

BAB II

TINJAUAN PUSTAKA

2.1 Darmajaya

Darmajaya merupakan salah satu perguruan tinggi swasta yang di Jl. ZA Pagar Alam No. 93, Bandar Lampung dengan luas 2,5 hektar. Darmajaya berdiri pada tahun 1995 di bawah naungan yayasan Alfian Husein, dan saat ini IIB Darmajaya memiliki 15 Program pendidikan aktif dengan 11 program pendidikan sarjana dan 4 pascasarjana. Program Sarjana pada IIB Darmajaya meliputi (Akuntansi, Bisnis Digital, Desain Interior, Desain Komunikasi Visual, Hukum Bisnis, Manajemen, Pariwisata, Pendidikan Teknologi Informasi, Sains Data, Sistem Informasi, Sistem Komputer dan Teknik Informatika) dengan program pendidikan Pascasarjana yaitu (Manajemen, Manajemen Teknologi, dan Teknik Informatika).

2.2 Teknik Informatika

Teknik Informatika merupakan salah satu program studi yang terdapat pada perguruan tinggi Institut Informatika dan Bisnis Darmajaya yang didirikan pada tahun 1997, dengan salah satu misi program pendidikannya adalah Menyelenggarakan proses pendidikan dan pembelajaran yang berkualitas dengan metode terkini untuk menghasilkan sumberdaya manusia / lulusan yang berkarakter dan berwawasan nasional – internasional.

Jurusan Teknik Informatika di IIB Darmajaya memiliki tiga peminatan yaitu: *Multimedia, Artificial Intelligence, dan Mobile Technology.*

2.3 Tingkat Kepuasan

Mengukur kepuasan siswa dengan pembelajaran dapat mengukur sejauh mana proses tersebut pembelajaran berlangsung, dengan tindakan yang dilakukan yang dapat menentukan jalannya penilaian atau Perbaikan proses pembelajaran dilakukan (Triayudi & Soepriyono, 2022). Atas dasar ini, itu perlu proses pengukuran kepuasan. Proses pengukuran kepuasan mahasiswa pembelajaran terjadi terlebih dahulu dengan pengumpulan data. Setelah pengumpulan data

kemudian melakukan pengolahan data yang dilakukan untuk mendapatkan hasil yang diharapkan di dalam pengolahan data yang dilakukan mengalami masalah tersendiri, hal ini karena tidak penggunaan metode atau aturan pemrosesan data. Kesalahan dalam pemrosesan data selanjutnya.

Hasil yang diperoleh juga tidak sesuai dengan tujuan yang dikejar. Jadi untuk mengakhiri masalah harus diselesaikan sesuai dengan proses menggunakan metode atau teknik secara terpisah ketika metodenya adalah *data mining*.

2.4 Proses Pembelajaran

Belajar adalah investasi paling penting bagi negara manapun apalagi bagi negara berkembang untuk secara aktif membangun negaranya. Belajar tidak lepas dari tanggung jawab pendidik, bagaimana pendidik itu mengubah pengetahuan posesif dengan materi pendidikan yang ada, serta memperhatikan metode pengajaran yang mudah diterima oleh siswa. siswa untuk mencapai tujuan yang diinginkan. tujuan yang diinginkan Untuk lulus, dalam proses mengajar, pendidik harus melakukan suatu kegiatan evaluasi proses pembelajaran yang dilakukan.

2.5 Mahasiswa

Mahasiswa adalah konsumen/pelanggan suatu perguruan tinggi (universitas), maka konsep kepuasan mahasiswa dapat disamakan dengan kepuasan konsumen yaitu derajat persepsi setelah membandingkan kinerja atau hasil yang dirasakannya dengan harapannya, sedangkan Wilkie mendefinisikan kepuasan konsumen sebagai respons emosional ketika mengevaluasi pengalaman konsumen tentang produk atau layanan (Rahmawati, 2013b).

2.6 Artificial Intelligence

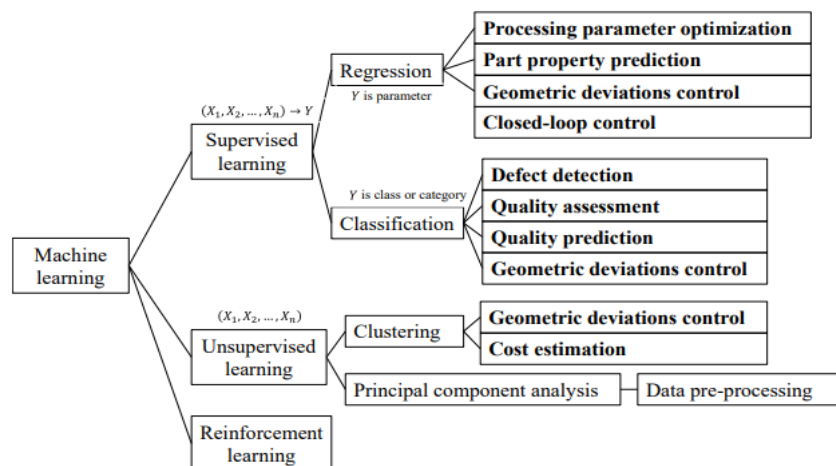
Artificial Intelligence (AI) atau dalam bahasa Indonesia dikenal sebagai Kecerdasan Buatan, adalah suatu bidang ilmu komputer yang bertujuan untuk membuat mesin atau sistem komputer dapat melakukan tugas yang biasanya memerlukan kecerdasan manusia (Novamizanti & Siadari, 2022). AI mencakup pengembangan algoritma dan model komputasional yang dirancang untuk memungkinkan mesin untuk "belajar" dari data, mengenali pola, mengambil keputusan, dan melakukan tugas-tugas yang cerdas (Markiewicz & Zheng, 2020).

2.6.1 Machine Learning

Machine Learning adalah bagian dari Kecerdasan Buatan (AI) yang melibatkan penggunaan algoritme dan model statistik untuk memungkinkan komputer atau mesin belajar dari dan membuat prediksi atau keputusan berdasarkan data tanpa diprogram secara eksplisit (Carleo dkk., 2019). Dengan kata lain, ini adalah metode atau pendekatan di mana mesin dapat secara otomatis belajar dari pengalaman, meningkat seiring waktu, dan melakukan tugas tanpa diprogram secara eksplisit untuk setiap tugas tertentu.

Pembelajaran Mesin melibatkan penggunaan sejumlah besar data, yang dimasukkan ke dalam algoritme dan model untuk melatih mesin. Selama proses pelatihan, mesin mengidentifikasi pola, tren, dan hubungan dalam data dan menyesuaikan parameter modelnya untuk mengoptimalkan kinerjanya. Setelah mesin dilatih, ia dapat menggunakan pengetahuan yang dipelajarinya untuk membuat prediksi atau keputusan pada data baru yang tidak terlihat (Zhou dkk., 2017).

Menurut (Meng dkk., 2020) *Machine Learning* dapat dikategorikan menjadi tiga jenis seperti pada gambar berikut:



Gambar 2. 1 Jenis *Machine Learning*

1. *Supervised Learning*

Dalam jenis Pembelajaran Mesin ini, model dilatih pada data berlabel, di mana data masukan disertai dengan label keluaran yang sesuai. Model belajar dari data berlabel untuk membuat prediksi atau mengklasifikasikan data baru yang tidak terlihat ke dalam kategori yang telah ditentukan sebelumnya.

2. *Unsupervised Learning*

Dalam jenis Pembelajaran Mesin ini, model dilatih pada data yang tidak berlabel, di mana data masukan tidak memiliki label keluaran yang menyertainya. Model belajar mengidentifikasi pola atau struktur dalam data, seperti pengelompokan atau pengurangan dimensi.

3. *Reinforcement Learning*

Dalam Pembelajaran Mesin jenis ini, model belajar dari interaksi dengan lingkungan dan menerima umpan balik dalam bentuk penghargaan atau hukuman. Model belajar mengambil tindakan di lingkungan untuk memaksimalkan imbalan kumulatif dari waktu ke waktu.

2.6.2 *Data Mining*

Data mining adalah suatu proses ekstraksi dan analisis data yang kompleks, besar, atau beragam untuk mengidentifikasi pola, tren, atau hubungan yang bermanfaat, serta mendapatkan informasi yang berharga atau pengetahuan yang dapat digunakan untuk pengambilan keputusan. *Data mining* melibatkan penggunaan teknik dan metode analisis statistik, matematika, dan komputer untuk menggali dan menginterpretasi data yang ada, mengenali pola tersembunyi atau hubungan yang dapat digunakan untuk mengambil keputusan yang lebih baik atau mengambil tindakan yang tepat (Miranda, 2011).

Data mining sering digunakan dalam berbagai bidang, termasuk bisnis, keuangan, kesehatan, pemasaran, ilmu sosial, ilmu komputer, dan banyak lagi (Fatimah & Saidah, 2021; Khairani & Irwansyah, 2018; Mahmud &

Hartanto, 2020; Nuryani & Darwis, 2021). *Data mining* dapat membantu menggali wawasan yang berharga dari data yang besar, rumit, atau beragam, yang tidak dapat ditemukan dengan metode analisis tradisional (Jollyta dkk., 2020). Metode *data mining* melibatkan langkah-langkah seperti pemilihan data, preprocessing data, pemilihan metode analisis, evaluasi model, dan interpretasi hasil. Menurut (Wanto dkk., 2020) salah satu teknik umum yang digunakan dalam *data mining* adalah regresi.

2.7 Metode Regresi

Metode regresi digunakan untuk memprediksi nilai numerik atau kontinu dari suatu variabel berdasarkan pola yang ditemukan pada data yang telah diprediksi sebelumnya (Wanto dkk., 2020). Metode ini bertujuan untuk menemukan persamaan matematis yang menggambarkan hubungan antara variabel independen (prediktor) dan variabel dependen (target) dalam bentuk fungsi regresi. Metode regresi dapat digunakan untuk melakukan prediksi atau estimasi nilai variabel dependen berdasarkan nilai variabel independen yang diberikan.

Dalam metode regresi, *Random Forest Regression* merupakan salah satu model regresi berdasarkan teknik *supervised learning* menggunakan fitur-fitur dari suatu *history data*. Hubungan antara fitur dan target akan diwakili oleh serangkaian kondisi yang terkait dan diatur dalam struktur seperti pohon dari atas ke bawah (Diana Tri Susetianingtias dkk., 2022).

Penerapan metode *Random Forest Regression* dalam penelitian kepuasan mahasiswa terhadap proses pembelajaran pada jurusan Teknik Informatika dapat membantu mengidentifikasi faktor-faktor yang berpengaruh terhadap kepuasan mahasiswa, termasuk hubungan non-linear yang mungkin ada antara variabel prediktor dan variabel target. Dengan memahami pengaruh faktor-faktor ini secara formal, dapat diambil langkah-langkah yang lebih tepat dalam meningkatkan kualitas pembelajaran di jurusan tersebut, sehingga dapat meningkatkan kepuasan mahasiswa secara efektif.

2.8 Coefficient of Determination-R²

Coefficient of Determination atau *R² Score* digunakan untuk mengevaluasi performa model pembelajaran mesin berbasis regresi yang telah dibentuk. Model prediksi yang dibentuk dengan menggunakan atribut dari variabel terikat dan variabel bebas, nantinya *Coefficient of Determination* atau *R² Score* akan digunakan untuk memeriksa seberapa baik performa model yang telah dibentuk (Ardiyah Nurul, 2019). Dengan rumus rasio deviasi dapat dilihat pada persamaan 2.1

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \quad (2.1)$$

Keterangan:

SS res: jumlah kuadrat dari kesalahan residual.

SS tot: jumlah total kesalahan.

2.9 Perangkat Lunak

Perangkat Lunak merupakan seluruh perintah yang digunakan untuk memproses informasi. Perangkat lunak dapat berupa program maupun prosedur yang didalamnya merupakan kumpulan perintah yang dimengerti oleh komputer sedangkan prosedur adalah perintah yang dibutuhkan oleh pengguna dalam memproses informasi (Swara dkk., 2016).

2.10 Streamlit

Streamlit merupakan *library* yang ada di bahasa pemrograman *Python* yang dapat dengan mudah untuk *deploy machine learning* maupun *data science*. *Streamlit* sangat cocok bagi mereka yang terjun di dunia data dan mereka ingin membuat sebuah *project* yang tergabung dalam sebuah tim (Irawansyah, t.t.)

2.11 Visual Studio Code

Visual Studio Code adalah kode editor sumber yang dikembangkan oleh Microsoft untuk Windows, Linux dan macOS. Ini termasuk dukungan untuk debugging, kontrol git yang tertanam dan GitHub, penyorotan sintaksis, penyelesaian kode cerdas, snippet, dan refactoring kode. Ini sangat dapat

disesuaikan, memungkinkan pengguna untuk mengubah tema, pintasan keyboard, preferensi, dan menginstal ekstensi yang menambah fungsionalitas tambahan (Bacodkk., 2022).

2.12 Python


Python adalah bahasa pemrograman interpretatif multiguna dengan filosofi perancangan yang berfokus pada tingkat keterbacaan kode. *Python* diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif (Syahrudin & Kurniawan, 2018)



2.13 Rancangan Perangkat Lunak

2.13.1 Use Case Diagram

Use Case diagram merupakan deskripsi peringkat tinggi bagaimana perangkat (aplikasi) akan digunakan oleh penggunanya. Selanjutnya, *Use Case* tidak hanya sangat penting pada tahap analisis, tetapi juga sangat penting untuk perancangan, untuk mencari kelas-kelas yang terlibat dalam aplikasi, serta untuk melakukan pengujian (Arfida dkk., t.t.). *Use case* diagram merupakan pemodelan untuk kelakuan sistem informasi yang akan dibuat. *Use case* mendeskripsikan sebuah interaksi antara satu atau lebih aktor dengan sistem informasi yang akan dibuat. *Use case* digunakan untuk mengetahui fungsi apa saja yang ada di dalam sebuah sistem informasi dan siapa saja yang berhak menggunakan fungsi-fungsi itu. Berikut adalah simbol-simbol yang ada pada *use case* yang dapat dilihat pada tabel 2.1.

Tabel 2. 1 Simbol *Use Case Diagram*



Simbol	Nama	Keterangan
	<i>Actor</i>	Menspesifikasikan himpunan peran yang pengguna mainkan ketika berinteraksi dengan <i>use case</i>

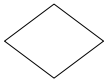


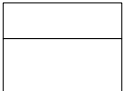
Simbol	Nama	Keterangan
----->	<i>Dependency</i>	Hubungan dimana perubahan yang terjadi pada suatu elemen mandiri (<i>Independent</i>) akan mempengaruhi elemen yang bergantung padanya elemen yang tidak mandiri
—————>	<i>Generalization</i>	Hubungan dimana objek anak (<i>descendent</i>) berbagi perilaku dan struktur data dari objek yang ada di atasnya objek induk(<i>ancestor</i>)
-----<<include>>----->	<i>Include</i>	Menspesifikasikan bahwa <i>use case</i> sumber secara <i>eksplisit</i>
-----<<extend>>----->	<i>Extend</i>	Menspesifikasikan bahwa <i>use case</i> target memperluas perilaku dan <i>use case</i> sumber pada suatu titik yang diberikan
—————	<i>Association</i>	Apa yang menghubungkan antara objek satu dengan objek lainnya
	<i>System</i>	Menspesifikasikan paket yang menampilkan sistem secara terbatas
	<i>Usecase</i>	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan suatu hasil yang terukur bagi suatu aktor

2.13.2 Activity Diagram

Activity Diagram adalah teknik untuk mendeskripsikan logika prosedural, proses bisnis dan aliran kerja dalam banyak kasus. *Activity diagram* mempunyai peran seperti halnya *flowchart*, akan tetapi perbedaannya dengan *flowchart* adalah *activity diagram* bisa mendukung perilaku paralel sedangkan *flowchart* tidak bisa (Hasugian & Shidiq, 2012). Dalam kata lain *activity diagram* menggambarkan berbagai aliran aktivitas dalam sistem yang sedang dirancang, bagaimana masing-masing aliran berawal, decision yang mungkin terjadi, dan bagaimana mereka berakhir. *Activity diagram* juga dapat menggambarkan proses paralel yang mungkin terjadi pada beberapa eksekusi. *Activity diagram* merupakan *state diagram* khusus, dimana sebagian besar *state* adalah *action* dan sebagian besar transisi di-trigger oleh selesainya *state* sebelumnya (*internal processing*) (Darmawan dkk., 2018). Oleh karena itu *activity diagram* tidak menggambarkan *behavior internal* sebuah sistem (interaksi antar subsistem) secara eksak, tetapi lebih menggambarkan proses - proses serta jalur - jalur aktivitas dari level atas ke umum. Sebuah aktivitas dapat direalisasikan oleh satu *use case* atau lebih. Aktivitas menggambarkan proses yang berjalan, sementara *use case* menggambarkan bagaimana *actor* menggunakan sistem untuk melakukan aktivitas. Berikut adalah simbol-simbol yang ada pada *activity diagram* yang dapat dilihat pada tabel 2.2.

Tabel 2. 2 Simbol *Activity Diagram*

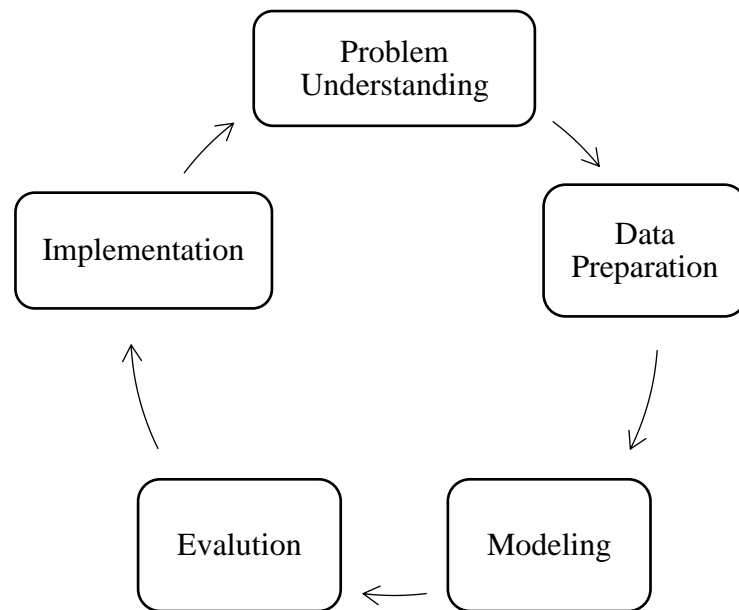
Simbol	Nama	Keterangan
	Status Awal	Sebuah diagram aktivitas memiliki sebuah status awal
	Aktivitas	Aktivitas yang dilakukan sistem, aktivitas biasanya diawali dengan kata kerja

Simbol	Nama	Keterangan
	Percabangan	Percabangan dimana ada pilihan aktivitas yang lebih dari satu.
	Penggabungan/join	Penggabungan dimana yang mana aktivitas lebih dari satu aktivitas lalu digabungkan menjadi satu
	Status Akhir	Status akhir yang dilakukan sistem, sebuah diagram aktifitas memiliki sebuah status akhir
	Swimlane	Swimlane memisahkan organisasi bisnis yang bertanggung jawab terhadap aktivitas yang terjadi.

2.14 Metode Pengembangan *Machine Learning Life Cycle*

Berdasarkan studi literature yang dilakukan oleh peneliti maka, dapat diberikan analisa bahwa dalam penelitian yang akan dilakukan ini terdapat kerangka pemikiran yang dijadikan untung menunjang keberlangsungannya penelitian ini yang disebut dengan *Machine Learning Life Cycle*.

Machine Learning Life Cycle merupakan salah satu model pengembangan model *machine learning* pada kecerdasan buatan (AI). Tujuan dari *machine learning life cycle* adalah untuk menemukan solusi untuk masalah yang diberikan dengan menerapkan model *Machine Learning*. *Machine Learning Life Cycle* dibagi menjadi lima tahap utama yaitu:



Gambar 2. 2 *Machine Learning Life Cycle*

Pada tahapan pembuatan model klasifikasi *Machine Learning* menggunakan *Machine Learning Life Cycle* memiliki tahapan-tahapan sebagai berikut:

1. *Problem Understanding*

Tahap ini melibatkan pemahaman yang komprehensif tentang masalah yang akan dipecahkan, termasuk identifikasi jenis masalah, pemahaman konteks penggunaan model, serta penetapan batasan dan tujuan yang jelas.

2. *Persiapan Data*

Pada tahap ini dilakukan persiapan data untuk mempersiapkan data yang akan digunakan pada penelitian agar kualitas data yang digunakan meningkat dan memperbaiki proses pemodelan (Nasrullah, 2018). Terdapat empat langkah utama pada persiapan data yaitu:

- *Mengumpulkan Data*

Berdasarkan spesifikasi yang telah ditentukan dalam langkah *problem understanding* maka dilakukannya pengumpulan data dari sumber yang telah ditentukan. Biasanya data

disimpan dalam bentuk big data ataupun dataset tertentu yang sesuai dengan sistem pengolahan data yang akan dipergunakan (Dewi dkk., 2021).

- **Memilih dan Memilah Data**

Data dipilih berdasarkan kualitas data dan tujuan penelitian yang akan dilakukan. Pemilihan data dilakukan berdasarkan record ataupun atribut (fitur). Pemilihan berdasarkan atribut dilakukan dengan menghilangkan atribut yang dianggap tidak layak untuk dipergunakan karena sifat atribut tersebut ataupun karena mirip sifatnya dengan atribut lain ataupun tidak ada kaitannya dengan fitur label (Iriadi & Nuraeni, 2016).

- **Membersihkan Data**

Melakukan pra-pemrosesan data seperti normalisasi, penghilangan outlier, dan pemilihan fitur jika diperlukan (Rahmawan, 2020).

- **Merekonstruksi Data**

Merekonstruksi data dilakukan agar data yang akan digunakan untuk penelitian akan cocok dengan algoritma yang akan dipakai dengan menggunakan teknik transformasi data. Transformasi data adalah proses perubahan data ke jenis lain. Data numerik diubah ke dalam bucket (kelompok data), perubahan ke data kategorial, atau data kategorial diubah menjadi berbagai atribut dummy, dll (Sidabutar, 2020).

3. Membuat Model

Membentuk suatu model dari data yang sudah diperbaiki/disiapkan dengan menggunakan pengembangan Machine Learning dengan menentukan algoritma yang akan dipakai, penentuan algoritma ini diharapkan dapat memberikan hasil penelitian terbaik meski dengan jumlah percobaan penelitian yang sedikit.

Pembangunan skenario ini dilakukan agar usaha yang dilakukan sedikit mungkin untuk mendapatkan model terbaik. Langkahnya adalah sebagai berikut:

- Melakukan pemilihan algoritma yang sesuai, berdasarkan hasil analisa permasalahan yang telah dilakukan dan data hasil wawancara dan kuisioner, maka pada penelitian ini algoritma yang digunakan adalah algoritma regresi non linear.
- Membagi data: Pembagian data latih (dipergunakan untuk membentuk model) dengan data uji (dipergunakan untuk mengukur performansi dari model yang telah dibuat), dengan perbandingan 80:20 atau 70:30 antara data latih dengan data uji yang dipergunakan.

4. *Evaluation*

Tahap ini dilakukan sebagai pengukuran kinerja model menggunakan metrik evaluasi yang sesuai. Terdapat dua poin utama yang menjadi evaluasi sebagai berikut:

- **Evaluasi Model**
Pada langkah ini, evaluasi model dilakukan untuk mengukur tingkat akurasi hasil penelitian dengan algoritma regresi non linear (Spjuth dkk., 2021).
- **Evaluasi Proses**
Evaluasi pekerjaan yang telah dilakukan untuk mengidentifikasi area di mana proses dapat ditingkatkan, dan membuat model menjadi lebih baik. Setiap langkah yang telah diambil diperiksa untuk menentukan apakah ada tindakan yang dilakukan agar lebih efektif daripada yang diperlukan. Jika tidak ada yang dianggap terlewatkan, model yang dibuat telah menggunakan data yang diperoleh dengan semaksimal mungkin (Spjuth dkk., 2021).

5. *Implementation*

Dalam tahap ini dilakukan dengan mengimplementasikan model dalam lingkungan produksi atau aplikasi yang diinginkan, termasuk integrasi model ke dalam sistem produksi, pengujian ulang model di lingkungan produksi, pengelolaan versi model, serta pemeliharaan model untuk memastikan kinerjanya tetap optimal seiring waktu (Spjuth dkk., 2021).

2.15 Penelitian Terkait

Tabel 2. 3 Penelitian Terdahulu

Judul, Penulis dan Tahun	Jumlah dan Atribut Dataset	Algoritma	Validasi		Open Source Dataset	Akurasi
Mengukur Tingkat Kepuasan Mahasiswa dalam Pembelajaran dengan Naïve Bayes (Siregar & Rofik, 2021)	159 data dengan 31 atribut	Naïve Bayes	Confusion Matrix	-	-	86%
Algoritma Decision Tree (C4.5) untuk memprediksi kepuasan Mahasiswa Terhadap Kinerja Dosen Politeknik TEDC Bandung (Yuliana & Pratomo, 2017)	90 data dengan 5 atribut	Decision Tree	Confusion Matrix	-	-	96,99%
Model Klasifikasi Kepuasan Mahasiswa Teknik Terhadap Sarana Pembelajaran Menggunakan Data Mining (Lestari dkk., 2020)	948 data	Decision tree C4.5, Naïve	Confusion Matrix	-	-	98%

Judul, Penulis dan Tahun	Jumlah dan Atribut Dataset	Algoritma	Validasi		Open Source Dataset	Akurasi
	dengan 8 atribut	Bayes, dan Support Vector Machine				
Analisis Faktor-faktor yang mempengaruhi Kepuasan Mahasiswa (Rahmawati, 2013a)	64 data dengan 7 atribut			-	-	

Berdasarkan tabel diatas Penelitian sebelumnya yang telah dilakukan mengenai faktor-faktor yang mempengaruhi kepuasan mahasiswa terhadap proses pembelajaran pada jurusan Teknik Informatika masih terbatas dan belum sepenuhnya menggunakan pendekatan *data mining* dan metode regresi. Oleh karena itu, penelitian ini dapat menjadi kontribusi baru dalam menggali lebih dalam faktor-faktor yang mempengaruhi tingkat kepuasan mahasiswa pada jurusan Teknik Informatika menggunakan pendekatan *data mining* dan metode regresi.