

## **BAB II. LANDASANTEORI**

### 2.1 Uang Kuliah Tunggal

Uang Kuliah Tunggal atau disingkat UKT menurut Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 39 Tahun 2017, [2]. *“Uang Kuliah Tunggal yang selanjutnya disingkat UKT adalah biaya yang ditanggung setiap mahasiswa berdasarkan kemampuan ekonominya”* [2]. Uang kuliah tunggal sebagaimana yang dimaksud terdiri atas beberapa kelompok yang ditentukan berdasarkan kelompok kemampuan ekonomi masyarakat. Uang kuliah tunggal ini ditetapkan berdasarkan biaya kuliah tunggal dikurangi biaya yang ditanggung oleh pemerintah.

*“Biaya Kuliah Tunggal yang selanjutnya disingkat BKT adalah keseluruhan biaya operasional yang terkait langsung dengan proses pembelajaran mahasiswa per semester pada program studi di PTN”*[2].

Dengan diberlakukannya UKT ini, perguruan tinggi negeri tidak dibolehkan memungut uang pangkal atau pungutan lainnya dari mahasiswa baru yang telah diterima diperguruan tinggi tersebut.

### 2.2 Data Mining

Menurut Jiawei Han, Micheline Kamber, Jian Pei dengan buku yang berjudul *Data Mining Concepts and Technique*, Data Mining (Penambangan data) adalah bidang yang dinamis dan berkembang cepat dengan kekuatan besar. Pada bagian ini, kami secara singkat menguraikan masalah utama dalam penelitian

penambangan data, mempartisi mereka menjadi lima kelompok: metodologi penambangan, interaksi pengguna, efisiensi dan skalabilitas, keragaman tipe data, dan data mining dan masyarakat. Banyak dari masalah ini telah dibahas dalam penelitian dan pengembangan data mining baru-baru ini sampai batas tertentu dan sekarang dianggap sebagai persyaratan data mining; yang lain masih dalam tahap penelitian. Isu-isu tersebut terus merangsang penyelidikan dan peningkatan lebih lanjut dalam penambangan data [3].

### 2.3 Metodologi Penambangan

Para peneliti telah giat mengembangkan metodologi penambangan data baru. Ini melibatkan penyelidikan jenis-jenis baru pengetahuan, penambangan dalam ruang multidimensi, mengintegrasikan metode dari disiplin ilmu lain, dan pertimbangan dari ancaman di antara objek data. Selain itu, metodologi penambangan harus mempertimbangkan masalah-masalah seperti ketidakpastian data, kebisingan, dan ketidaklengkapan. Beberapa metode penambangan mengeksplorasi bagaimana langkah-langkah khusus pengguna dapat digunakan untuk menilai ketertarikan pola yang ditemukan serta memandu proses penemuan. Mari kita lihat berbagai aspek metodologi penambangan ini.

Menambang berbagai macam dan jenis pengetahuan baru: Penambangan data mencakup spektrum yang luas dari analisis data dan tugas-tugas yang diketahui, mulai dari analisis karakterisasi dan diskriminasi hingga asosiasi dan analisis korelasi, klasifikasi, regresi, pengelompokan, analisis pencilan, analisis urutan, dan tren serta analisis evolusi. Tugas-tugas ini dapat menggunakan database yang

sama dengan cara yang berbeda dan membutuhkan pengembangan berbagai teknik penambangan data. Karena beragamnya aplikasi, tugas-tugas penambangan baru terus berlanjut, membuat bidang data menjadi dinamis dan cepat tumbuh. Contohnya, untuk penemuan pengetahuan yang efektif dalam jaringan informasi, pengelompokan dan pemeringkatan yang terintegrasi dapat mengarah pada penemuan kelompok dan peringkat objek yang berkualitas tinggi dalam jaringan besar. Menambang pengetahuan dalam ruang multidimensi: Saat mencari pengetahuan dalam kumpulan data besar, kita dapat menjelajahi data dalam ruang multidimensi. Artinya, kita dapat mencari pola yang menarik di antara kombinasi dimensi (atribut) pada berbagai tingkat abstraksi. Penambangan semacam itu dikenal sebagai penambangan data multidimensi (eksplorasi). Dalam banyak kasus, data dapat dikumpulkan atau dilihat sebagai kubus data multidimensi. Menambang pengetahuan dalam ruang kubus secara substansial dapat meningkatkan kekuatan dan fleksibilitas penambangan data. Penambangan data — upaya lintas disiplin: Kekuatan penambangan data dapat ditingkatkan secara substansial dengan mengintegrasikan metode baru dari berbagai disiplin ilmu. Misalnya, untuk menambang data dengan teks bahasa alami, masuk akal untuk menggabungkan metode penambangan data dengan metode pencarian informasi dan pemrosesan bahasa alami. Sebagai contoh lain, pertimbangkan menambang bug perangkat lunak dalam program besar. Bentuk penambangan ini, yang dikenal sebagai penambangan bug, mendapat manfaat dari penggabungan pengetahuan rekayasa perangkat lunak ke dalam proses penambangan data. Meningkatkan kekuatan penemuan di lingkungan jaringan: Sebagian besar objek data berada di

lingkungan yang terhubung atau saling berhubungan, apakah itu Web, hubungan basis data, file, atau dokumen. Tautan semantik di beberapa objek data dapat digunakan untuk mendapatkan keuntungan dalam penambangan data. Pengetahuan yang diperoleh dalam satu set objek dapat digunakan untuk mendorong penemuan pengetahuan dalam set objek yang “terkait” atau terhubung secara semantik. Menangani ketidakpastian, kebisingan, atau ketidaklengkapan data: Data sering mengandung kebisingan, kesalahan, pengecualian, atau ketidakpastian, atau tidak lengkap. Kesalahan dan kebisingan dapat membingungkan proses penambangan data, yang mengarah ke derivasi pola yang salah. Pembersihan data, pemrosesan ulang data, pendeteksian awal dan penghapusan, dan penalaran ketidakpastian adalah contoh teknik yang perlu diintegrasikan dengan proses penambangan data. Evaluasi pola dan penambangan dengan pola atau kendala: Tidak semua pola yang dihasilkan oleh proses penambangan data menarik. Apa yang membuat suatu pola menarik dapat bervariasi dari pengguna ke pengguna. Oleh karena itu, teknik diperlukan untuk menilai ketertarikan pola yang ditemukan berdasarkan pada tindakan subjektif. Ini memperkirakan nilai pola sehubungan dengan kelas pengguna yang diberikan, berdasarkan keyakinan atau harapan pengguna. Selain itu, dengan menggunakan langkah-langkah menarik atau kendala yang ditentukan pengguna untuk memandu proses penemuan, kami dapat menghasilkan pola yang lebih menarik dan mengurangi ruang pencarian.

*Teori yang lain menyebutkan bahwa “Data Mining berisi pencarian trend atau pola yang diinginkan dalam database besar untuk membantu pengambilan*

keputusan di waktu yang akan *dating*. Pola-pola ini dikenali oleh perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan berwawasan yang kemudian dapat dipelajari dengan lebih teliti, yang mungkin saja menggunakan perangkat pendukung keputusan” (Hermawati, 2013).

Sedangkan menurut Siallagan[4]

- a. Mining adalah mengekstrak atau menggali pengetahuan dari kumpulan data yang sangat besar.”
- b. Data yang diekstraksi menjadi informasi yang berguna, dimana tidak diharapkan, tidak dikenal dan implisit.
- c. Eksplorasi & analisis, dari sekumpulan data yang sangat besar untuk memperoleh pola-pola data yang berart secara otomatis atau semi otomatis.
- d. Menemukan pola yang valid, baru, berguna dan dapat dipahami manusia melalui proses analisis *database* yang besar secara semi otomatis.

Menurut A.Prasetyo[5] pengertian data mining secara naratif mempunyai beberapa maksud:

- a. Pencarian otomatis pola dalam basis data yang besar, menggunakan teknik komputasional campuran dari *machine learning*, statistik, dan pengenalan pola.
- b. Pengekstrakan implisit non-trivial, yang sebelumnya belum diketahui secara potensial adalah informasi berguna dari data.
- c. Ilmu pengekstrakan informasi yang berguna dari set data atau basis data besar.
- d. Eksplorasi otomatis atau semi otomatis dan analisis data dalam jumlah besar, dengan tujuan untuk menemukan pola yang bermakna.

- e. Proses penemuan informasi otomatis dengan mengidentifikasi pola dan hubungan “tersembunyi” dalam data.

*Data Mining* merupakan teknologi yang sangat berguna untuk membantu perusahaan - perusahaan menemukan informasi yang sangat penting dari gudang data (*Data warehouse*) mereka. Dengan *data mining* dapat meramalkan tren dan sifat-sifat perilaku bisnis yang sangat berguna untuk mendukung pengambilan keputusan penting (Kusnawi, 2007)[6].

*Data mining* melibatkan integrasi teknik dari berbagai disiplin ilmu seperti *database* dan *data warehouse* teknologi, statistik, pembelajaran mesin (*machine learning*), kinerja tinggi komputasi, pengenalan pola, jaringan saraf (*neural network*), visualisasi data, informasi pengambilan, gambar dan pemrosesan sinyal, dan analisis data spasial atau temporal (Han *et al.*, 2012)[3].

*Data warehousing* dan *data mining* merupakan dasar-dasar arsitektural bagi sistem-sistem pendukung keputusan. Keduanya memiliki hubungan simbiotik dimana *data warehouse* menyiapkan tahapan untuk kegiatan *data mining* yang efektif. Teknologi *data warehouse* telah memungkinkan sebuah organisasi untuk mengelola dan menyimpan data bisnis dalam volume yang sangat besar dalam bentuk yang dapat dianalisa. Kematangan dalam bidang kecerdasan buatan telah pula menciptakan sekumpulan teknik mesin pembelajaran (*machine learning*) yang berguna untuk mengotomatisasi kegiatan-kegiatan penting dan melelahkan guna mengungkapkan pola-pola dalam database. (Sitompul, 2008)[7]

### 2.3.1. Operasi *data mining*

Operasi *data mining* menurut sifatnya dibedakan menjadi dua yaitu:

#### a. Prediksi (*predictiondriven*)

Yaitu untuk menjawab pertanyaan apa dan sesuatu yang bersifat remang-remang atau transparan. Operasi prediksi digunakan untuk validasi hipotesis, *querying* dan pelaporan, analisis multidimensi; *OLAP (Online Analytic Processing)* serta analisis statistik.

#### b. Penemuan (*discoverydriven*)

Yaitu bersifat transparan dan untuk menjawab pertanyaan “mengapa?”. Operasi penemuan digunakan untuk analisis data eksplorasi, pemodelan prediktif, segmentasi *database*, analisis keterkaitan (*link analysis*) dan deteksi deviasi.

### 2.3.2. Tahapan dalam datamining

Dalam penggunaan *data mining* terdapat proses dalam penggunaannya. Adapun tahapannya dapat dilihat pada Gambar 2.1.



**Gambar 2.1.** Tahap-tahap penambangan data. (Han, *et al.*, 2006)

Pada tahapan yang terlihat pada gambar 2.1 dapat diuraikan sebagai berikut:

- a) Memahami *domain* aplikasi untuk mengetahui dan menggali pengetahuan awal serta apa sasaran pengguna
- b) Membuat target data-set yang meliputi pemulihan data dan fokus pada sub-set data.
- c) Pembersihan dan transformasi data meliputi eliminasi derau, *outliners*, *missing value*, serta pemilihan fitur dan reduksi dimensi.
- d) Penggunaan algoritma *data mining* yang terdiri dari asosiasi, sekuensial, klasifikasi, klusterisasi, dll.
- e) Interpretasi, evaluasi dan visualisasi pola untuk melihat apakah ada sesuatu yang baru dan menarik dan dilakukan iterasi jika diperlukan.

### 2.3.3. Teknik *datamining*

Beberapa teknik dan sifat dalam *data mining* adalah sebagai berikut :

- a) Klasifikasi (*Classification*)

Menentukan sebuah *record* data baru ke salah satu dari beberapa kategori atau kelas yang telah didefinisikan sebelumnya.

- b) Klusterisasi (*Clustering*)

Mempartisi data-set menjadi beberapa sub-set atau kelompok sedemikian rupa sehingga elemen-elemen dari suatu kelompok tertentu memiliki set property yang dishare bersama, dengan tingkat similaritas tinggi dalam satu kelompok dan tingkat similaritas antar kelompok yang rendah.

- c) Kaidah Asosiasi (*Association Rules*)

Mendeteksi kumpulan atribut-atribut yang muncul bersamaan dalam frekuensi yang sering, dan membentuk sejumlah kaidah dari kumpulan-kumpulan tersebut.

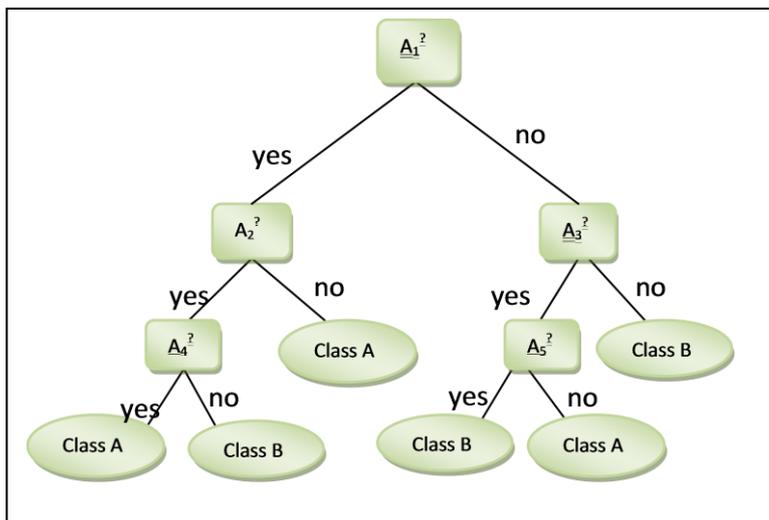
d) Pencarian pola sekuensial (*Sequence Mining*)

Mencari sejumlah *event* yang secara umum terjadi bersama-sama.

e) Regresi (*Regression*)

Memprediksi nilai dari suatu variabel kontinu yang diberikan berdasarkan nilai dari variabel yang lain, dengan mengasumsikan sebuah model ketergantungan linier atau nonlinier.

#### 2.4 Metode Decision Tree



**Gambar 2.2.** Pohon Keputusan

*Decision Tree* atau pohon keputusan adalah pohon yang digunakan sebagai prosedur penalaran untuk mendapatkan jawaban dari masalah yang dimasukkan.

Pohon yang dibentuk tidak selalu berupa pohon biner. Jika semua fitur dalam data set menggunakan dua macam nilai kategorikal maka bentuk pohon yang didapatkan berupa pohon biner. Jika dalam fitur berisi 2 macam nilai kategorikal atau menggunakan tipe numerik maka bentuk pohon yang didapatkan biasanya tidak berupa pohon biner[4].Metode ini merupakan salah satu metode yang ada

pada teknik klasifikasi dalam *data mining*. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Pohon keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target[1]. *Decision tree* adalah metode diskriminasi nonlinier yang menggunakan sekumpulan variabel independen untuk membagi sampel ke dalam kelompok-kelompok yang lebih kecil secara bertahap. Prosedur tersebut dilakukan secara iteratif di setiap cabang pohon, yakni memilih variabel independen yang memiliki asosiasi terkuat dengan variabel dependen menurut kriteria tertentu.

Menurut Prasetyo[7], *Decision Tree* mempunyai tiga pendekatan klasik:

- a) Pohon klasifikasi, digunakan untuk melakukan prediksi ketika ada data baru yang belum diketahui label kelasnya. Pendekatan ini yang paling banyak digunakan.
- b) Pohon regresi, ketika hasil prediksi dianggap sebagai nilai nyata yang mungkin akan didapatkan. Misalnya kasus harga minyak, kenaikan harga rumah, prediksi inflasi tiap tahun dan sebagainya.
- c) CART (atau C&RT), ketika masalah klasifikasi dan regresi digunakan bersama-sama.

Ada banyak pilihan algoritma untuk menginduksi *decision tree*, seperti Hunt, CART (C&RT), ID3, C4.5, SLIQ, SPRINT, QUEST, DTREG, THAID, CHAID, dan sebagainya.

### 2.4.1. Algoritma C4.5

Algoritma C4.5 diperkenalkan oleh Quinlan (1996) sebagai versi perbaikan dari ID3. C4.5 adalah algoritma yang sudah banyak dikenal dan digunakan untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. Hasil dari proses klasifikasi yang berupa aturan-aturan dapat digunakan untuk memprediksi nilai atribut bertipe diskret dari *record* yang baru.

Algoritma C4.5 merupakan suksesor dari ID3 menggunakan rasio gain untuk memperbaiki *information gain*. Rasio gain merupakan normalisasi dari *information gain* dengan memperhitungkan nilai *entropy* dari probabilitas *subset* setelah dilakukan proses partisi. Atribut dengan nilai rasio gain tertinggi dipilih sebagai atribut node akar [6].

Secara sistematis Gain dihitung dengan Persamaan 1 dibawah ini:

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n * \text{Entropy}(S_i) \dots\dots (1)$$

Keterangan:

- a) S : himpunan Kasus
- b) A : atribut
- c) n : jumlah partisi atribut A
- d) |S<sub>i</sub>| : jumlah kasus pada partisi ke-i
- e) |S| : jumlah kasus dalam S

### Menghitung Nilai Entropy

Untuk menghitung jumlah entropy dengan cara menghitung hasil penjumlahan dari keseluruhan partisi pada suatu kasus dari min p ke i dikali Log2 dari p ke I dimana p ke I adalah proporsi dari s ke I terhadap S, sehingga bisa dibuat persamaan sebagai berikut:

$$\text{Entropy}(S) = \sum_{i=1}^n -p_i * \log_2 p_i \dots\dots\dots (2)$$

Keterangan:

S : himpunan kasus

A : fitur

n : jumlah partisi S

$p_i$  : proporsi dari  $S_i$  terhadap S

Perincian algoritma ( langkah 1)

- a) Menghitung jumlah kasus seluruhnya
- b) Menghitung Entropy dari semua kasus
- c) Lakukan penghitungan Gain utk setiap atributnya

Langkah selanjutnya adalah tinggal melakukan hal yang sama seperti pada rumus diatas dengan tidak mengikut sertakan atribut yang gainnya sudah terpilih sampai didapat hasil berupa decision tree.

#### 2.4.2. *Confusion Matrix*

*Confusion matrix* adalah suatu metode yang digunakan untuk melakukan perhitungan akurasi pada konsep *data mining*. Evaluasi menggunakan *confusion matrix* menghasilkan nilai *accuracy*, *recall* dan *precision*. Akurasi dalam klasifikasi adalah presentase ketepatan record data yang diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi. *Recall* merupakan jangkauan algoritma terhadap nilai yang sebenarnya, sedangkan *precision* merupakan ketepatan algoritma terhadap nilai yang telah ditemukan algoritma[12].

Nilai Prediksi	Nilai Sebenarnya	
	True	False
True	$TP$ <i>(True Positive)</i> <i>Correct Result</i>	$FP$ <i>(False Positive)</i> <i>Unexpected Result</i>
False	$FN$ <i>(False Negative)</i> <i>Missing Result</i>	$TN$ <i>(True Negative)</i> <i>Correct Absence of Result</i>

**Tabel 1: Model Confusion Matrik**

Keterangan untuk tabel 1 dinyatakan sebagai berikut:

1. *True Positive* (TP), yaitu jumlah dokumen dari kelas *true* yang benar dan diklasifikasikan sebagai kelas *true*.
2. *True Negative* (TN), yaitu jumlah dokumen dari kelas *false* yang benar diklasifikasikan sebagai kelas *false*.
3. *False Positive* (FP), yaitu jumlah dokumen dari kelas *false* yang salah diklasifikasikan sebagai kelas *true*.
4. *False Negative* (FN) yaitu jumlah dokumen dari kelas *true* yang salah diklasifikasikan sebagai kelas *false*.

Perhitungan akurasi dinyatakan dalam persamaan 3

$$\text{Accuracy} = \frac{Tp+Tn}{Tp+Fn+Fp+Tn} \times 100\% \dots\dots\dots (3)$$

Perhitungan Recall dinyatakan dalam persamaan 4

$$\text{Recall} = \frac{Tp}{Rp+Fn} \dots\dots\dots (4)$$

## 2.5 Rapid Miner

*RapidMiner* merupakan perangkat lunak yang bersifat *open source*. *RapidMiner* adalah sebuah *tools* untuk melakukan analisis terhadap *data mining*, *text mining* dan analisis prediksi. *RapidMiner* dibangun dengan menggunakan Bahasa *java*, sehingga dapat dijalankan di semua sistem operasi. *RapidMiner* sebelumnya bernama YALE (*Yet Another Learning Environment*) yang dikembangkan pada tahun 2001 oleh RalfKlinkenberg, Ingo Mierswa, dan Simon Fischer di *Artificial Intelligence Unit* dari *University of Dortmund*.

Beberapa Fitur dari *RapidMiner*, antara lain:

- a) Banyaknya algoritma *data mining*, seperti *decision tree* dan *self-organization map*.
- b) Bentuk grafis yang canggih, seperti tumpang tindih diagram histogram, *tree chart* dan *3D Scatter plots*.
- c) Banyaknya variasi *plugin*, seperti *text plugin* untuk melakukan analisis teks.
- d) Menyediakan prosedur *data mining* dan *machine learning* termasuk: ETL (*extraction, transformation, loading*), *data preprocessing*, *visualisasi*, *modelling* dan evaluasi.
- e) Proses data mining tersusun atas operator-operator yang *nestable*, dideskripsikan dengan XML, dan dibuat dengan GUI
- f) Mengintegrasikan proyek *data mining Weka* dan statistika R

## 2.6 Sistem Pendukung Keputusan

Sistem pendukung keputusan adalah sistem interaktif yang membantu pengambilan keputusan melalui penggunaan data dan model-model keputusan

untuk memecahkan masalah-masalah yang sifatnya semi terstruktur dan tidak terstruktur. Tahap-tahap dalam pengambilan keputusan terdiri dari:

a) Tahappemahaman

Merupakan proses penelusuran dan pendeteksian dari lingkup problematika serta proses pengenalan masalah. Data masukan diperoleh, diproses dan diuji dalam rangka mengidentifikasi masalah.

b) Tahapperancangan

Proses pengembangan dan pencarian alternatif tindakan atau penyelesaian yang dapat diambil. Hal tersebut merupakan representasi kejadian nyata untuk mengetahui keakuratan model dalam meneliti masalah yang ada.

c) Tahappemilihan

Pemilihan terhadap berbagai alternatif penyelesaian yang dimunculkan pada tahap perencanaan agar ditentukan atau diperhatikan kriteria-kriteria berdasarkan tujuan yang akan dicapai.

d) Tahapimplementasi

Penerapan terhadap rancangan sistem yang telah dibuat pada tahap perancangan serta pelaksanaan alternatif tindakan yang telah dipilih pada tahap pemilihan.

## 2.7 Penelitian Terdahulu

Dalam melakukan penelitian, penulis membutuhkan beberapa bahan penelitian yang sudah pernah dilakukan peneliti-peneliti lainnya Terkait dengan Uang Kuliah Tunggal mengenai masalah teknik data mining

a) Budianto Karim<sup>1)</sup>, Steven R. Sentinuwo<sup>2)</sup>, Alwin M. Sambul<sup>3)</sup> dengan

judul:[4]

”Penentuan Besaran Uang Kuliah Tunggal untuk Mahasiswa Baru di Universitas Sam Ratulangi Menggunakan Data Mining 2017”.

Dari hasil pengujian dihasilkan bahwa Dari hasil pengujian dapat diketahui tingkat akurasi dari algoritma C4.5 menggunakan confusion matrix dengan jumlah data 80 menunjukkan bahwa tingkat akurasinya 50%, yang terdiri dari 50% untuk data training dan 50% untuk data testing. Sedangkan untuk hasil pengujian dengan jumlah data 115 menunjukkan bahwa tingkat akurasinya sebesar 52.63%. Dimana jumlah data sangat berpengaruh dalam proses pembentukan model dan proses pengujian model.

b) Muhammad Arif Rahman (2015) dengan judul:

*“Algoritma C45 Untuk Menentukan Mahasiswa Penerima Beasiswa (Studi Kasus : Pps Iain Raden Intan Bandar Lampung)[8]”.*

Didapat kesimpulan bahwa dari sampel yang diambil yaitu sebanyak 40 (empat puluh mahasiswa) calon penerima beasiswa pada Program Pascasarjana IAIN Raden Intan Lampung dapat dilihat pada pohon keputusan yaitu sebanyak 18 (delapan belas) mahasiswa yang tidak layak menjadi penerima beasiswa karena mempunyai IPK <3,00, kemudian sebanyak 8 (delapan) mahasiswa yang tidak layak menjadi penerima beasiswa karena mempunyai masa kerja <5 tahun dan pekerja Non PNS, sehingga dihasilkan sebanyak 14 (empat belas) mahasiswa yang layak menjadi penerima beasiswa karena telah memenuhi kriteria penerima beasiswa dari segi IPK, pekerjaan dan masa kerja yang telah ditentukan sebelumnya.

c) Eka Pandu Cynthia<sup>1)</sup>, Edi Ismanto<sup>2)</sup> Tahun 2018 dengan Judul:

*“Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji[9]”*

Didapat Kesimpulan bahwa: Dari hasil percobaan pencarian pohon hasil keputusan dari data penjualan gerai makanan cepat saji menggunakan algoritma C4.5 dihasilkan nilai entropy dan gain tertinggi yaitu 1,501991 pada atribut-atribut Menu Makanan pada perhitungan manual. Sedangkan menggunakan aplikasi Rapidminer diperoleh hasil pohon keputusan seperti terlihat pada Gambar 3.2. Harga – Jumlah Terjual – Menu Makanan (Rice Bento = Kurang Laris, Dada = Laris) dengan bobot (weight) masing-masing atribut : Harga (0,738), Jenis Menu (0,067), Jumlah Terjual (0,156), Status Penjualan (0,040).