

## ANALISIS MINAT BELAJAR SISWA PADA LEMBAGA PENDIDIKAN PENDIDIKAN INDONESIA AMERIKA PERUMNAS 3 BEKASI DENGAN METODE C4.5

Novi Wulandari, Prilly Etikasari  
STMIK Pranata Indonesia

[Wulandari.novi2@yahoo.com](mailto:Wulandari.novi2@yahoo.com), [Prilly.etikasari@gmail.com](mailto:Prilly.etikasari@gmail.com)

*Naskah di terima 4 Maret 2019*

### ABSTRACT

*Research for decision support system is done to determine the percentage of the most dominant factors in the program offered by the American Indonesian Education Institute (LPIA) Bekasi 3, as a benchmark for selecting the most desirable programs to suit students' wants and needs at the same time provide existing program recommendations to respond to student expectations and can be used as an LPIA strategy.. The analytical method used in this study is quantitative research using the C4.5 algorithm, namely to obtain the best attributes for separating data and developing leaf nodes on decision trees according to the criteria of student learning interest with RapidMiner calculations. The population in this study are service users both children and adults with samples taken as many as 300 respondents using the sampling technique produced in training data. So that the dominant factors that are most in demand for the selection of programs that are in accordance with the wishes and needs of students can be obtained. By using the C4.5 algorithm, it can be concluded that in predicting the level of accuracy using the C4.5 algorithm, an accuracy rate of 69.17% is obtained and for algoritma C4.5 classification on the highest value vector performance chart with the Premium Application program category.*

**Keywords:** *Decision Support System, Decision Tree, c4.5 algorithm*

### ABSTRAK

Penelitian sistem informasi pendukung keputusan minat belajar siswa dilakukan untuk menentukan persentase variabel terhadap faktor yang paling dominan pada program yang ditawarkan Lembaga Pendidikan Indonesia Amerika (LPIA) Perumnas 3 Bekasi, sebagai tolak ukur terhadap pemilihan program yang paling diminati agar sesuai dengan keinginan dan kebutuhan siswa sekaligus memberikan rekomendasi program yang ada untuk merespon ekspektasi siswa dan dapat dijadikan sebuah strategi LPIA. Metode analisis yang digunakan dalam penelitian ini adalah penelitian kuantitatif dengan menggunakan algoritma C4.5 yaitu untuk memperoleh atribut terbaik untuk memisahkan data dan mengembangkan simpul daun pada pohon keputusan sesuai dengan kriteria-kriteria minat belajar siswa dengan perhitungan RapidMiner. Populasi dalam penelitian ini adalah pengguna jasa baik kalangan anak-anak maupun dewasa dengan sampel yang diambil sebanyak 300 responden menggunakan teknik sampling yang dihasilkan pada training data. Sehingga dapat diperoleh faktor dominan yang paling banyak diminati terhadap pemilihan program yang sesuai dengan keinginan dan kebutuhan siswa. Dengan menggunakan algoritma C4.5 dapat disimpulkan bahwa dalam melakukan prediksi tingkat ketepatan dengan menggunakan algoritma C4.5 didapat tingkat akurasi sebesar 69.17% dan untuk klasifikasi algoritma C4.5 pada chart performance vektor nilai tertinggi dengan kategori program Aplikasi Premium.

**Kata Kunci:** *Sistem Pendukung Keputusan, Pohon Keputusan, Algoritma c4.*

## PENDAHULUAN

Minat merupakan salah satu faktor penentu dalam keberhasilan pendidikan. Dampak dari adanya minat belajar dapat menumbuhkan metode baru dalam belajar peserta didik. Belajar dikatakan berhasil jika dapat menumbuhkan sikap, tingkah laku dan cara berfikir dalam memecahkan permasalahan-permasalahan yang dihadapi [1]. Minat belajar ini yang akan menjadi landasan utama dalam menumbuhkan keinginan belajar seseorang.

Minat belajar merupakan dorongan batin yang tumbuh dari seorang siswa untuk meningkatkan kebiasaan belajar. Minat belajar akan tumbuh saat siswa memiliki keinginan untuk meraih nilai terbaik, atau ingin memenangkan persaingan dalam belajar dengan siswa lainnya. Minat belajar juga dapat dibangun dengan menetapkan cita-cita yang tinggi sesuai dengan bakat dan kemampuan siswa [2].

Dalam pembelajaran, minat siswa untuk mau belajar adalah hal yang harus ditemukan terlebih dahulu. Minat yang muncul dalam melakukan pembelian menciptakan suatu motivasi yang terus terekam dalam benaknya dan menjadi suatu keinginan yang sangat kuat yang pada akhirnya ketika seorang siswa harus memenuhi kebutuhannya, maka siswa akan mengaktualisasikannya apa yang ada dalam benaknya itu [3].

Ada banyak sekali cara yang dilakukan siswa untuk belajar, salah satunya yaitu dengan mengikuti pendidikan non-formal yaitu kursus untuk menyesuaikan kebutuhan yang dimilikinya. Lembaga Pendidikan Indonesia Amerika (LPIA) adalah salah satu institusi pendidikan. Mereka menekankan kecakapan dan kompetensi praktis pengoperasian, dalam program-program yang

ditawarkannya LPIA telah turut serta menyediakan dan menyiapkan Sumber Daya Manusia kompeten dan kompetitif bagi kebutuhan pendidikan, dunia kerja dan industri. Dibandingkan dengan yang lain, LPIA bukan hanya mengajarkan ilmu pengetahuan kepada para siswanya, tetapi kemampuan penggunaannya pula.

Dengan adanya berbagai program-program yang ditawarkan LPIA diharapkan sebagai tolak ukur penelitian terhadap pemilihan program yang paling diminati agar sesuai dengan keinginan dan kebutuhan siswa. Sehingga dengan memahami minat siswa pula LPIA dapat merancang strategi yang tepat untuk merespon ekspektasi konsumen dan dapat dijadikan sebagai sebuah strategi LPIA dengan pesaingnya. Ada beberapa metode yang dapat digunakan untuk permasalahan penentuan minat belajar siswa, yaitu dengan Algoritma K-Nearest Neighbord (K-NN), Naive Bayes, dan Algoritma C4.5.

Algoritma K-Nearest Neighbord (K-NN) merupakan suatu metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Sehingga K-Nearest Neighbor (K-NN) merupakan salah satu metode pengklasifikasian data berdasarkan similaritas dengan label data [4].

Naive Bayes adalah suatu penerapan dari teorema Bayesian. Algoritma Naive Bayes juga didasarkan pada suatu asumsi yang digunakan untuk menyederhanakan atribut dengan mengkondisikan setiap atribut tidak terkait satu sama lain. Naive Bayes juga bisa digolongkan pengklasifikasian menggunakan aturan statistika. Naive bayes memiliki tingkat akurasi yang tergolong tinggi dan kecepatan komputasi yang relative cepat saat digunakan data berdimensi besar [5].

Algoritma C4.5 merupakan algoritma klasifikasi dengan teknik Decision tree (pohon keputusan) yang terkenal dan

disukai karena memiliki kelebihan-kelebihan. Kelebihannya adalah dapat mengolah data numerik (kontinyu) dan diskret, dapat menangani nilai atribut yang hilang, dapat menghasilkan aturan aturan yang mudah diinterpretasikan dan tercepat diantara algoritma-algoritma yang lain. Decision tree adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Dengan masing-masing rangkaian pembagian, anggota himpunan hasil menjadi mirip satu dengan yang lain [6].

Dari beberapa model metode klasifikasi yang ada menurut [7] bahwa algoritma klasifikasi decision tree merupakan algoritma klasifikasi yang memiliki tingkat akurasi paling tinggi dibandingkan algoritma klasifikasi naive bayes dan nearest neighbour. Dengan menggunakan algoritma C4.5 Pada penelitian ini, diharapkan dapat membantu peneliti dalam menggunakan aturan klasifikasi untuk menentukan minat belajar siswa dengan atribut dan record sebagai suatu parameter yang dibuat dalam kriteria pembentukan pohon keputusan terhadap pemilihan program belajar yang paling banyak diminati siswa.

## TINJAUAN PUSTAKA

### 1.1 Sistem Informasi

“Sistem informasi adalah suatu sistem di dalam suatu organisasi yang mempertemukan kebutuhan pengolahan transaksi harian, mendukung operasi, bersifat manajerial dan kegiatan strategi dari suatu organisasi dan menyediakan pihak luar tertentu dengan laporan-laporan yang diperlukan.” [8].

“Sistem informasi merupakan gabungan dari empat bagian utama. Keempat bagian utama tersebut

mencakup perangkat lunak (software), perangkat keras (hardware). Infrastruktur, dan Sumber Daya Manusia (SDM) yang terlatih. Keempat bagian utama ini saling berkaitan untuk menciptakan sebuah sistem dapat mengolah data menjadi informasi yang bermanfaat.” [9].

Berdasarkan penjelasan tersebut dapat disimpulkan bahwa definisi sistem dan informasi yang telah dijelaskan diatas, maka dapat dinyatakan bahwa sistem informasi merupakan gabungan dari empat bagian utama. Keempat bagian utama tersebut mencakup perangkat lunak (Software), perangkat keras (Hardware), infrastruktur, dan Sumber Daya Manusia (SDM) yang telah terlatih.

### 1.2 Sistem Pendukung Keputusan

Sistem Pendukung Keputusan (SPK) digunakan sebagai alat bantu bagi para pengambil keputusan untuk memperluas kapabilitas para pengambil keputusan. Menurut [10] Konsep Sistem Pendukung Keputusan pertama kali dikemukakan oleh Michael S.Scott Morton pada tahun 1970 dengan istilah Management Decision System. Konsep tersebut ditandai dengan dibuatnya sebuah sistem komputer yang interaktif dalam mengumpulkan data dan menggunakan model untuk mendukung keputusan penyelesaian masalah-masalah yang tidak terstruktur.

Secara umum SPK dapat diartikan sebagai sistem yang dimaksudkan untuk mendukung para pengambil keputusan manajerial dalam situasi keputusan semi terstruktur. SPK dimaksudkan untuk menjadi alat bantu bagi para pengambil keputusan untuk memperluas kapabilitas mereka, namun tidak untuk menggantikan penilaian mereka. SPK ditujukan untuk keputusan-keputusan yang memerlukan penilaian atau pada keputusan-keputusan yang sama sekali tidak dapat didukung oleh algoritma.

Terdapat tiga komponen Sistem Pendukung Keputusan adapun komponen tersebut sebagai berikut:

- a) Subsistem Manajemen Basis Data (Data Base Management Subsystem) merupakan subsistem data yang terorganisasi dalam suatu basis data. Data yang merupakan suatu sistem pendukung keputusan dapat berasal dari luar maupun dalam lingkungan. Untuk keperluan SPK, diperlukan data yang relevan dengan permasalahan yang hendak dipecahkan melalui simulasi.
- b) Subsistem Manajemen Basis Model (Model Base Management Subsystem), salah satu keunggulan SPK adalah kemampuan untuk mengintegrasikan akses data dan model-model keputusan. Hal ini dapat dilakukan dengan menambahkan model-model keputusan ke dalam sistem informasi yang menggunakan database sebagai mekanisme integrasi dan komunikasi diantara model-model. Karakteristik ini menyatukan kekuatan pencarian dan pelaporan data dari PDE dan pengembangan disiplin manajemen.
- c) Subsistem dialog merupakan bagian dari SPK yang dibangun untuk memenuhi kebutuhan representasi dan mekanisme kontrol selama proses analisa dalam SPK ditentukan dari kemampuan berinteraksi antara sistem yang terpasang dengan user. Pemakai terminal dan sistem perangkat lunak merupakan komponen – komponen yang terlibat dalam subsistem dialog yang mewujudkan komunikasi antara user dengan sistem tersebut. Komponen dialog menampilkan keluaran sistem bagi pemakai dan menerima masukkan dari pemakai ke dalam SPK.

Berdasarkan pengertian tersebut, dapat diambil kesimpulan bahwa Sistem Penunjang Keputusan (SPK) adalah suatu system informasi berbasis

komputer yang melakukan pendekatan untuk menghasilkan berbagai alternative keputusan untuk membantu pihak tertentu dalam menangani permasalahan dengan menggunakan data dan model, sehingga suatu SPK hanya memberikan alternative keputusan dan selanjutnya diserahkan kepada user untuk mengambil keputusan.

### 1.3 Algoritma Pohon Keputusan C4.5

Metode ini Salah satu metode data mining yang umum digunakan adalah pohon keputusan. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan rule. Pohon keputusan adalah salah satu metode klasifikasi yang paling populer karena mudah untuk diinterpretasi oleh manusia. Konsep dari pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan [11].

Berikut merupakan tahapan bagaimana suatu permasalahan dirumuskan menjadi Algoritma Pohon Keputusan C4.5.

- a) Data training di input.
- b) Hitung gain ratio, split info dan entropy dari masing-masing atribut data training yang ada. Entropi (Entropy) adalah mengukur ketidakpastian suatu variabel acak.
- c) Buat simpul akar dari pemilihan atribut yang memiliki gain ratio terbesar. Gain Ratio adalah pengukuran kehadiran dan ketidakhadiran suatu variabel untuk membuat klasifikasi yang benar dalam class .
- d) Hitung gain ratio, split info dan entropy dari masing-masing atribut dengan menghilangkan atribut yang telah dipilih sebelumnya.
- e) Buat simpul internal dari pemilihan atribut yang memiliki gain ratio terbesar. Cek apakah semua atribut sudah dibentuk pada pohon.

- f) Jika belum, maka ulangi proses d dan e, jika sudah maka lanjut pada proses berikutnya.
- g) Selanjutnya pemangkasan pohon untuk menghilangkan cabang-cabang yang tidak perlu. Kemudian aturan keputusan di generate mengikuti pohon yang telah dibentuk sebelumnya.

Berikut tampilan gambar alur flowchart pada algoritma pohon keputusan C4.5



Gambar 1. Flowchart Algoritma Decision Tree

## 2 METODE PENELITIAN

### 2.1 Variabel Penelitian

Pengguna Variabel-variabel yang akan digunakan untuk proses data mining ini ditentukan berdasarkan tujuan penelitian. Variabel dalam penelitian ini merupakan kriteria-kriteria daftar program belajar siswa LPIA Perumnas 3 yang diperoleh dari studi lapangan. Instrumen yang digunakan dengan menggunakan data berupa aplikasi file Microsoft Excel. Setiap variabel

memiliki nilai atau kategori sebagai berikut :

- a) Variabel Akademik dikelompokkan dalam dua kategori yaitu Bahasa Inggris dan Komputer.
- b) Variabel Jenis Kelamin dikelompokkan dalam 2 kategori yaitu laki-laki dan perempuan.
- c) Variabel Usia dikelompokkan dalam 4 kategori yaitu usia 6-15 tahun, usia 16-30 tahun, usia 31-45 tahun, dan usia 46-60 tahun.
- d) Variabel Program dikelompokkan dalam 8 kategori dari 2 program yaitu:
  1. Bahasa Inggris meliputi TOEFL, Conversation, Elementary, dan Basic.
  2. Komputer meliputi Programming, Graphic Desain, Aplikasi Premium, dan Aplikasi Perkantoran.
- e) Variabel Lama Belajar dikelompokkan menjadi 2 kategori yaitu 2 jam dan 1.5 jam .

Terdapat 300 sampel data konsumen yang diambil minat program belajarnya yang dibagi menjadi atribut dan beberapa kriteria yaitu :

- a) Atribut meliputi program bahasa inggris dan komputer yaitu :
  1. Bahasa Inggris meliputi TOEFL, Conversation, Elementary, dan Basic.
  2. Komputer meliputi Programming, Graphic Desain, Aplikasi Premium, dan Aplikasi Perkantoran.

- b) Kriteria meliputi jenis kelamin, usia, lama belajar, dan akademik.

Dari beberapa kategori yang dikelompokkan dari masing-masing variabel. Maka dengan penggunaan sistem klasifikasi pohon keputusan menggunakan algoritma decision tree aturan atau data dapat diketahui untuk menentukan minat belajar siswa pada program belajar yang terdapat di LPIA Perumnas 3.

## 2.2 Metode Penentu Subyek Penelitian

Tabel Pada tahap perancangan penelitian yaitu menentukan metode penentu subyek dengan penerapan metode decision tree yang digunakan untuk melihat hasil sebuah sistem pendukung keputusan yang telah diimplementasikan dengan kriteria yang telah ditentukan dalam klasifikasi minat belajar siswa pada LPIA Perumnas 3 Bekasi dengan subyek populasi dan sample penelitian sebagai berikut :

### a) Populasi Penelitian

Populasi merupakan sekumpulan subyek penentu yang menjadi pusat perhatian dalam penelitian terkandung informasi yang ingin diketahui. Subyek disebut dengan satuan analisis. Satuan analisis memiliki kesamaan perilaku atau karakter yang ingin diteliti. Populasi dalam penelitian minat belajar siswa yang memiliki kriteria:

1. Sudah pernah menjadi mitra belajar LPIA Perumnas 3 Bekasi.
2. Populasi sebanyak 300 record .

### b) Sample Penelitian

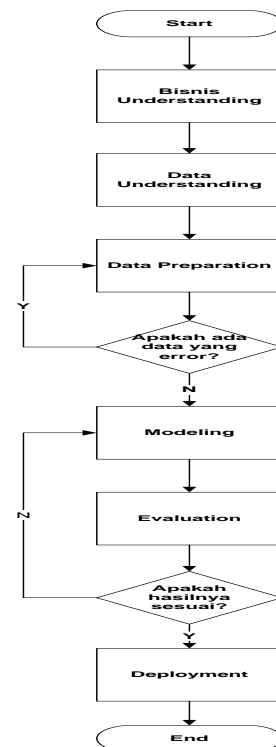
Sample penelitian dalam penelitian ini yaitu populasi yang dianggap mewakili populasi tersebut sehingga informasi yang dihasilkan oleh sample mewakili keseluruhan populasi. Dengan jumlah populasi sebanyak 300 record diambil jumlah sample yang dibagi menjadi dua sample data secara acak meliputi data training (80%) sebanyak 240 record dan data testing (20%) sebanyak 60 record.

## 2.3 Instrumen Penentuan Data

Metode Data mengalami beberapa proses pengolahan. Sebelum diterapkan algoritma data mining terhadap sebuah data-set, perlu dilakukan pengolahan data awal yang bertujuan untuk mendapatkan data-set yang dapat diolah dengan cepat dan menghasilkan kesimpulan yang tepat. Berikut Instrumen yang digunakan dalam Instrumen pengumpulan data awal yang terdiri dari :

### a) Alur Proses Data Mining

Data mining adalah proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstrak pengetahuan (knowledge) secara otomatis. Data mining berisi pencarian trend atau pola yang diinginkan dalam database besar untuk membantu pengambilan keputusan di waktu mendatang. Berikut ilustrasi alur proses data mining.



Gambar 2. Flowchart Data Mining

### b) Tahapan Data Mining

#### 1. Business Understanding

Penelitian akan dilakukan di LPIA Perumnas 3 Bekasi berdasarkan identifikasi masalah yang telah tertera pada Bab I, yang akan difokuskan untuk mencari analisa mengenai minat belajar konsumen, sesuai dengan kriteria yang dibutuhkan oleh LPIA Perumnas 3 Bekasi. Hasil akhir penilaiannya adalah atribut program yang terdapat dalam Bahasa Inggris dan Komputer

. Dan pembuatan Analisis menggunakan metode *decision tree* dengan algoritma C4.5 agar penyeleksian dapat diketahui dengan cepat dan akurat. Data yang akan digunakan dalam penelitian berupa data siswa masuk tahun periode Maret 2014 – Juli 2018.

2. Data Understanding

(a) Tahap Pengumpulan Data (Aggregation)

Proses pengumpulan data (Aggregation) adalah proses mengkombinasikan dua atau lebih atribut-atribut atau objek-objek ke dalam satu atribut tunggal atau objek. Dalam tahapan pengumpulan data dilakukan dengan mengambil data langsung ke LPIA Perumnas 3. Data yang digunakan dalam penelitian adalah data siswa yang menggunakan jasa pendidikan pada program belajar dengan melalui file daftar belajar siswa LPIA Perumnas 3.

Data yang didapat dari LPIA Perumnas 3 Bekasi adalah daftar siswa belajar tahun 2014-2015 yang akan mempengaruhi dalam klasifikasi minat belajar siswa, jumlah data sebanyak 300 record yang akan dijadikan node (akar) pada pohon keputusan yang akan dibangun. Variabel yang menjadi dasar minat belajar siswa kursus yaitu TOEFL, Conversation, Elementary, Basic, Programming, Graphic Desain, Aplikasi Premium, dan Aplikasi Perkantoran. Kemudian untuk variabel tujuan yaitu variabel yang dijadikan sebagai kriteria minat belajar. Variabelnya yaitu jenis kelamin, usia, lama belajar, dan akademik. Kandidat data tersebut didapat dari bagian administrasi dan beberapa penyesuaian agar penilaian yang dilakukan lebih akurat yang ditunjukkan pada Tabel 1.

Tabel 1. Daftar Program Belajar Siswa LPIA Perumnas 3

No	Tanggal Daftar	Nama	Tanggal Lahir	Jenis Kelamin	Usia	Program	Lama Belajar	Akademik
1	01 Januari 2017	Yuliana Hilal Rafifa Kharisma	05 November 2015	Laki-laki	12 tahun	Conversation	2 jam	Bahasa Inggris
2	12 Januari 2017	Yehmas Andalin	22 April 1993	Laki-laki	26 tahun	TOEFL	1,5 jam	Bahasa Inggris
3	13 Januari 2017	Krisna Melarani	04 April 2020	Perempuan	9 tahun	Basic	1,5 jam	Bahasa Inggris
4	14 Januari 2017	Atis Widianto	14 Mei 1994	Laki-laki	24 tahun	Conversation	2 jam	Bahasa Inggris
5	31 Januari 2017	Emilie Panjaitan	25 Desember 1994	Perempuan	23 tahun	Conversation	2 jam	Bahasa Inggris
6	09 Februari 2017	Rahelien Hermelis Ardiwang	15 April 1993	Laki-laki	20 tahun	Conversation	2 jam	Bahasa Inggris
7	10 September 2017	Hendra Kotto	05 Januari 1989	Laki-laki	29 tahun	TOEFL	1,5 jam	Bahasa Inggris
8	11 Februari 2017	Darwisya Komara	16 November 2013	Laki-laki	14 tahun	Basic	1,5 jam	Bahasa Inggris
9	11 November 2017	Mohamad Reza Achikidjaja	12 Februari 2013	Laki-laki	15 tahun	Basic	1,5 jam	Bahasa Inggris
10	28 Februari 2017	Vijaya Rakasa Andrian Saputra	14 Oktober 2010	Laki-laki	9 tahun	Basic	1,5 jam	Bahasa Inggris
↓	↓	↓	↓	↓	↓	↓	↓	↓
299	12 Mei 2018	Ma Yullanto	30 Juli 2001	Laki-laki	17 tahun	Aplikasi Perkantoran	1,5 jam	Komputer
300	29 Juni 2018	Kebrian Dika Prabhita	10 Februari 1989	Perempuan	29 tahun	Aplikasi Perkantoran	1,5 jam	Komputer

(b) Tahap Pembersihan Data (Cleaning Data)

Dalam pembersihan data tidak semua atribut dari data digunakan, hanya atribut-atribut yang dianggap penting yang digunakan, untuk itu perlu dilakukan proses pembersihan data. Pada proses ini dilakukan pemeriksaan data, menghilangkan redundansi data, memperbaiki kesalahan data dan menghilangkan data yang tidak diperlukan sehingga menyisakan data yang penting yang akan digunakan untuk tahap selanjutnya.

Agar mudah diolah data rekapan yang masih merupakan data mentah harus diolah kembali berdasarkan klasifikasi minat belajar siswa yang telah ditentukan pada tahap sebelumnya kedalam Microsoft Excel.

Pada file daftar belajar siswa yang telah dikumpulkan dilakukan cleaning data, yaitu dengan menghilangkan record-record yang noise atau tidak lengkap dan record yang berulang, atribut yang tidak diperlukan juga bisa dihilangkan seperti atribut Tanggal Daftar dan Tanggal Lahir. Hal ini dilakukan karena atribut tersebut tidak berpengaruh terhadap pengolahan data pada proses konsep entropy reduction dan information gain.

(c) Tahap Transformasi Data

Pada tahap ini merupakan tahap data yang disederhanakan atau dikelompokkan. Data yang digunakan dalam penelitian ini adalah data dari



hasil daftar program belajar siswa LPIA Perumnas 3 yang dilakukan pada tahun 2014 sampai 2018. Sumber data berupa aplikasi file *Microsoft Excel* dengan jumlah populasi sebanyak 300 record data siswa. Data akan di transformasikan untuk mempermudah dalam pembacaan proses *data mining*.

(d) Tahap *Training Data*

Dalam penelitian ini data yang digunakan sejumlah 300 record, yang kemudian dibagi menjadi dua bagian yaitu sebagai *data training* dan *data testing*. *Persentase* untuk pembagian *data training* dan *data testing* yang digunakan yaitu 80 : 20, yang menghasilkan 240 data sebagai *data training*.

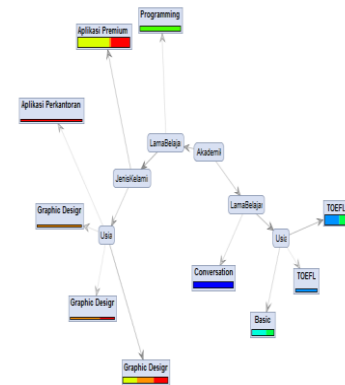
Hasil perhitungan gain dan Entropy pada data training maka di dapat hasil yang diuraikan dalam bentuk tabel. Berikut uraian hasil perhitungan yang didapat pada tabel 2.

Tabel 2. Data Perhitungan Node 1

NOE	KATEGORI	Banyak Kelas	TOEFL	Conersation	Reading	Basic	Programming	Graphic Design	Aplikasi Premium	Aplikasi Perkantoran	Contoh	Gain
	TOTAL	240	39	32	28	21	20	19	45	36	2.924	
	Jenis Kabin										2.953	0,025
	Pesepanan	119	17	14	16	6	4	2	35	21	2.163	
1	Ubi	65	3	1	11	20	2	0	1	5	1.994	0,542
	6530 ubi	100	34	25	17	1	9	14	37	25	2.511	
	3245 ubi	39	4	7	0	0	0	2	5	4	2.444	
	4540 ubi	1	0	1	0	0	0	3	2	2	1.063	
	Lama Belajar										1.812	0,570
	2.0m	12	0	2	0	0	0	0	0	0	1.812	
	1.5.0m	108	39	0	26	21	0	19	45	36	2.325	
	Alat/Perk										1.855	0,599
	Basic	120	39	32	28	21	0	0	0	0	1.855	
	Komputer	120	0	0	0	0	20	19	45	36	1.855	
											1.855	0,599

Data Perhitungan Node 1 merupakan nilai gain terbesar dan terlihat atribut akademik mempunyai nilai tertinggi yaitu 0,9998 sehingga akan dijadikan node (akar) pertama pada pohon

keputusan minat belajar siswa. Berikut ilustrasi Pohon Keputusan Node 1.



Gambar 3. Pohon Keputusan Node 1

(e) Tahap *Testing Data*

Pembagian data yang digunakan *data testing* yaitu 20%, yang menghasilkan sebagai *data testing* yaitu 60 record.

3. *Data Preparation*

Setelah tahap pengolahan data, perlu dilakukan tahap persiapan (preprocessing), dengan tujuan untuk mencari field data yang kosong (missing value) dan data dengan format yang tidak sesuai (noise), agar akurasi dari hasil pengolahan data dapat meningkat.

4. *Modelling*

Fase ini menentukan tehnik data mining yang digunakan, menentukan tools data mining, tehnik data mining, algoritma data mining, menentukan parameter dengan nilai yang optimal.

Pada penelitian ini tehnik data mining yang akan digunakan adalah klasifikasi dan algoritma yang akan digunakan adalah C4.5. Langkah berikutnya yaitu mengimplementasikan data training yang telah dipersiapkan kedalam algoritma C4.5 untuk pembuatan pohon keputusan. Pengolahan data akan dilakukan



dengan bantuan rapidminer. Setelah dilakukan pemrosesan data training sehingga menghasilkan beberapa aturan dan akan membentuk sebuah pohon keputusan pada algoritma C4.5 .

5. Evaluation

Pada tahap evaluasi merupakan tahap interpretasi terhadap permodelan data mining berdasarkan pemahaman klasifikasi yang terbentuk sesuai sasaran yang dicapai.

**HASIL DAN PEMBAHASAN**

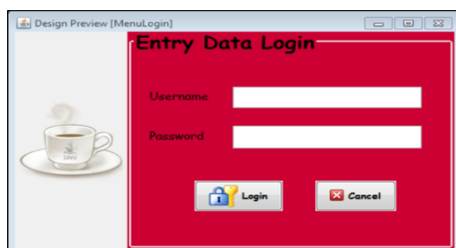
**2.4 Evaluation**

Dapat diambil kesimpulan bahwa untuk prediksi Conversation, akurasi model mencapai 100% , prediksi TOEFL mencapai 67.86% , prediksi Basic mencapai 62.50% , prediksi Elementary mencapai 0% , prediksi Programmer mencapai 100% , prediksi Aplikasi Premium mencapai 60.34% , prediksi Graphic Design mencapai 44.74% , sedangkan prediksi Aplikasi Perkantoran mencapai 100% dan nilai akurasi rata-ratanya mencapai 69.17%.

**4.2 Program Aplikasi**

1) Implementasi Menu Login

Tampilan menu login merupakan halaman awal yang dapat diakses pengguna dari aplikasi minat belajar konsumen. Pada menu login pada sistem. Tampilan menu login digunakan untuk admin atau user dalam melakukan login guna mengerjakan Create, Read, Update, Delete (CRUD) semua data. Login hanya dapat dilakukan oleh seorang user atau admin yang mempunyai akses masuk . Berikut uraian yang terdapat pada Gambar 4.



Gambar 4. Implementasi Antarmuka Menu Login

2) Implementasi Menu Minat Belajar Konsumen

Tampilan menu minat belajar konsumen merupakan halaman yang hanya bisa diakses setelah melakukan login. Fungsinya untuk CRUD minat belajar konsumen. Dengan penginputan data yang disesuaikan pada kebutuhan yang didapat pada data-set dan disesuaikan dengan hasil rule model pada rapid miner . Berikut uraian yang terdapat pada Gambar 5.



Gambar 5. Implementasi Minat Belajar Konsumen

**PENUTUP**

**Kesimpulan**

Berdasarkan pada hasil penelitian yang berlangsung pada analisis data dan fakta selama proses penelitian pada Lembaga Pendidikan Indonesia Amerika (LPIA) Perumnas 3 Bekasi, maka diperoleh kesimpulan sebagai berikut :

- a) Berdasarkan hasil penilaian untuk nilai akurasi klasifikasi algoritma C4.5 senilai 69.17% dan untuk klasifikasi algoritma C4.5 pada chart performance vektor dengan kategori program Aplikasi Premium. Klasifikasi tersebut akan menghasilkan rules sebagai dasar penentuan minat belajar siswa pada

- Lembaga Pendidikan Indonesia Amerika (LPIA) Perumnas 3 Bekasi.
- b) Rules yang didapatkan dari pohon keputusan, diimplementasikan ke dalam sebuah sistem pendukung keputusan, yang dapat membantu untuk mempermudah pengambilan keputusan terhadap minat belajar siswa yang bersangkutan dalam pemilihan program yang ditawarkan.
- c) Dengan adanya penerapan data mining algoritma C4.5 di harapkan mampu memberikan solusi perbaikan dalam kriteria minat belajar siswa guna memudahkan merancang strategi yang tepat untuk merespon ekspektasi siswa pada Lembaga Pendidikan Indonesia Amerika (LPIA) Perumnas 3 Bekasi.

#### DAFTAR PUSTAKA

- [1] S. P. Astuti, “Pengaruh Kemampuan Awal dan Minat Belajar Terhadap Prestasi Belajar Fisika,” *J. Form.*, vol. 5, no. 1, p. 69, 2015.
- [2] M. Gusniwati, “Pengaruh Kecerdasan Emosional Dan Minat Belajar Terhadap Penguasaan Konsep Matematika Siswa SMAN di Kecamatan Kebon Jeruk,” *J. Form.*, vol. 5, no. 1, p. 32, 2015.
- [3] Y. U. Salim and S. Widaningsih, “Analisis Minat Beli Konsumen Terhadap Program Kursus Bahasa Inggris Di Lembaga Kursus Bahasa Northern Lights Education Center ( NLEC ) Bandung,” *e-Proceeding Manag.*, vol. 3, no. 2, p. 523, 2017.
- [4] A. Rohman, “Komporasi Metode Klasifikasi Data Mining Untuk Prediksi Penyakit Jantung,” *J. Neo Tek.*, vol. 2, no. 2, p. 23, 2016.
- [5] Y. P. Astuti, U. Sudibyoy, A. W. Kurniawan, and Y. Rahayu, “Algoritma Naive Bayes Dengan Fitur Seleksi Untuk Mengetahui Hubungan Variabel Nilai dan Latar Belakang Pendidikan,” *J. Simetris*, vol. 9, no. 1, p. 599, 2018.
- [6] A. N. Wahyono, “Penerapan Algoritma C4.5 Untuk Prediksi Tingkat Kompetensi Karyawan PT Multistrada Arah Sarana,” *Joutica*, vol. 3, no. 1, p. 1, 2018.
- [7] D. Sartika and D. I. Sensuse, “Perbandingan Algoritma Klasifikasi Naive Bayes , Nearest Neighbour , dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian,” *Jatiji*, vol. 1, no. 2, p. 159, 2017.
- [8] Jogianto, *Analisis dan Desain Sistem Informasi Pendekatan Terstruktur, Teori dan Praktek Aplikasi Bisnis*. Yogyakarta: CV. Andi Offset, 2017.
- [9] I. P. A. E. Pratama, *Sistem Informasi dan Implementasi Teori dan Konsep Sistem Informasi Disertai Berbagai Contoh Praktiknya Menggunakan Perangkat Lunak*. Bandung: Informatika Bandung, 2014.
- [10] I. N. A. A. Dwijayadi, I. M. A. Wirawan, and D. G. H. Divayana, “Pengembangan Sistem Pendukung Keputusan Penentuan Hotel Di Kecamatan Buleleng Dengan Metode Analytic Hierarchy Process ( AHP ) Dan Technique for Others Reference By Similarity To Ideal Solution ( TOPSIS ),” vol. 7, no. 1, p. 3, 2018.
- [11] Y. Kustiyahningsih and E. Rahmanita, “Aplikasi Sistem Pendukung Keputusan Menggunakan Algoritma C4.5 untuk Penjurusan SMA,” *J. Simantec*, vol. 5, no. 2, pp. 101–108, 2016.