

## **BAB II**

### **LANDASAN TEORI**

#### **2.1 Sistem Informasi Terpadu**

Sistem Informasi Terpadu merupakan suatu sistem berbasis teknologi informasi yang dirancang untuk mengintegrasikan berbagai proses, data, dan layanan dalam satu kesatuan yang saling terhubung dan terkoordinasi. Sistem ini memungkinkan pengelolaan informasi secara terpusat sehingga data yang dihasilkan dari satu bagian dapat digunakan secara langsung oleh bagian lain tanpa perlu dilakukan pengolahan ulang. Dengan adanya integrasi tersebut, Sistem Informasi Terpadu bertujuan untuk meningkatkan efisiensi, efektivitas, serta akurasi dalam pengelolaan informasi, khususnya pada organisasi atau institusi yang memiliki kompleksitas layanan yang tinggi, seperti perguruan tinggi.

Dalam konteks perguruan tinggi, Sistem Informasi Terpadu berfungsi sebagai sarana pendukung utama dalam pelaksanaan kegiatan akademik dan administrasi. Sistem ini mencakup berbagai modul layanan, antara lain pengelolaan data mahasiswa, pengisian Kartu Rencana Studi (KRS), manajemen perkuliahan, pengolahan nilai, hingga layanan administrasi akademik lainnya. Seluruh modul tersebut saling terhubung dalam satu platform sehingga memudahkan pengguna, baik mahasiswa, dosen, maupun tenaga kependidikan, dalam mengakses informasi dan layanan secara cepat dan akurat.

Penerapan Sistem Informasi Terpadu juga berperan penting dalam mendukung transparansi, konsistensi data, serta pengambilan keputusan yang

lebih tepat. Dengan sistem yang terintegrasi, risiko terjadinya kesalahan data, keterlambatan informasi, dan duplikasi proses dapat diminimalkan. Oleh karena itu, Sistem Informasi Terpadu tidak hanya berfungsi sebagai alat bantu operasional, tetapi juga sebagai komponen strategis dalam meningkatkan kualitas layanan dan tata kelola institusi pendidikan secara keseluruhan serta Sistem informasi memiliki nilai tambah jika sistem tersebut menghasilkan informasi berdasarkan data nyata yang diolah dan bermanfaat bagi penerimanya(Kinanti et al., 2021).

## **2.2 Kepuasan Pengguna**

kepuasan pengguna merupakan salah satu faktor yang mempengaruhi keberhasilan suatu sistem, semakin berkualitas produk atau jasa yang diberikan, maka kepuasan pelanggan akan semakin tinggi(M. Ali Akbar, Fajriyah, 2024).

Kepuasan pengguna menggambarkan keselarasan antara harapan seseorang dan hasil yang diperoleh dengan adanya suatu sistem dimana tempat orang tersebut berpartisipasi dalam pengembangan sistem informasi. Kepuasan pengguna sistem informasi merupakan tolak ukur tingkat keberhasilan sebuah sistem informasi, kepuasan tersebut merupakan sebuah penilaian yang menyangkut apakah kinerja suatu sistem informasi terbilang baik atau buruk, dan apakah sistem informasi yang digunakan cocok atau tidak dengan tujuan penggunaannya(Abidin & Nuryana, 2023).

Dalam konteks analisis kepuasan pengguna aplikasi Sistem Informasi Terpadu Universitas Labuhanbatu, penerapan algoritma C4.5 memiliki peran strategis karena mampu mengolah data kuesioner pengguna secara sistematis

untuk menghasilkan pola hubungan antar atribut penilaian yang tidak terlihat secara langsung. Dengan membangun pohon keputusan, algoritma C4.5 dapat mengidentifikasi faktor mana yang paling berpengaruh terhadap tingkat kepuasan pengguna, misalnya kemudahan penggunaan, kecepatan akses, keandalan sistem, tampilan antarmuka, serta ketepatan informasi yang disajikan. Setiap atribut dievaluasi melalui perhitungan entropy dan information gain sehingga diperoleh bobot pengaruh yang objektif terhadap kategori kepuasan. Melalui pendekatan ini, analisis kepuasan tidak hanya bersifat deskriptif, tetapi menghasilkan model klasifikasi yang mampu menjelaskan keterkaitan logis antarvariabel. Dengan demikian, algoritma C4.5 tidak hanya membantu dalam menentukan tingkat kepuasan pengguna, tetapi juga memberikan wawasan mendalam mengenai aspek-aspek sistem yang perlu diperbaiki atau ditingkatkan berdasarkan perilaku data pengguna secara nyata.

### **2.3 Indikator Kepuasan Pengguna**

Indikator kepuasan pengguna pada penelitian ini suatu faktor penting dalam mengevaluasi kinerja dan efektivitas suatu sistem informasi, termasuk Sistem Informasi Terpadu Universitas Labuhanbatu (SITU). Dalam konteks ini, indikator kepuasan pengguna mencakup beberapa dimensi utama yang mempengaruhi persepsi pengguna terhadap sistem, seperti kemudahan penggunaan, kecepatan akses, keandalan sistem, dan tampilan antarmuka. Kemudahan penggunaan mencerminkan sejauh mana pengguna merasa nyaman dan mudah berinteraksi dengan sistem tanpa menghadapi kesulitan teknis. Kecepatan akses menggambarkan seberapa cepat pengguna dapat mengakses dan

memanfaatkan berbagai layanan yang tersedia dalam sistem. Keandalan sistem, di sisi lain, berkaitan dengan sejauh mana sistem dapat berfungsi dengan stabil dan konsisten, tanpa adanya gangguan teknis yang merugikan pengguna.

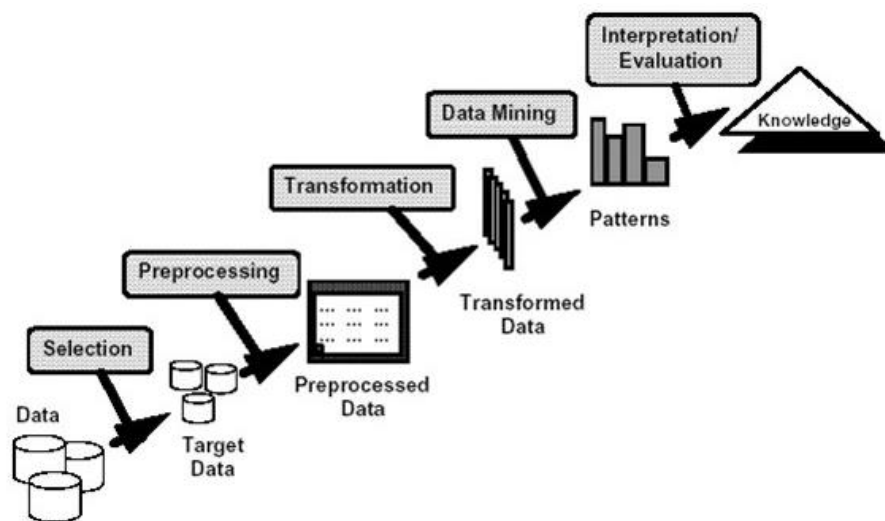
Tampilan antarmuka juga merupakan aspek penting yang memengaruhi kepuasan pengguna. Sebuah antarmuka yang menarik dan mudah dinavigasi akan meningkatkan kenyamanan pengguna, sementara antarmuka yang rumit atau tidak responsif dapat menurunkan kepuasan.

Penting untuk diingat bahwa kepuasan pengguna tidak hanya ditentukan oleh faktor-faktor teknis tersebut, tetapi juga oleh persepsi pengguna terhadap pengalaman mereka dan perasaan seseorang setelah menggunakannya lalu membandingkan kinerja atau hasil yang dirasakan dengan harapannya dalam menggunakan sistem. Oleh karena itu, evaluasi terhadap tingkat kepuasan pengguna perlu dilakukan secara menyeluruh dengan mempertimbangkan seluruh dimensi ini. Dalam penelitian ini, penerapan algoritma C4.5 dalam analisis data pengguna diharapkan dapat memberikan gambaran yang lebih jelas mengenai faktor-faktor yang memengaruhi kepuasan pengguna, sehingga sistem dapat terus dikembangkan sesuai dengan kebutuhan dan ekspektasi pengguna (Winantu et al., 2023).

#### **2.4 Knowledge Discovery in Database (KDD)**

*Knowledge discovery in databases* (KDD) adalah metode teknis yang berguna untuk mencari dan mengidentifikasi pola (pattern) dalam data, pola yang sudah ditemukan bersifat sah dan baru sehingga dapat bermanfaat dan dapat dimengerti (Dewi et al., 2022).

KDD merupakan proses mencari informasi yang lebih bernilai, lebih mudah dipahami dan baru dari penyimpanan data yang besar dan kompleks. Proses KDD menafsirkan hasil yang diperoleh dari sekumpulan data dengan menggabungkan dengan ilmu lainnya. Proses KDD dimulai dengan menetapkan tujuan dan diakhiri dengan evaluasi (Widaningsih, 2019).



**Gambar 2.1** Proses Tahapan Knowledge Discovery in Database (KDD)

Berikut penjelasan mengenai proses tahapan dalam *Knowledge Discovery in Database* (KDD) :

a. Data Selection

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *Data Mining*, kata memilih data-data seperti apa saja yang kita butuhkan untuk proses lebih lanjut dan kemudian data disimpan dalam suatu berkas, terpisah dari basis data

operasional sehingga memberikan kemudahan untuk penggunaan berikutnya.

b. Pre-processing (Cleaning)

Pada umumnya data yang diperoleh, baik dari database suatu perusahaan maupun eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut data yang tidak relevan itu juga lebih baik dibuang karena keberadaannya bisa mengurangi mutu atau akurasi dari hasil *Data Mining* nantinya.

c. Data Transformation

Beberapa teknik *Data Mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa teknik standar seperti analisis asosiasi dan klastering hanya bisa menerima input katagorikal. Karenanya data berupa angka numerik yang berlanjut bisa dibagi-bagi menjadi beberapa interval. Proses ini sering disebut binning. Disini juga dilakukan pemilihan data yang diperlukan oleh teknik *Data Mining* yang dipakai. Transformasi dan pemilihan data ini juga menemukan kualitas dari hasil *Data Mining* nantinya karena ada beberapa karakteristik dari teknik-teknik *Data Mining* tertentu tergantung pada tahap ini.

d. *Data Mining*

*Data Mining* adalah proses menjadi pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik,

metode atau algoritma dalam *Data Mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat tergantung pada tujuan dan proses KDD secara keseluruhan. Data yang digunakan untuk bisa menjadi sebuah model yang baik idealnya mencukupi sebagai data riset. Semakin banyak data dan semakin sedikit kesalahan (error) maka semakin bagus model yang dijadikan patokan.

e. Interpretation (Evaluation)

Pola informasi yang dihasilkan dari proses *Data Mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut interpretasi. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta pola atau informasi yang ada sebelumnya (Shaumi et al., 2022).

## 2.5 *Data Mining*

*Data mining* merupakan proses perulangan dan memerlukan interaksi manusia dalam prosesnya untuk menemukan pola atau model baru yang dapat digeneralisasi untuk masa yang akan datang, dan bermanfaat jika digunakan untuk melakukan suatu tindakan (Ririn Ardilla, Sundari Retno Andani, 2024).

*Data Mining* merupakan suatu cara mencari hubungan yang signifikan antara variabel dalam set data yang besar dengan memanfaatkan teknologi yang sesuai dan dalam penelitiannya memakai studi literatur yang berbeda. Beberapa teknik yang ada dalam *data mining* salah satunya yaitu classification (Nenitrolina Ndruru, 2024).

Melalui teknik klasifikasi, penelitian ini mampu membangun model pengelompokan data berdasarkan atribut-atribut yang relevan. Model tersebut digunakan untuk mengidentifikasi pola dan struktur data yang terbentuk dari hubungan antarvariabel. Hasil pengelompokan ini memberikan gambaran objektif dan terukur mengenai faktor-faktor yang saling berkaitan dalam menentukan tingkat kepuasan pengguna terhadap Sistem Informasi Terpadu Universitas Labuhanbatu.

## **2.6 Algoritma C4.5**

Algoritma C4.5 adalah salah satu algoritma untuk mengubah fakta yang besar menjadi pohon keputusan (decision tree) yang merepresentasikan aturan (rule)(Adriansa et al., 2022).

Pada akhir tahun 1970 sampai awal tahun 1980 J, Ross Quinlan, Metode pohon keputusan yang disebut ID3 dikembangkan oleh seorang peneliti di bidang pembelajaran mesin (Iterative Dichotomiser). Quinlan kemudian memperbaiki ID3 dengan mengembangkan metode C4.5, terkadang dikenal sebagai pohon keputusan.

Algoritma ID3 telah berkembang menjadi algoritma C4.5. Algoritme C4.5 sekarang memiliki prinsip operasi fundamental yang sama dengan algoritme ID3 sebagai hasil dari kemajuan ini. Perlu diingat bahwa algoritma C4.5 menemukan nilai entropi adalah langkah pertama dalam prosedur ini. Atribut tersebut kemudian dipilih menggunakan nilai gain ratio(Sahadan et al., 2023).

Tahapan dari algoritma C4.5 adalah sebagai berikut:

- a. Menghitung nilai Entropy

Entropy digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan keluaran atribut.

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

Keterangan :

S : Himpunan Kasus

A : Atribut

n : Jumlah Partisi S

Pi : Proporsi dari Si terhadap S(Wulandari et al., 2024).

b. Menghitung nilai Gain untuk masing-masing atribut

Information Gain adalah ukuran untuk menentukan atribut terbaik dalam membagi data dengan cara mengurangi tingkat ketidakpastian. Atribut dengan nilai gain tertinggi dipilih sebagai dasar pembentukan pohon keputusan.

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

Keterangan:

S : Himpunan kasus

A : Attribute

n : Jumlah partisi atribut A

|Si| : Jumlah kasus pada partisi ke i (Si)

|S| : Jumlah kasus dalam S

Melalui perhitungan nilai entropy dan gain, algoritma ini mampu menentukan atribut mana yang memiliki pengaruh terbesar dalam membedakan kategori kepuasan pengguna. Proses pemilihan atribut dengan gain tertinggi memungkinkan pohon keputusan yang dihasilkan dapat menampilkan urutan logis dari faktor-faktor yang paling menentukan, seperti kemudahan penggunaan, kecepatan akses, keandalan sistem, dan tampilan antarmuka aplikasi.

- c. Atribut yang memiliki Gain tertinggi dipilih menjadi akar (*root*) dan atribut yang memiliki nilai Gain yang terendah dari akar (*root*) dipilih menjadi cabang (*branches*)
- d. Menghitung lagi nilai Gain tiap-tiap atribut dengan tidak mengikutsertakan atribut yang terpilih menjadi akar (*root*) ditahap sebelumnya.
- e. Atribut yang memiliki Gain tertinggi dipilih menjadi cabang (*branches*)
- f. Mengulangi langkah ke-4 dan ke-5 sampai dengan dihasilkan nilai Gain = 0 untuk semua atribut yang tersisa (Prasetyaningrum & Susanti, 2023).

## 2.7 Alat Bantu Program/Tools Pendukung

Dalam penelitian yang melibatkan pengolahan data, penggunaan alat bantu program (tools) merupakan komponen penting untuk memastikan proses kerja berjalan sistematis, terukur, dan dapat dipertanggungjawabkan. Pada penelitian ini, tools pendukung diperlukan untuk menunjang rangkaian kegiatan pengolahan dan analisis data dalam penerapan algoritma C4.5 untuk menentukan tingkat kepuasan pengguna Aplikasi Sistem Informasi Terpadu Universitas Labuhanbatu.

Adapun tools utama yang digunakan dalam penelitian ini adalah Microsoft Excel dan RapidMiner(Lestari et al., 2023).

### 2.7.1 Microsoft Office Excel

Microsoft Office Excel adalah sebuah program aplikasi lembar kerja spreadsheet yang dibuat dan didistribusikan oleh Microsoft Corporation untuk sistem operasi Microsoft Windows dan Mac.OS. Aplikasi ini memiliki fitur kalkulasi dan pembuatan grafik yang, dengan menggunakan strategi marketingMicrosoft yang agresif, menjadikan Microsoft Excel sebagai salah satu program komputer yang populer digunakan di dalam computer mikro hingga saat ini(Hasan, 2019).



**Gambar 2.2 Microsoft Excel**

Microsoft Excel tidak hanya berfungsi sebagai aplikasi pengolah angka, tetapi juga dapat dimanfaatkan sebagai alat bantu dalam pengelolaan dan pengolahan data penelitian. Dalam konteks penelitian berbasis data, Microsoft Excel banyak dimanfaatkan sebagai alat bantu awal (preprocessing tool) untuk melakukan rekapitulasi data, pembersihan data (data cleaning), serta transformasi data sebelum dilakukan analisis lanjutan menggunakan metode atau perangkat lunak khusus. Dalam penelitian ini, Microsoft Excel digunakan untuk membantu

proses rekapitulasi data kuesioner, pengecekan konsistensi data, serta penyusunan data ke dalam format yang rapi sebelum dilakukan analisis lebih lanjut menggunakan metode *Data Mining*. Dengan demikian, Excel berperan sebagai perangkat pendukung pada tahap persiapan data agar proses pemodelan dan analisis berikutnya dapat dilakukan secara lebih sistematis.

### 2.7.2 RapidMiner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator *Data Mining*, termasuk operator untuk input, output, data preprocessing dan visualisasi. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi (Manullang et al., 2021).



**Gambar 2.3 RapidMiner**

Untuk melakukan evaluasi terhadap tingkat kepuasan pengguna, dibutuhkan metode analisis yang mampu mengolah data secara sistematis dan objektif. Salah satu pendekatan yang dapat digunakan adalah *Data Mining*, yaitu proses penggalian informasi atau pola dari kumpulan data yang besar dan kompleks. Dalam konteks penelitian ini, *Data Mining* digunakan untuk menganalisis hasil survei pengguna terhadap Sistem Informasi Terpadu Universitas Labuhanbatu.

## 2.8 Penelitian Terdahulu

Beberapa penelitian terdahulu digunakan sebagai referensi untuk memperkuat kajian serta menjadi dasar dalam penyusunan proposal penelitian ini, yaitu sebagai berikut:

**Tabel 2.1 Penelitian Terdahulu**

No	Judul	Nama Penulis	Tahun	Hasil
1	Klasifikasi Kepuasan Pengguna Layanan Aplikasi Shopee Menggunakan Metode Decision Tree C4.5	Ballya Vicky Haekal, Lin Ernawati, Nurul Chamidah	2021	Penelitian ini menggunakan algoritma C4.5 untuk menganalisis tingkat kepuasan pengguna Shopee, dengan hasil akurasi 97%, recall 96.9% untuk kelas "Ya", dan 100% untuk kelas "Tidak", serta precision 100% untuk

				kelas "Ya" dan 66.6% untuk kelas "Tidak"
2	Prediksi Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Model Decision Tree	Badroe Zaman, Lenny Margaretta Huizen, Muhammad Basyier Ardima	2023	Menggunakan algoritma C4.5, penelitian ini mengukur kepuasan mahasiswa terhadap pelayanan akademik dengan akurasi 94.23%, recall 94.12% untuk kelas "Ya", dan 100% untuk kelas "Tidak". Precision untuk kelas "Ya" adalah 100%, dan 25% untuk kelas "Tidak"
3	Klasifikasi <i>Data Mining</i> Pada Tingkat Kepuasan Pengunjung Maccahaya Waterboom dengan Algoritma C.45	Ismail, Rezky Erwin Syah, M. Afdal Tahir	2024	Penelitian ini menggunakan algoritma C4.5 untuk mengklasifikasikan tingkat kepuasan pengunjung Maccahaya Waterboom, dengan hasil akurasi 93.75%. Variabel yang

				paling berpengaruh terhadap kepuasan adalah kebersihan (X5) dan fasilitas (X2)
4	Analisis Algoritma C4.5 dalam Mengukur Tingkat Kepuasan Staf terhadap Kinerja Teknisi Komputer	Siti Aliyah, Efani Desi, Fitri Pranita Nasution, Fithry Tahel	2025	Penelitian ini menggunakan algoritma C4.5 untuk menganalisis kepuasan staf terhadap kinerja teknisi komputer di Universitas Potensi Utama, dengan hasil akurasi 82% dan atribut yang mempengaruhi kepuasan adalah Reliability, Empathy, dan Tangibles
5	Penerapan Algoritma C4.5 Untuk Analisis Tingkat Kepuasan Santriwan Santriwati Pesantren Darul Ulum Mambaul	Muhammad Ikhsan Aji, Rakhmat Kurniawan	2025	Penelitian ini menggunakan algoritma C4.5 untuk mengklasifikasikan tingkat kepuasan santri, dengan hasil akurasi 53%. Variabel yang paling berpengaruh terhadap

	Huda			kepuasan adalah keamanan, kenyamanan ruang belajar, dan kualitas interaksi sosial. Algoritma C4.5 mampu mengidentifikasi faktor utama yang memengaruhi kepuasan santri
6	Penerapan Algoritma C4.5 dalam Menentukan Tingkat Kepuasan Pengguna Aplikasi Sistem Informasi Terpadu Universitas Labuhanbatu	Penelitian ini	2025	Masih dalam proses penelitian