

BAB 3

METODOLOGI PENELITIAN

3.1 Subjek Penelitian

Subjek penelitian yang digunakan dalam penelitian ini yaitu LMS *Openlearning* pada mata kuliah dasar umum (MKDU) “Kewarganegaraan” semester ganjil tahun 2020/2021 dan 2021/2022 untuk memperoleh data riwayat perkuliahan *online*.

3.2 Lokasi Penelitian

Penelitian ini dilakukan di Universitas Muhammadiyah Kalimantan Timur (UMKT) yang berlokasi di Jl. Ir. H. Juanda No.15 Samarinda.

3.3 Teknik Pengumpulan Data

Teknik pengumpulan data melalui 2 tahapan yaitu :

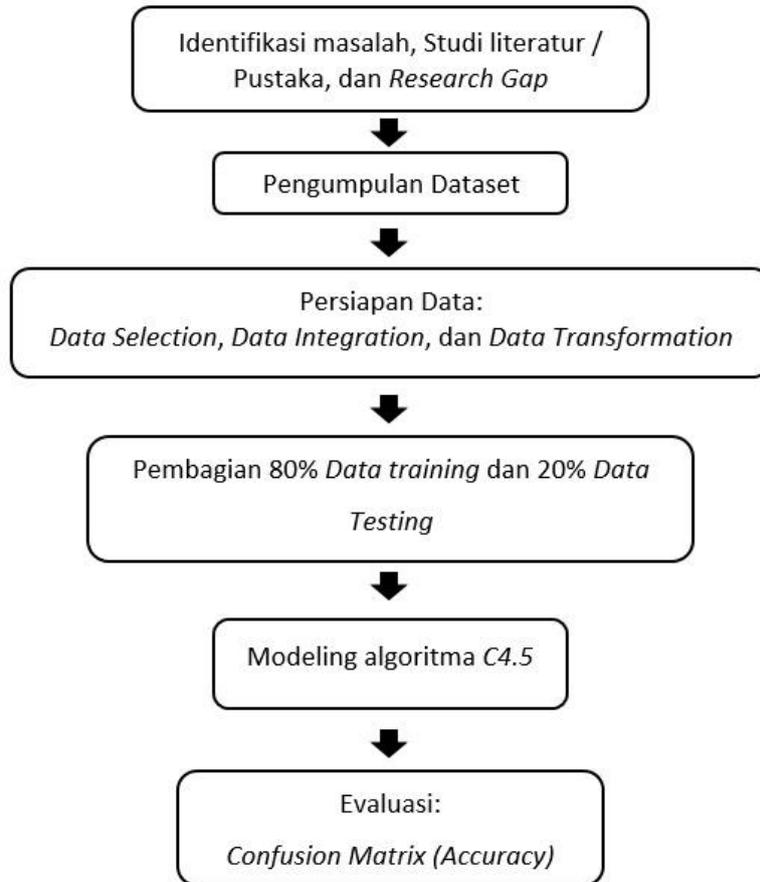
1. Observasi

Observasi dilakukan melalui pengamatan langsung tentang struktur data yang dapat diambil dari riwayat perkuliahan *online* pada mata kuliah “Kewarganegaraan” semester ganjil tahun 2020/2021 dan 2021/2022. Tahapan ini menghasilkan kumpulan *dataset* riwayat perkuliahan *online* mahasiswa UMKT yang didapatkan dari *platform OpenLearning* dan Bagian Administrasi Akademik (BAA) UMKT untuk digunakan pada proses pemodelan.

2. Studi Dokumen

Studi dokumen merupakan kegiatan pengambilan data dokumen-dokumen yang berkaitan dengan penelitian yang sedang dilaksanakan yaitu mempelajari jurnal, buku, serta referensi lainnya guna mendukung penelitian tersebut.

3.4 Alur Penelitian



Gambar 3. 1 Alur Penelitian

1. Persiapan Data

Pada penelitian ini proses prediksi terhadap performa perkuliahan *online* mahasiswa UMKT, dengan tujuan untuk dilakukan dengan metode *data mining* menggunakan algoritma C4.5 menggunakan data riwayat perkuliahan *online* mahasiswa pada *OpenLearning* dan Bagian Administrasi Akademik (BAA) Universitas Muhammadiyah Kalimantan Timur. Pada tahap persiapan data pada penelitian ini data yang digunakan diperoleh dari *platform OpenLearning* dan BAA UMKT. *Dataset* awal yang didapatkan dari *OpenLearning* adalah seperti yang terlihat pada tabel 3.1.

Tabel 3. 1 Keterangan Atribut Data *OpenLearning*

No	Atribut	Keterangan
1	<i>Profile name</i>	Id mahasiswa pada sistem <i>OpenLearning</i>
2	<i>Learner name</i>	Nama mahasiswa
3	<i>Learner email</i>	Email mahasiswa
4	<i>Enrolment ID</i>	Id pendaftaran <i>OpenLearning</i>
5	<i>Institution Membership ID</i>	Id anggota institusi
6	<i>Enrolment date</i>	Tanggal daftar
7	<i>Completion date</i>	Tanggal menyelesaikan mata kuliah
8	<i>Time spent on course</i>	Lama waktu mahasiswa berada di mata kuliah
9	<i>Progress</i>	Presentase kemajuan mahasiswa
10	<i>% Course completed</i>	Persentase kemajuan mahasiswa menyelesaikan mata kuliah
11	<i>Certificate ID</i>	Id sertifikat
12	<i>Comments</i>	Banyaknya komentar mahasiswa selama perkuliahan
13	<i>Kudos</i>	Penghargaan
14	<i>Enrolment cost</i>	Biaya pendaftaran
15	Tugas	Nilai rata-rata tugas 1-10
16	Quiz	Nilai quiz
17	UTS	Nilai ujian tengah semester

Tabel 3. 2 Keterangan Atribut Data BAA

No	Atribut	Keterangan
1	NIM	Nomor induk mahasiswa
2	Nama Mahasiswa	Nama Mahasiswa
3	Nilai Akhir	Nilai akhir mahasiswa
4	Bobot	Bobot nilai berdasarkan standar penilaian
5	Simbol	Skala Penilaian

Tahapan persiapan data yang ketiga setelah data dikumpulkan, data-data tersebut perlu diidentifikasi, dipilih, dibersihkan, dan dibangun ke dalam format yang diinginkan. Persiapan data disebut juga dengan data *pre-processing*. Data yang akan digunakan sebagai *data training* adalah data mahasiswa semua program studi Universitas Muhammadiyah Kalimantan Timur (UMKT) yang mengikuti mata kuliah Kewarganegaraan pada semester ganjil tahun 2020/2021 dan 2021/2022. Data yang digunakan merupakan data yang diperoleh dari sistem *OpenLearning* dan BAA UMKT dengan atribut *Profile name*, *Learner name*, *Learner email*, *Enrolment ID*, *Institution Membership ID*, *Enrolment date*, *Completion date*, *Time spent on course*, *Progress*, *Course completed*, *Certificate ID*, *Kudos*, *Enrolment cost* pada sistem

OpenLearning, serta Nomor Induk Mahasiswa, Program Studi, Nama Mahasiswa, Nilai Akhir, Bobot, Simbol pada bagian BAA UMKT.

a. Seleksi dan Integrasi Data

Tahap seleksi dan integrasi dari atribut dan data yang dikumpulkan diperlukan untuk meningkatkan akurasi dan efisiensi dari algoritma yang digunakan. Kumpulan data dipilih dan dikurangi dengan mencari korelasi atau hubungan antar atribut yang paling kuat untuk mendapatkan atribut yang cukup dan hanya memiliki informasi yang diperlukan. Berikut adalah hasil dari proses seleksi dan integrasi data.

Tabel 3. 3 Atribut Seleksi dan Integrasi Data

No	Atribut	Tipe Data	Keterangan
1.	<i>Time spent on course</i>	<i>Numeric</i>	Lama waktu mahasiswa berada di mata kuliah
2.	<i>Course completed</i>	<i>Numeric</i>	Persentase kemajuan mahasiswa menyelesaikan mata kuliah
3.	Tugas	<i>Nominal</i>	Nilai rata-rata tugas 1-10
4.	UTS	<i>Nominal</i>	Nilai ujian tengah semester
5.	Quiz	<i>Polynomial</i>	Nilai quiz
6.	Simbol	<i>Binomial</i>	Skala penilaian / digunakan sebagai label target

b. Transformasi Data

Tahapan transformasi data dilakukan untuk mengubah format, struktur, atau nilai pada *dataset* mentah yang diperoleh. Tahapan ini dilakukan untuk mencapai data riwayat perkuliahan *online* mahasiswa UMKT yang bersih dan siap untuk dimasukkan ke pemodelan.

c. Reduksi Data

Tahapan reduksi data dilakukan untuk menghindari data yang tidak seimbang. Dimana kelompok kelas data yang lebih sedikit disebut minoritas, dan kelompok kelas data yang lebih besar disebut dengan kelompok mayoritas.

d. Pembersihan Data (Data Cleaning)

Tahapan data cleaning dilakukan untuk membersihkan data dari *noise* seperti data yang kosong, nilai ekstrim, dan nilai data yang salah.

Tahapan ini dilakukan untuk mencapai data riwayat perkuliahan *online* mahasiswa UMKT yang bersih dan siap untuk dimasukkan ke pemodelan.

2. Pemodelan

Tahap pemodelan yang digunakan pada penelitian ini adalah algoritma C4.5 untuk memprediksi performa perkuliahan *online* mahasiswa UMKT. Berikut adalah langkah-langkah penerapan algoritma C4.5 untuk membentuk pohon keputusan :

- a. Mempersiapkan data *training* yang dapat diambil dari data *history* yang pernah terjadi sebelumnya dan dikelompokkan dalam kelas-kelas tertentu.
- b. Menghitung nilai *entropy* dengan rumus

$$Entropy(S) \equiv \sum_{i=1}^n -p_i * \log_2 p \quad (3.1)$$

- c. Menentukan akar dari pohon dengan menghitung nilai *gain* yang tertinggi dari masing-masing atribut menggunakan rumus.

$$Gain(S, A) \equiv Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (3.2)$$

- d. Ulangi langkah ke-2 hingga semua record terpartisi
- e. Proses partisi pohon keputusan akan berhenti saat :
 - 1) Semua record dalam simpul N mendapat kelas yang sama.
 - 2) Tidak ada atribut di dalam record yang dipartisi lagi.
 - 3) Tidak ada record di dalam cabang yang kosong.

3. Evaluasi

Tahap evaluasi merupakan tahapan yang akan dilakukan untuk mengukur keakuratan dan kualitas *data training* yang dilakukan oleh pemodelan dari rasio data latih yang diberikan. Adapun teknik *confusion matrix* yang digunakan adalah *accuracy*.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.3)$$

3.5 Peralatan Yang Digunakan

Untuk membantu aktivitas penelitian yang dilakukan, kebutuhan peralatan yang digunakan yaitu *hardware* dan *software* sebagai berikut :

1. Perangkat Keras (*Hardware*)
 - a. Laptop Asus X441B
 - b. Processor AMD A9-9420 RADEON R5
 - c. RAM 4 GB
 - d. Mouse
2. Perangkat Lunak (*Software*)
 - a. Windows 2010
 - b. Pemrograman Python
 - c. *Library Sci-kit Learn (sklearn)*

3.6 Jadwal Penelitian

Tabel 3. 4 Jadwal Penelitian

No	Jenis Penelitian	Bulan / 2022					
		Jan	Feb	Mar	Apr	Mei	Jun
Tahap Pra Penelitian							
1.	Menentukan Judul Penelitian						
2.	Menyusun Rumusan Masalah						
3.	Mencari Data Pendukung						
4.	Menyusun Metode Penelitian						
5.	Menyusun Proposal Penelitian						
6.	Seminar Proposal						
Tahap Penelitian							
1.	Pengumpulan data						
2.	Analisis Data						
3.	Pengujian <i>Data Training</i> dan <i>Data Testing</i> dengan Algoritma C4.5						
Tahap Akhir Penelitian							
1.	Penyusunan Laporan						
2.	Evaluasi						

Tabel 3. 5 Jadwal Penelitian (Lanjutan)

No	Jenis Penelitian	Bulan / 2022					
		Jan	Feb	Mar	Apr	Mei	Jun
3.	Penyempurnaan Laporan						
4.	Seminar Hasil						