jurnal Rekayasa Teknologi Nusa Putra. Vol. 1, No. 2, Februari 2015: Hal 39 – 43.
- [8] A. Puspita, M.Wahyudi. “Algoritma C4.5 Berbasis Decision Tree untuk Prediksi Kelahiran Bayi Prematur”, Konferensi Nasional Ilmu Pengetahuan dan Teknologi (KNIT) 2015, 8 Agustus 2015, Bekasi, Indonesia.
- [9] L. Atika, “Sistem Penunjang Keputusan Penilaian Kinerja Pemilihan Dosen Berprestasi Menggunakan Metode AHP”, Jurnal Imiah MATRIK Vol.12. No.3, Desember 2010: 1-10.
- [10] Y.Kustiyarningsih, N.Syafa’ah, “Sistem Pendukung Keputusan Untuk Menentukan Jurusan Pada Siswa SMA Menggunakan Metode KNN dan Smart”, Madura: 2015.