

# Analisa Kinerja Algoritma C4.5 Dalam Menentukan Pola Dominasi Mainstream Mahasiswa

Dewi Anggraeni, Tika Christy

Sistem Informasi, Sekolah Tinggi Manajemen Informatika dan Komputer Royal, Kisaran, Indonesia  
Email: <sup>1</sup>anggraeni1987@gmail.com, <sup>2\*</sup>tikaja18@gmail.com

## Abstrak

Program studi adalah kesatuan kegiatan pendidikan dan pembelajaran yang memiliki kurikulum dan metode pembelajaran. Setiap program studi mempunyai kurikulum, yang dimana kurikulum adalah seperangkat rencana pengaturan mencapai capaian pembelajaran lulusan, bahan kajian, proses dan penilaian yang digunakan sebagai pedoman penyelenggaraan program studi. Mainstream matakuliah pilihan adalah suatu pilihan matakuliah pilihan yang dapat menjadi peminatan mahasiswa dalam menentukan minat dan bakat. Algoritma C4.5 adalah salah satu metode algoritma yang memiliki decision tree atau pohon keputusan, dengan memanfaatkan data dari akademik tentang pemilihan mainstream matakuliah. Berdasarkan hasil penelitian, analisa algoritma C4.5 dalam menentukan pola dominasi mahasiswa menunjukkan bahwa 73% dapat dijadikan acuan sebagai bahan keputusan dalam menentukan mainstream matakuliah pilihan pada mahasiswa.

**Kata Kunci:** Mainstream Matakuliah, Data Mining, Algoritma C4.5, STMIK Royal

## Abstract

The study program is a unit of education and learning activities that has a curriculum and learning method. Each study program has a curriculum, in which the curriculum is a set regulatory plans to achieve graduate learning outcomes, study material, processes and assessments that are used as guidelines for the implementation of study program. The main stream of elective and talents. C4.5 algorithm is one method of algorithm that has a decision tree or decision tree, by utilizing data from academics about the selection of the mainstream courses. Bases on the result of the study, the analysis of C4.5 algorithm in determining the pattern of student dominance shows the 73% can be used as a reference material in determining student interests and talents

**Keywords:** Mainstream Courses, Data Mining, C4.5 Algorithm, STMIK Royal

## 1. PENDAHULUAN

Mainstreams matakuliah adalah suatu pilihan peminatan matakuliah yang diterap kepada mahasiswa di STMIK Royal Kisaran. Tujuan penentuan mainstream matakuliah ini adalah agar mahasiswa dapat mengetahui minat dan bakat yang dimiliki selama masa perkuliahan dan dapat menjadi acuan untuk menyelesaikan skripsi atau tugas akhir. Berdasarkan data yang peneliti peroleh berupa nilai hasil study mahasiswa semester 1 sampai dengan semester 5.

Pada Kondisi mahasiswa saat ini kurangnya kepercayaan diri atas kemampuan minat dan bakat yang dimiliki dalam menentukan mainstream matakuliah. Dalam hal ini [1] banyak kasus dijumpai bahwa jurusan yang tidak sesuai dengan kemampuan, kepribadian, minat dan bakat dapat mempengaruhi mahasiswa dalam mengikuti perkuliahan. [2] dampak pada ketidaksesuaian bidang peminatan dengan minat dan keterampilan mahasiswa tersebut, akibatnya banyak mahasiswa yang mengalami kesulitan ketika menyelesaikan tugas akhir. Sistem adalah suatu kumpulan atau himpunan unsur atau variabel yang saling berhubungan, berinteraksi, dan ketergantungan antara satu unsur dengan unsur yang lainnya. Database adalah suatu data dari beberapa tabel yang berisikan data [3]. Untuk membuat rencana studi, mahasiswa perlu mengetahui hubungan keterkaitan materi antar matakuliah [4]. Dengan demikian mahasiswa bias lebih berfokus kepada minat dan bakatnya.

Tujuan pada penelitian ini adalah untuk mengetahui kinerja model Algoritma C4.5 dalam pengolahan informasi data mahasiswa menjadi pohon keputusan, selain itu tujuan mengklasifikasikan yaitu untuk menemukan model dari training set yang membedakan atribut kedalam kategori atau kelas yang sesuai, model tersebut kemudian digunakan untuk mengklasifikasikan atribut yang kelasnya belum diketahui sebelumnya [5]. Model algoritma C4.5 dijadikan sebagai acuan atau dasar pengambilan keputusan pada pemilihan mata kuliah pilihan untuk mahasiswa.

Algoritma C4.5 adalah salah satu algoritma yang memiliki decision tree, dengan memanfaatkan data dari akademik tentang pemilihan mainstream matakuliah, terdapat 4 mainstream matakuliah yaitu mainstream logika dan pemrograman, mainstreams Basis data dan Manajemen Informasi. Selain itu algoritma C4.5 merupakan algoritma yang digunakan untuk membuat sebuah keputusan dalam bentuk pohon keputusan dengan menetapkan kriteria-kriteria yang menjadi dasar pembentukan keputusan [6]

Tahap pelaksanaan penelitian dapat dilihat pada alur State of the art Penelitian dan Road Map Penelitian dibawah ini :



Pada tahap ini adalah awal dari penelitian, dimulai dari objek peneliti. Observasi lakukan secara dilaksanakan di kantor prodi. Proses yang dilakukan data dari prodi STMIK berupa data nilai mahasiswa, data pemetaan matakuliah, data mahasiswa, data matakuliah, data mainstream matakuliah.

## 2. Kajian Metode

Setelah melakukan observasi dan studi pendahuluan, maka dilakukan kajian metode. Tujuan Kajian metode dilakukan untuk mencari metode yang sesuai dengan hasil observasi dan Studi pendahuluan. Berdasarkan hasil wawancara dan data akademik mahasiswa.

## 3. Tahapan Menentukan Tujuan

Berdasarkan kajian metode yang dilakukan, tahap selanjutnya menentukan tujuan. Tahap tujuan penelitian ini untuk memperjelas yang menjadi sasaran dalam penelitian. tujuan dari penelitian ini adalah bahwa hasil penelitian ini dapat menjadi informasi awal untuk mahasiswa guna mengetahui matakuliah yang akan dipilih pada semester 6.

## 4. Pengumpulan Data dan Pengujian

Proses pengumpulan data dilakukan di STMIK Royal Kisaran. Data yang akan digunakan dalam penelitian ini adalah data mahasiswa yang berupa data nilai hasil studi semester 1 sampai dengan semester 5, data matakuliah, data pemetaan matakuliah. Selanjutnya data akan diproses untuk data training dan data testing, Untuk pengujian data training dan data testing menggunakan algoritma C4.5, sedangkan untuk implementasi menggunakan software aplikasi rapid miner.

# 3. ANALISA DAN PEMBAHASAN

## 3.1 Analisa Data

Analisa Data dilakukan pada data mahasiswa STMIK Royal dari mulai semester 1 sampai dengan semester 5. Data penelitian ini adalah data berdasarkan data pemilihan mainstream mahasiswa program studi sistem informasi tahun ajaran 2016/2017 sebanyak 100 data.

## 3.2 Selection Data

Sejumlah data yang perlu dilakukan sebelum tahap seleksi informasi. Untuk menentukan masa studi mahasiswa dilakukan analisa algoritma C4.5 untuk proses analisa data sebelum data digunakan di data mining [5]

### a. Data Matakuliah

Berdasarkan kurikulum program studi sistem informasi, yang diterapkan didalam pendidikan tinggi STMIK Royal, terdapat 56 matakuliah dan 144 sks, dimana setiap semester mahasiswa diwajibkan mengambil matakuliah maksimal 9 matakuliah dengan jumlah sks maksimal 24 sks.

**Tabel 1.** Matakuliah Semester 1

No.	Kode MK	Nama Matakuliah	SKS
1	SIKK101	Pengantar Teknologi Informasi	3
2	SIKK102	Dasar Akuntansi	2
3	SIKK103	Program Aplikasi	2
4	SIKK104	Algoritma Dan Pemrograman	2
5	SIKK105	Praktek Algoritma Dan Pemrograman	2
6	SIKK106	Kalkulus	2
7	SIKK101	Bahasa Inggris	2
8	SIPK101	Pendidikan Agama	2
9	SIPK102	Pendidikan Kewarganegaraan	2

**Tabel 2.** MataKuliah Semester 2

No.	Kode MK	Nama Matakuliah	SKS
1	SIPB202	Bahasa Inggris 2	2
2	SIKK207	Matematika Diskrit	2
3	SIPB203	Bahasa Indonesia	2
4	SIKK208	Konsep Sistem Informasi	3
5	SIKK209	Design Grafis	2

6	SIKB201	Sistem Basis Data	3
7	SIKK211	Program Aplikasi Akuntansi	2
8	SIPK203	Kepemimpinan Dan Keterampilan Interpersonal	2
9	SIKK212	Manajemen Umum	2

**Tabel 3.** MataKuliah Semester 3

No.	Kode MK	Nama Matakuliah	SKS
1	SIKK310	Struktur Data	3
2	SIKK313	Sistem Berkas / Arsip dan Akses	3
3	SIKK314	Komunikasi Data	2
4	SIKK315	Sistem Operasi	3
5	SIKK316	Interaksi Manusia dan Komputer	2
6	SIKK317	Desain Web	3
7	SIKK318	Pemrograman Berorientasi Objek	3
8	SIKK319	Pengetahuan Bisnis	2

**Tabel 4.** MataKuliah Semester 3

No.	Kode MK	Nama Matakuliah	SKS
1	SIKB402	Analisis dan Perancangan	4
2	SIKK420	Statistik	3
3	SIKB403	Manajemen Sistem Basis Data	3
4	SIKK421	Sistem Informasi Manajemen	3
5	SIKB404	Analisis Proses Bisnis	2
6	SIKK422	Jaringan Komputer	3
7	SIPB404	Perilaku Organisasi	2

#### b. Data Penelitian

Data penelitian yang digunakan pada penelitian ini adalah nilai data mahasiswa semester 1 sampai dengan semester 5. Berikut ini data penelitian dengan jumlah mahasiswa 100.

**Tabel 5.** Data Penelitian

NIM	Nama Mahasiswa	IPK Semester 1	IP Semester 2	IP Nilai Semester 3	IP Nilai Semester 4	IP Nilai Semester 5	Jenis Kelamin	Mainstream
15220241	ADE ZAHWANA PUTRA	2.56	3.58	3.98	3.60	2.58	Laki-laki	Mobile Programming
15220254	ADI GUNAWAN NASUTION	2.7	3.98	3.60	3.58	2.59	Laki-laki	Mobile Programming
15220262	AL KHUDRI	2.56	3.78	3.58	3.58	2.98	Laki-laki	Mobile Programming
15220266	ARUM RUSMALIAH DEWI	2.45	3.78	3.58	3.98	2.14	Laki-laki	Mobile Programming
15220267	AYU WULANDARI	2.12	3.65	3.98	3.78	2.69	Laki-laki	Mobile Programming
15220257	DENY YULIA	2.36	3.5	3.78	3.78	2.58	Laki-laki	Mobile Programming
15220264	DESY KARLINA ZEIN	2.58	3.56	3.78	3.65	2.59	Laki-laki	Mobile Programming
15220246	DEVI PURWANINGSIH	2.59	3.59	3.65	3.5	2.98	Laki-laki	Mobile Programming
15220253	DEWI LESTARI	2.98	3.87	3.5	3.60	2.14	Laki-laki	Mobile Programming
15220335	EKA DEWI SARTIKA	2.14	3.87	3.56	3.58	2.69	Laki-laki	Mobile Programming
15220247	INDRIYANI	2.69	3.98	3.5	3.58	2.58	Laki-laki	Mobile Programming
15220270	ITA	2.01	3.55	3.56	3.98	2.59	Laki-laki	Mobile Programming
15220482	JULIANI SIAHAAN	2.63	3.66	3.59	3.78	2.98	Laki-laki	Mobile Programming
15220272	MUTIA DANI	2.74	3.58	3.87	3.78	2.14	Laki-laki	Mobile Programming
15220316	NUR HAKIKI	2.48	3.58	3.87	3.65	2.69	Laki-laki	Mobile Programming
15220252	NURAFNI IRVANI	2.63	3.98	3.98	3.3	2.69	Laki-laki	Analisis dan Perancangan E-Bisnis
15220245	NURLELI	3.83	3.78	3.54	3.60	2.69	Laki-laki	Analisis dan Perancangan E-Bisnis
15220261	RIZKY WAHYUNI	3.55	3.78	3.59	3.58	2.69	Laki-laki	Analisis dan Perancangan E-Bisnis

### 3.3 Transformasi

Sebelum melakukan proses data mining, Transformasi data terlebih dahulu untuk dipilih. Proses transformasi dilakukan untuk mencari basis data yang akan digunakan. Proses transformasi dilakukan dengan cara membuat klasifikasi pada masing-masing atribut. Sebelum melakukan transformasi data, terlebih dahulu dilakukan proses cleaning data. klasifikasi atribut nilai terbagi menjadi 3 bagian, yaitu :

**Tabel 6.** Klasifikasi Nilai

No.	IPK Semester	Kelas
1	IPK $\geq$ 3.5	Tinggi
2	IPK $\geq$ 2.75	Sedang
3	IPK $<$ 2.75	Kecil

Berdasarkan hasil transformasi pada klasifikasi nilai table 6 dapat disimpulkan seperti terlihat pada tabel 7 berikut ini :

**Tabel 7.** Transformasi Data

IP Semester 1	IP Semester 2	IP Nilai Semester 3	IP Nilai Semester 4	IP Nilai Semester 5	Jenis Kelamin	Mainstream
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Mobile Programming
kecil	tinggi	tinggi	tinggi	sedang	Laki-laki	Analisis dan Perancangan E-Bisnis
tinggi	tinggi	tinggi	tinggi	sedang	Laki-laki	Analisis dan Perancangan E-Bisnis
tinggi	tinggi	tinggi	tinggi	sedang	Laki-laki	Analisis dan Perancangan E-Bisnis
tinggi	tinggi	tinggi	tinggi	sedang	Laki-laki	Analisis dan Perancangan E-Bisnis
tinggi	tinggi	tinggi	tinggi	sedang	Laki-laki	Analisis dan Perancangan E-Bisnis
tinggi	tinggi	tinggi	tinggi	sedang	Laki-laki	Analisis dan Perancangan E-Bisnis
tinggi	tinggi	tinggi	tinggi	sedang	Laki-laki	Analisis dan Perancangan E-Bisnis
tinggi	tinggi	sedang	tinggi	tinggi	Laki-laki	Analisis dan Perancangan E-Bisnis
tinggi	tinggi	sedang	tinggi	tinggi	Perempuan	Analisis dan Perancangan E-Bisnis

### 3.4 Data Mining

Data mining adalah suatu proses yang dilakukan untuk mencari pola atau informasi dengan menggunakan beberapa metode data mining.

#### Algoritma C4.5

Algoritma C4.5 adalah salah satu algoritma yang memiliki decision tree, dengan memanfaatkan data dari akademik tentang pemilihan mainstream matakuliah, terdapat 3 mainstream matakuliah yaitu mainstream mobile dan programming dan data warehouse, analisa dan perancangan E-Bisnis. [1] algoritma C4.5 dalam penerapannya diharapkan mampu menjadi alat pendukung keputusan. Tujuan Analisa kinerja algoritma C4.5 adalah dapat menjadi keputusan dalam menentukan mainstream matakuliah pilihan [7]. Data Training adalah data yang akan diolah dalam pembentukan algoritmanya yang kemudian akan diuji dengan data testing yang hasilnya berupa pohon keputusan (Decision Tree). [1] Algoritma C4.5 akurat diterapkan untuk penentuan kesesuaian jurusan mahasiswa dengan tingkat keakuratan 93,31 % dan akurasi rekomendasi jurusan sebesar 82,64%.

Berdasarkan data transformasi, selanjutnya akan dilakukan analisa untuk menghasilkan sebuah pohon keputusan dengan menggunakan algoritma C4.5, secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

1. Mempersiapkan data training. Data training didapat dari data history yang pernah terjadi sebelumnya dan sudah dikelompokkan dalam kelas-kelas tertentu.
2. Menghitung akar dari pohon. Akar akan diambil dengan cara menghitung nilai gain dari masing-masing atribut. Nilai gain yang paling tinggi akan menjadi akar yang pertama. Sebelum menghitung nilai gain dari atribut, terlebih dahulu hitung nilai entropy.

$$Entropy(S) = \sum_{i=1}^n P_i \log_2 P_i$$

Keterangan :  
 S = Himpunan Kasus  
 n = Jumlah Partisi S  
 P<sub>i</sub> = Proporsi S<sub>i</sub> Terhadap S

Entropy (IP Semester 1)

$$= \left(-\frac{59}{100} * \log_2 \frac{59}{100}\right) + \left(-\frac{30}{100} * \log_2 \frac{30}{100}\right) + \left(\frac{11}{100} * \log_2 \frac{11}{100}\right)$$

$$= 0.814704$$

3. Kemudian hitung nilai gain dengan menggunakan rumus sebagai berikut :

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \left|\frac{S_i}{S}\right| Entropy(S_i)$$

Keterangan :  
 S =Himpunan kasus  
 A =Fitur  
 n =Jumlah Partisi  
 |S<sub>i</sub>|=Proporsi S<sub>i</sub> terhadap S  
 |S|=Jumlah Kasus Dalam S

Gain (IP Semester 1)

$$= 0.814704 - \left(\frac{50}{100} * 0.691312\right) + \left(\frac{30}{100} * 0.774785\right) + \left(\frac{20}{100} * 0.75\right)$$

$$= 0.851483$$

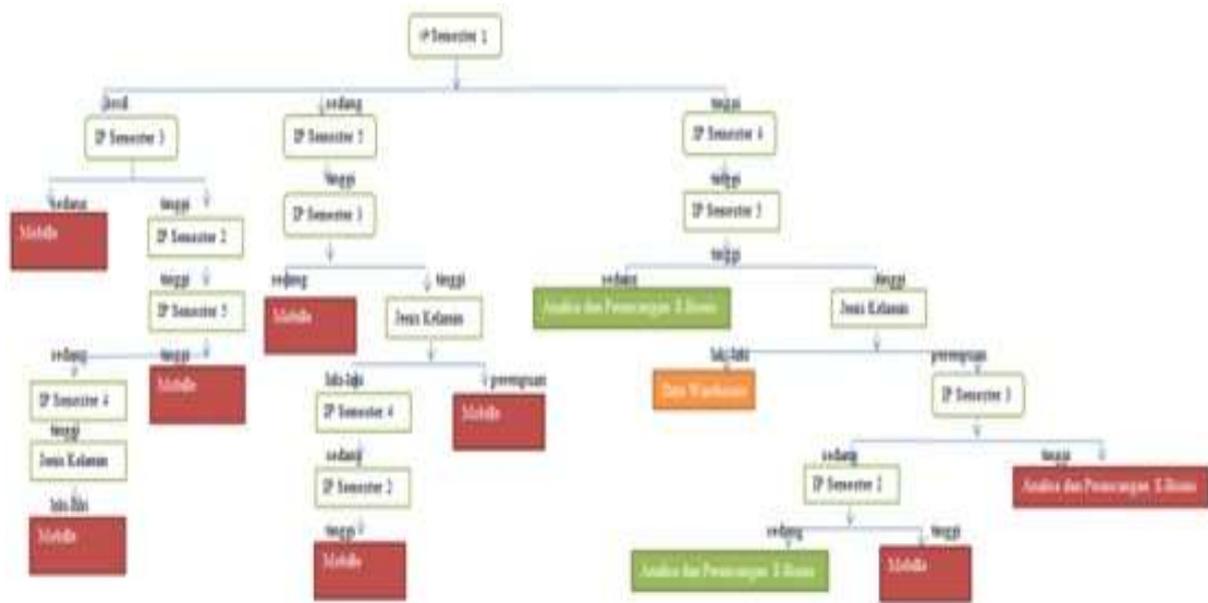
4. Merancang Decision tree menggunakan algoritma C4.5.

Berdasarkan tabel data penelitian, maka langkah selanjutnya adalah menentukan *Entropi* dan *Gain* untuk menentukan akar atau node. Berikut ini adalah hasil perhitungan entropi dan gain pada node 1. Terlihat pada tabel 8.

**Tabel 8. Node 1**

Atribut	Kelas	Jumlah Kasus	NMP	NDW	NAP	Entropi	Gain
		100	59	30	11	0.814704	
IP Semester 1	Tinggi	50	20	20	10	0.691312	0.851483
	Sedang	30	15	10	5	0.774785	
	Kecil	20	10	5	5	0.75	
IP Semester 2							0.762067
	Tinggi	57	20	27	10	0.741499	
	Sedang	43	20	20	3	0.860508	
IP Semester 3	Kecil	0	0	0	0	0	0.753142
	Tinggi	58	25	25	8	0.758295	
	Sedang	42	20	20	2	0.900593	
IP Semester 4	Kecil	0	0	0	0	0	0.244911
	Tinggi	85	25	45	10	0.781685	
	Sedang	15	5	5	5	0.63093	
IP Semester 5	Kecil	0	0	0	0	0	0.366484
	Tinggi	78	35	35	8	0.807467	
	Sedang	22	10	10	2	0.825477	
Jenis Kelamin	Kecil	0	0	0	0	0	0.921732
	Laki-laki	49	20	20	9	0.706578	
	Perempuan	51	20	27	4	0.888728	

Sesuai dengan tabel diatas, terlihat bahwa atribut gain tertinggi adalah IP Semester 1 yaitu 0.85148. lakukan perhitungan yang sama sehingga dapat menemukan nilai entropi dan gain selanjutnya. Maka atribut IP Semester 1 dapat dijadikan akar atau Node 1. Berdasarkan hasil perhitungan nilai entropi dan gain pada node 1 dapat digambarkan *decision tree* atau *pohon keputusan* secara manual seperti gambar dibawah ini :



**Gambar 4. Pohon Keputusan**

#### 4. KESIMPULAN

Algoritma C4.5 adalah salah satu metode algoritma yang memiliki decision tree atau pohon keputusan, dengan memanfaatkan data dari akademik, mengenai pemilihan mainstream matakuliah. Berdasarkan hasil penelitian, analisa algoritma C4.5 dalam menentukan pola dominasi mahasiswa menunjukkan bahwa 73% dapat dijadikan acuan sebagai bahan keputusan dalam menentukan mainstream matakuliah pilihan pada mahasiswa. Dari hasil penelitian ini, mahasiswa mengetahui minat dan bakat sesuai dengan nilai matakuliah. Sehingga mahasiswa untuk menyelesaikan tugas akhir atau skripsi bisa tepat pada waktunya atau lebih cepat. Setiap program studi mempunyai kurikulum, yang dimana kurikulum adalah seperangkat rencana pengaturan mencapai capaian pembelajaran lulusan, bahan kajian, proses dan penilaian yang digunakan sebagai pedoman penyelenggaraan program studi. Mainstream matakuliah pilihan adalah suatu pilihan matakuliah pilihan yang dapat menjadi peminataan mahasiswa dalam menentukan minat dan bakat.

#### 5. UCAPAN TERIMA KASIH

Ucapan terima kasih sebesar-besarnya atas pendanaan penelitian dan publikasi yang dibiayai oleh Direktorat Riset dan Pengabdian Masyarakat Direktorat Jenderal Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi dan Pendidikan Tinggi sesuai dengan Kontrak Penelitian Tahun Anggaran 2019.

#### REFERENCES

- [1] L. Swastina, "Penerapan Algoritma C4.5 Untuk Penentuan Jurusan Mahasiswa," vol. 2, no. 1, 2013.
- [2] F. Nasari, S. Informasi, P. Keputusan, and D. Selection, "Penerapan algoritma c4.5 dalam pemilihan bidang peminatan program studi sistem informasi di stmik potensi utama medan," pp. 30–34, 2014.
- [3] D. Anggraeni, A. Z. Syah, and S. Informasi, "Tips Dan Trik Membangun Relationship Dan Query Dalam," vol. 1, no. 2, 2018.
- [4] A. I. Tarigan, "Model Optimasi Pemetaan Mata Kuliah Berprasyarat Untuk Rencana Studi Mahasiswa ( Studi Kasus Program Studi Matematika Fmipa Ut )."
- [5] P. Studi, T. Informatika, U. Harapan, and M. Sumatera, "Decision Tree Penentuan Masa Studi Mahasiswa Prodi Teknik Informatika ( Studi Kasus : Fakultas Teknik dan Komputer Universitas Harapan Medan )," vol. 5341, no. April, pp. 16–24, 2018.
- [6] M. S. Sinambela, E. Rosely, F. I. Terapan, and U. Telkom, "THE DECISION SUPPORT SYSTEM FOR SUBJECT SPECIALIZATION STUDENTS IN," vol. 2, no. 3, pp. 858–866, 2016.
- [7] N. Hijriana, "Penerapan metode decision tree algoritma c4.5 untuk klasifikasi mahasiswa berprestasi," pp. 39–43, 2008.
- [8] Rizaldi, R., Anggraeni, D., & Syah, A. Z. (2018). TIPS DAN TRIK MEMBANGUN RELATIONSHIP DAN QUERY DALAM DATABASE. Jurdimas (Jurnal Pengabdian Kepada Masyarakat) Royal, 1(2), 45-50.