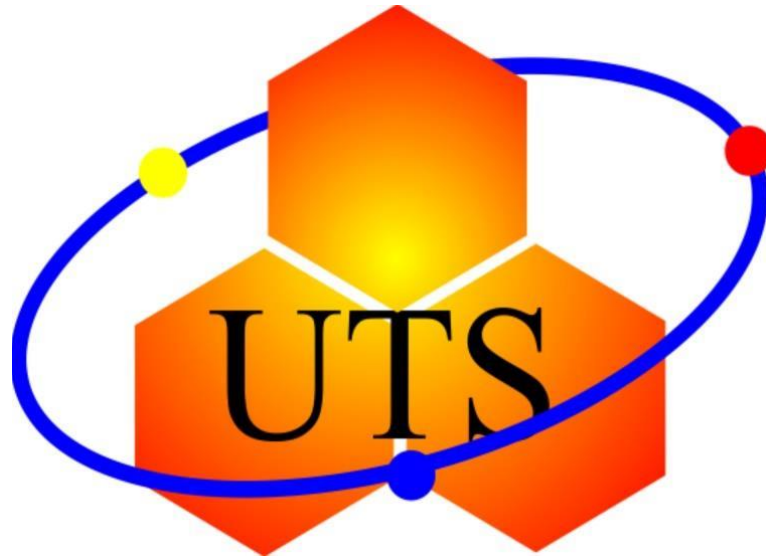


**PENERAPAN ALGORITMA DT C4.5 UNTUK MEMPREDEKSI
LULUS DAN TIDAKNYA MENTORING MAHASISWA PADA
UNIVERSITAS TEKNOLOGI SUMBAWA**



Oleh :

Nurmala 20.01.013.069

Asti Oktaviani 20.01.013.003

Nurul Mufliha Puasa 20.01.013.014

Armanita Aning 20.01.013.002

Ambar Wati 20.01.013.001

Rismayanti 20.01.013.025

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS REKAYASA SISTEM
UNIVERSITAS TEKNOLOGI SUMBAWA
2022**

ABSTRACT

Mentoring adalah suatu hubungan antara seseorang yang lebih berpengalaman dengan seseorang yang kurang pengalamannya. Literatur menunjukkan bahwa program mentoring akan memberikan dukungan untuk kesuksesan akademik dan pengembangan karakter diri. Program mentoring mulai dilaksanakan pada semester genap 2019 – 2020 dan merupakan hal yang baru untuk mahasiswa. Perlu untuk mengetahui gambaran program mentoring dalam bentuk data naratif. Tujuan penelitian adalah untuk menggambarkan manfaat dari program mentoring. Desain penelitian adalah studi deskriptif kualitatif. Jumlah partisipan dalam penelitian adalah 8 mentor dan 55 mentee.

Data kualitatif naratif dianalisis dengan menggunakan analisis isi. Hasil analisis data naratif dari mentor tentang manfaat program mentoring menghasilkan tiga tema yaitu belajar bersama, keterampilan komunikasi yang efektif, dan rasa peduli kepada orang lain. Hasil analisis data naratif dari mentee tentang manfaat program mentoring menghasilkan tiga tema yaitu saling menghormati, rasa kekeluargaan, dan pembelajaran akademik yang efektif. Diharapkan program mentoring dapat dikembangkan lebih baik lagi khususnya untuk program perekrutan dan pelatihan yang berkelanjutan bagi mentor.

KATA PENGANTAR

Puji syukur kami ucapkan kehadiran Allah SWT yang telah melimpahkan rahmat, karunia, dan hidayah-Nya, sehingga kami dapat menyelesaikan laporan yang berjudul “Penerapan Algoritma DT C4.5 Untuk Memprediksi Lulus Dan Tidaknya Mentoring Mahasiswa Pada Universitas Teknologi Sumbawa” Penyusunan laporan ini digunakan untuk memenuhi tugas mata kuliah Kecerdasan Buatan yang dibimbing oleh Bapak Herfandi M.Kom. Pada kesempatan ini perkenankan kami mengucapkan terima kasih kepada: Bapak Herfandi M.Kom. selaku Dosen pengampuh mata kuliah Kecerdasan Buatan. Kami sadar dalam makalah ini masih terdapat kekurangan dan kesalahan, kami mohon maaf dan meminta kepada Bapak dosen, kiranya bersedia memberikan kritik dan saran untuk perbaikan selanjutnya. Sekian dari kami semoga tugas ini sesuai dengan apa yang diharapkan dan dapat bermanfaat bagi yang membacanya.

Sumbawa, 3 Januari 2022

Penyusun,

DAFTAR ISI

	Halaman
COVER	
ABSTRAK	i
KATA PENGANTAR	ii
BAB I PENDAHULUAN	1
1.1.Latar Belakang	1
1.2.Rumusan Masalah	1
1.3.Batasan Masalah.....	2-3
1.4.Tujuan Penelitian	3
1.5.Manfaat Penelitian	3
1.6.Sistematika Penulisan	4
BAB II LANDASAN TEORI	5
2.1 Tinjauan Pustaka	5
2.2 Dasar Teori.....	5-12
BAB III METODOLOGI PENELITIAN	
3.1 Pengumpulan Data	13
3.2 Pengelolahan Data Awal	1
BAB IV PEMBAHASAN	15
4.1 Data Understanding	15
4.2 Business Understanding.....	15
4.3 Data Preparation	15
BAB V PENUTUP	20
5.1 Kesimpulan	20
5.2 Saran.....	28
DAFTAR PUSTAKA	29

BAB I

PENDAHULUAN

1.1.Latar Belakang

Sistem informasi saat ini semakin berkembang dengan pesat dari waktu ke waktu yang mendorong penggunaan computer di berbagai sektor kehidupan. Mulai dari instansi pemerintah, industry, bisnis dan perbankan hingga sektor Pendidikan seperti Universitas. Komputer nampaknya telah menjadi suatu kebutuhan disetiap sektor kehidupan. Begitu pula dengan pendidikan yang ada di Universitas Teknologi Sumbawa, untuk mengembangkan perlu adanya perancangan sebuah sistem terkomputerisasi mentoring (monev). Monev sendiri adalah gabungan dari mentoring dan evaluasi yang merupakan salah satu kegiatan yang dilakukna di Universitas Teknologi Sumbawa yang dikhususkan bagi mahasiswa jalur beasiswa.

Usaha pencapaian tujuan program dalam sebuah mentoring harus selalu diupayakan oleh pihak kampus. Upaya tersebut dimulai dari menyusun rencana strategis jangka panjang 5 tahunan yang memuat penetapan visi, misi dan tujuan mentoring. Visi, misi dan tujuan strategis mentoring merupakan komitmen bersama seluruh mahasiswa di dalam kampus untuk mewujudkannya. Rencana strategis selanjutnya mengevaluasi dalam satu tahunan yaitu dengan menjabarkan Visi, misi dan tujuan menjadi sasaran jangka pendek dalam program kegiatan mentoring.

Berbagai program yang telah akan dilaksanakan memerlukan peningkatan kinerja pimpinan BPKO panitia dan mentor baik dalam sistem perencanaan, pelaksanaan maupun evaluasinya. Keberhasilan suatu program dapat dilihat dari kesesuaian antara perencanaan dan pelaksanaannya, terukur atau akuntabel hasilnya, serta ada keberlanjutan aktivitas yang merupakan dampak dari program itu sendiri. Melalui kegiatan mentoring dan evaluasi

(monev) maka keberhasilan, dampak dan kendala pelaksanaan suatu program dapat diketahui. Ditinjau dari aspek pelaksanaan, monev memerlukan keterampilan petugas. Petugas adalah seorang evaluator yang terampil untuk mengumpulkan berbagai data yang sesuai dengan tujuan mentoring dan evaluasi. Selain itu, kejujuran, keuletan, dan penguasaan pengetahuan tentang mentoring dan evaluasi menjadi tuntutan kualifikasi petugas. Bila ditinjau dari aspek sistem mentoring dan evaluasi, maka panitia mentor yang terlibat dalam kegiatan ini harus mampu merencanakan, menyiapkan, melaksanakan dan melaporkan seluruh kegiatan mentoring dan evaluasi.

Monev dilakukan oleh panitia mentor yang profesional, dan didukung dengan instrumen yang baku akan dapat diperoleh data objektif. Data objektif yang dianalisis dengan teknik yang tepat akan mendapatkan informasi yang terpercaya untuk dasar pengambilan keputusan manajemen. Sehingga keputusan yang diambil tepat untuk mencapai tujuan yang telah ditetapkan. Monev dilakukan setiap semester dengan menentukan syarat-syarat seperti nilai Indeks Prestasi (IP) minimal 3,00 (40%), aktif pada kegiatan kemahasiswaan (20%), Mengikuti kegiatan program pembinaan terpadu mahasiswa (20%), Memiliki Prestasi non akademik (20%). Apabila dari keempat syarat tersebut tidak mencapai skor 60% maka mahasiswa dinyatakan tidak lulus Mentoring.

Monev bertujuan untuk mengamati perkembangan dan kemajuan akademik serta menentukan lulus atau tidak lulus mahasiswa jalur beasiswa, apabila mahasiswa tidak lulus maka akan dikenakan tarif biaya yang telah ditentukan oleh pihak UPT Beasiswa Universitas Teknologi Sumbawa. Dalam hal ini penulis menerapkan algoritma C4.5 dalam proses pembuatan dataset klasifikasi Monev.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan diatas, dapat dirumuskan pertanyaan penelitian antara lain yaitu bagaimana mengoptimalkan proses mentoring dan evaluasi pembelajaran di Universitas Teknologi Sumbawa, pelaksanaan dan laporan hasil Monev serta menentukan metode penelitian dengan menggunakan klasifikasi dataset.

1.3. Batasan Masalah

Dengan adanya rumusan masalah tersebut, maka akan dibuat sistem informasi dengan batasan masalah sebagai berikut ;

1. Pembuatan sistem informasi mentoring mahasiswa berbasis web ini, difokuskan pada kebutuhan mahasiswa yang ada di Universitas Teknologi Sumbawa.
2. Penelitian ini hanya dalam lingkup mentoring kedisiplinan para anggota mentoring.
3. Output dari sistem informasi ini berupa laporan poin-point baik itu pelanggaran, kehadiran maupun point untuk tugas yang telah diberikan.

1.4. Tujuan Penelitian

Tujuan mentoring dan evaluasi monev ini adalah memberi pemahaman tentang langkah-langkah persiapan, perencanaan, pelaksanaan dan pelaporan hasil evaluasi program mentoring. Hasil monev merupakan informasi berharga yang dapat dijadikan pedoman bagi pimpinan BPKO untuk mengambil keputusan dalam pengembangan kegiatan mentoring tersebut.

1.5. Manfaat Penelitian

Manfaat dari penelitian pembuatan sistem informasi mentoring siswa berbasis web Universitas Teknologi Sumbawa adalah sebagai berikut:

1. Membantu pekerjaan BPKO untuk mentoring mahasiswa dengan mudah, serta dalam penyimpanan data dan perekapan data dengan menggunakan sistem terkomputerisasi.

2. Mahasiswa dapat melihat dan memantau poin yang dimilikinya kapan saja dimana saja melalui website.
3. Pelaporan mentoring mahasiswa di Universitas Teknologi Sumbawa akan rutin diterima oleh mahasiswa setiap satu semester.
4. Manfaat penelitian berisi tentang kegunaan hasil penelitian. Manfaat bersifat lebih praktis dan efisien.

1.6. Sistematika Penulisan

BAB I PENDAHULUAN

Menjelaskan tentang latar belakang penelitian, mengapa topic ini menjadi pilihan penulis, tujuan penelitian, batasan penelitian serta metodologi penelitian yang digunakan.

BAB II TINJAUAN PUSTAKA

Pada bab ini dijelaskan tentang data mining yang berfungsi untuk mendapatkan informasi baru dan dapat membantu dalam pengambilan keputusan. Algoritma C4.5 berfungsi untuk memprediksi nilai atribut bertipe diskret dari record yang baru.

BAB III METODOLOGI PENELITIAN

Pada bab ini menjelaskan tentang metode penelitian apa yang digunakan penulis dalam pembuatan laporan.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini membahas apa yang menjadi tujuan, hasil dari peneliti dan juga berisi tentang bagaimana data yang diperoleh memberikan solusi pada permasalahan yang ingin dipecahkan

BAB V PENUTUP

Bab ini berisi kesimpulan

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

2.1.1 Pengertian Data Mining

Saat ini Data Mining menjadi salah satu pusat perhatian para akademis maupun praktisi. Menurut (Suntoro, 2019) data mining adalah proses untuk mendapatkan informasi yang berguna dari basis data yang besar dan perlu diekstraksi agar menjadi informasi baru dan dapat membantu dalam pengambilan keputusan. Data mining adalah proses menganalisa data dari yang berbeda dan menyimpulkannya menjadi informasi atau pengetahuan atau pola yang penting untuk meningkatkan keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya (Witten, 2016).

2.1.2 Tipe Data Mining yaitu :

- a. Tipe Data Numerik adalah tipe data yang diperoleh dengan cara pengukuran, dimana jarak dua titik pada skala sudah diketahui.
- b. Tipe Data Kategorial adalah tipe data yang diperoleh dengan cara kategorisasi atau klasifikasi.
- c. Tipe Data Rentang Waktu adalah tipe yang diperoleh dengan cara menunjukan beberapa objek yang berbeda.

2.1.3 Manfaat Data Mining

Menurut Tan 2004 dalam (Vulandari S., 2017) Pemanfaatan dari data mining sendiri bisa dilihat dari dua sudut pandang, baik sudut pandang komersial dan sudut pandang keilmuan. Dari sudut pandang komersial, data mining bisa digunakan untuk menangani adanya peledakan dari volume data. Dengan melihat bagaimana menyimpannya, mengekstraknya dan memanfaatkannya. Tentunya berbagai ilmu komputasi dapat untuk menghasilkan informasi yang

dibutuhkan. Selain itu data mining juga bisa dimanfaatkan untuk menyelesaikan masalah dalam kebutuhan dibidang bisnis, misalnya :

- 1) Mengetahui hilangnya pelanggan dikarenakan adanya pesaing.
- 2) Memprediksi dan mengetahui siswa yang tidak lulus dan yang lulus.
- 3) Mengetahui item suatu produk yang memiliki kesamaan karakteristik.
- 4) Mengidentifikasi produk-produk yang sudah terjual dengan produk lainnya.

2.1.4 Tahap-Tahap Data Mining

Menurut Tan 2004 dalam (Vulandari S. M., Data mining teori dan aplikasi rapidminer, 2017), serangkaian proses tahapan memiliki tujuh 7 tahapan yaitu:

a. Pembersihan Data (data cleaning)

Pembersihan data adalah proses untuk menghilangkan data-data yang tidak relevan. Data-data yang dibuang terkadang dibandingkan terlebih dahulu dengan hipotesa yang telah dibuat. Sehingga pada proses selanjutnya dapat dengan mudah menemukan hasil yang diinginkan.

b. Integrasi data (data integration)

Integrasi data merupakan proses dalam menggabungkan data dari beberapa database kedalam satu database baru. Tidak sedikit data yang dibutuhkan diambil dari berbagai database atau teks file.

c. Seleksi data (data selection)

Data yang sudah ada di database seringkali tidak semuanya dibutuhkan, maka dari itu dibutuhkan penyeleksian data untuk data yang benar-benar dibutuhkan dalam proses selanjutnya.

d. Transformasi data (data transformation)

Data digabung atau diubah sesuai dengan proses yang digunakan dalam data mining. Karena beberapa format data

mining membutuhkan format data yang khusus dalam pemrosesannya.

e. Proses mining

Adalah proses menggali data dari sebuah database atau kumpulan data untuk memperoleh informasi yang tersembunyi dari data yang diolah

f. Evaluasi Pola (pattern evaluation)

Dalam proses ini adalah hasil dari teknik data mining berupa pola-pola yang akan diujia pada hipotesa yang sudah dibuat sebelumnya. Sehingga akan memperoleh kesimpulan-kesimpulan yang mendekati hasil atau hipotesa untuk proses selanjutnya.

g. Presentasi pengetahuan (knowledge presentation)

Ini termasuk dalam langkah akhir dari data mining dalam tahap ini saatnya untuk mempresentasikan hasil yang telah dilakukan dengan mengimplementasikan analisis yang didapat. Sehingga akan memperoleh kesimpulan real.

2.1.5 Fungsi Data Mining

Menurut (Hasket, 2017) Fungsi umum dalam data mining yaitu :

- a) Association, adalah proses untuk menemukan aturan asosiasi Antara suatu kombinasi item dalam suatu waktu.
- b) Sequence, adalah proses untuk menentukan aturan asosiasi Antara suatu kombinasi item dalam suatu waktu dan diterapkan lebih dari satu periode.
- c) Clustering, adalah proses pengelompokan sejumlah data/objek ke dalam kelompok data sehingga setiap kelompok berisi data yang mirip.
- d) Classification, adalah proses penemuan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui.

- e) Regression, adalah proses pemetaan data dalam suatu nilai prediksi.
- f) Forecasting, adalah proses pengestimasian nilai prediksi berdasarkan pola-pola di dalam sekumpulan data.
- g) Solution, adalah proses penemuan akar masalah dan problem solving dari persoalan bisnis yang diharapkan atau paling tidak sebagai informasi dalam pengambilan keputusan.

2.1.6 Tugas/Pekerjaan dalam data mining

Menurut (Rerung, 2018) Data mining dibagi menjadi beberapa kelompok berdasarkan tugas/pekerjaan yang dapat dilakukan yaitu :

1. Deskripsi: terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.
2. Estimasi: hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan baris data (record) lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.
3. Prediksi: hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.
4. Klasifikasi: terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga

kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

5. Pengklasteran: merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas obyek-obyek yang memiliki kemiripan. Klaster adalah kumpulan record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan record dalam klaster yang lain. Berbeda dengan klasifikasi, pada pengklasteran tidak ada variabel target. Pengklasteran tidak melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target, akan tetapi, algoritma pengklasteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal.
6. Asosiasi: Tugas asosiasi dalam data mining adalah untuk menemukan atribut yang muncul dalam satu waktu. Salah satu implementasi dari asosiasi adalah market basket analysis sebagaimana yang akan dibahas dalam penelitian ini.

2.1.7 Association Rules

Menurut (Agusta, 2015) Association Rule adalah suatu metode data mining yang bertujuan untuk mencari sekumpulan items yang sering muncul bersamaan. Umumnya association rule ini dianalogikan dengan keranjang belanjaan. Dari keranjang belanjaan para pengunjung supermarket akan dapat diketahui, barang apa saja yang sering dibeli bersamaan dan barang mana saja yang tidak. Menurut (Pracoyo, 2016) Analisis asosiasi atau association rule mining adalah teknik data mining untuk menemukan aturan suatu kombinasi item. Salah satu tahap analisis asosiasi yang menarik perhatian banyak

peneliti untuk menghasilkan algoritma yang efisien adalah analisis pola frekuensi tinggi (frequent pattern mining). penting tidaknya suatu asosiasi dapat diketahui dengan dua tolak ukur, yaitu : support dan confidence. Support (nilai penunjang) adalah persentase kombinasi item tersebut dalam database, sedangkan confidence (nilai kepastian) adalah kuatnya hubungan antar item dalam aturan asosiasi.

2.1.8 Algoritma C4.5

Algoritma 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 sendiri merupakan pengembangan dari algoritma ID3, dimana pengembangan dilakukan dalam hal, bisa mengatasi missing data, bisa mengatasi data kontinu dan pruning. Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk setiap cabang sampai semua
5. kasus pada cabang memiliki kelas yang sama.
6. Untuk memilih atribut akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti yang tertera dalam persamaan 1 berikut :

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_i \text{Entropy}(S_i) \text{ Dimana :}$$

S : himpunan kasus

A : atribut

N : jumlah partisi atribut A

$|S_i|$: jumlah kasus pada partisi ke- i

$|S|$: jumlah kasus dalam S

2.1.9 Tanagra

Tanagra adalah software Data Mining bebas untuk tujuan akademik dan penelitian ini mengusulkan beberapa metode data mining dari analisis eksplorasi data, pembelajaran statistik, pembelajaran mesin dan daerah database. Tanagra adalah "proyek *open source*" karena setiap peneliti dapat mengakses ke kode sumber, dan menambahkan algoritma sendiri, sejauh dia setuju dan sesuai dengan lisensi distribusi perangkat lunak.

Tujuan utama dari proyek Tanagra adalah memberikan peneliti dan mahasiswa yang mudah untuk menggunakan perangkat lunak data mining, sesuai dengan norma-norma yang hadir dari pengembangan perangkat lunak dalam domain ini (terutama dalam desain GUI dan cara untuk menggunakannya), dan memungkinkan untuk menganalisis baik data yang nyata atau sintetis.

Tujuan kedua Tanagra adalah untuk mengusulkan kepada peneliti arsitektur yang memungkinkan mereka untuk dengan mudah menambahkan metode penambangan data mereka sendiri, untuk membandingkan kinerja mereka. Tanagra bertindak lebih sebagai platform eksperimental untuk membiarkan mereka pergi ke penting dari pekerjaan mereka, pengeluaran mereka untuk berurusan dengan bagian menyenangkan dalam *programmation* semacam ini alat pengelolaan data (Julsam, 2009).

Tujuan ketiga dan terakhir, arah pengembang pemula, terdiri dalam menyebarkan metodologi yang mungkin untuk membangun perangkat lunak semacam ini. Mereka harus mengambil keuntungan dari akses gratis ke kode sumber, untuk melihat bagaimana perangkat lunak semacam ini dibangun, masalah untuk menghindari, langkah-langkah utama dari proyek ini, dan alat-alat dan perpustakaan kode yang digunakan untuk. Dengan cara ini,

Tanagra dapat dianggap sebagai alat pedagogis untuk belajar teknik pemrograman. (Mempersiapkan data yang menjadi sumber untuk *data mining* termasuk data “*cleaning*” untuk dipelajari polanya). Setelah menentukan definisi masalah, langkah berikutnya adalah mencari data yang mendukung definisi masalah. Menentukan porsi data yang digunakan men-training data mining berdasarkan algoritma data mining yang telah dibuat. Setelah persiapan data selesai dilakukan, langkah berikutnya adalah memberikan sebagian data kedalam algoritma data mining.

BAB III

METODE PENELITIAN

Pengumpulan data merupakan suatu hal yang sangat penting. Pengumpulan data ini merupakan suatu kegiatan yang terkait dengan keadaan real di lapangan (empiris). Hasil pengumpulan data ini digunakan untuk penyimpulan pada kegiatan penelitian, termasuk di dalamnya mentoring dan evaluasi (monev).

Pengumpulan data merupakan prosedur yang sistematis dan terstandarkan untuk memperoleh data dalam suatu penelitian. Dalam pemilihan metode penelitian, selalu tergantung pada permasalahan penelitian yang akan dipecahkan dan jenis data yang akan dikumpulkan. Masalah penelitian menentukan jenis data yang diperlukan, dan jenis data ini memandu pemilihan metode atau cara pengumpulan data. Pada monev metode pengumpulan data yang digunakan adalah pengumpulan data, pengolahan awal data, dan model.

3.1 Pengumpulan data

Pada pengumpulan data dijelaskan tentang bagaimana dan dari mana data dalam penelitian ini didapatkan, ada dua tipe dalam pengumpulan data, yaitu pengumpulan data primer dan pengumpulan data sekunder. Data primer adalah data yang dikumpulkan pertama kali untuk melihat apa yang sesungguhnya terjadi. Data sekunder adalah data yang sebelumnya pernah dibuat oleh seseorang baik di terbitkan atau tidak (Vercelis, 2009). Dalam pengumpulan data primer dalam penelitian ini menggunakan metode observasi dan interview, dengan menggunakan data-data yang berhubungan dengan mentoring semester 1 dan semester 2. Data monev selama ini tidak tersusun dengan baik, sehingga data yang setiap minggu tersebut hanya berfungsi sebagai perbandingan monev dan tidak dapat dimanfaatkan untuk perkembangan kampus. Oleh sebab itu kampus tersebut memerlukan sistem untuk mengolah data yang dapat menghasilkan data monev untuk

perkembangan kampus, sehingga data monev tersebut dapat menjadi acuan untuk mengembangkan dan meningkatkan perkembangan kampus UTS.

3.2 Pengolahan Awal Data

Jumlah data awal yang diperoleh dari pengumpulan data dalam 1 tahun yaitu bulan Agustus 2020 sd bulan juni 2021 yaitu berupa data kehadiran, namun tidak semua data dapat digunakan dan tidak semua atribut digunakan karena harus melalui beberapa tahap pengolahan awal data (*preparation data*). Untuk mendapatkan data yang berkualitas, beberapa teknik yang dilakukan yaitu (Kothari, 2004): data validation, data integration and transformation dan data *size reduction and discritization*. Sehingga perlu dilakukan perbandingan terhadap setiap kehadiran yang akan dilakukan analisis.

NO	NAMA	PRODI	POIN		POINT		
1	LENA LESTARI	EKBANG	52,5	15	13	80,5	LULUS
2	LINDA AMANDA SAPIRA	EKBANG	0	0	0	0	TIDAK LULUS
3	NANDA PUTRI	EKBANG	52,5	15	13	80,5	LULUS
4	NENDI ROHAYU	EKBANG	17,5	12,75	12	42,25	TIDAK LULUS
5	NUR HASANA	EKBANG	61,25	15	15	91,25	LULUS
6	ANIRAWATI	EKBANG	70	15	10	95	LULUS
7	SULIS YUNITA ULANI	EKBANG	0	12	9	21	TIDAK LULUS
8	DINDA MERLIANA	EKBANG	61,25	12	10	83,25	LULUS
9	EMA SAHRITA	EKBANG	70	12,75	9	91,75	LULUS
10	GINA HARLINA	EKBANG	61,25	12,75	9	83	LULUS
11	ISNAINI	EKBANG	26,25	10,5	8	44,75	TIDAK LULUS
12	CINDYA	EKBANG	70	12,75	9	91,75	LULUS
13	WITRI CAHYA ARINI	EKBANG	52,5	0	0	52,5	TIDAK LULUS
14	AINUN ZARIYAH PUA	AKUNTANSI	70	13,8	12	95,8	LULUS
15	ALMA SHARIKA	AKUNTANSI	70	14,25	13	97,25	LULUS
16	AYU DWI CAHYANI	AKUNTANSI	70	14,7	12	96,7	LULUS
17	DIAN NATULHIKMAH	AKUNTANSI	61,25	13,65	10	84,9	LULUS
18	DIAN SUNARSIH	AKUNTANSI	0	0	0	0	TIDAK LULUS
28	DWI YANA RAVIKA	MANAJEMEN	52,5	13,5	13	79	LULUS
20	EFA ROSIFA	MANAJEMEN	17,5	7,5	13	38	TIDAK LULUS
21	EKA RAHMATULYANI	MANAJEMEN	43,75	12,75	13	69,5	TIDAK LULUS
22	EKA SAPUTRIANI	MANAJEMEN	70	9	13	92	LULUS
23	ELI ERMAWATI	MANAJEMEN	70	12,75	13	95,75	LULUS
24	ANITA SAFITRI	ILKOM	61,25	12	10	83,25	LULUS
25	ASTRI FITRIANTI	ILKOM	43,75	12,45	10	66,2	TIDAK LULUS
26	DESTIAN CAHYANI	ILKOM	70	14,25	11	95,25	LULUS
27	BUNGA AYU NINGTYAS	THP	70	13,5	10	93,5	LULUS

58	NANDA PUTRI	EKBANG	17,5	12,75	12	42,25	TIDAK LULUS
59	NENDI ROHAYU	EKBANG	17,5	12,75	12	42,25	TIDAK LULUS
60	NUR HASANA	EKBANG	61,25	15	15	91,25	LULUS
61	ANIRAWATI	EKBANG	70	15	10	95	LULUS
62	SULIS YUNITA ULANI	EKBANG	0	12	9	21	TIDAK LULUS
63	DINDA MERLIANA	EKBANG	61,25	12	10	83,25	LULUS
64	EMA SAHRITA	EKBANG	70	12,75	9	91,75	LULUS
65	GINA HARLINA	EKBANG	61,25	12,75	9	83	LULUS
66	ISNAINI	EKBANG	26,25	10,5	8	44,75	TIDAK LULUS
67	CINDYA	EKBANG	17,5	12,75	12	42,25	TIDAK LULUS
68	WITRI CAHYA ARINI	EKBANG	52,5	0	0	52,5	TIDAK LULUS
69	AINUN ZARIYAH PUA	AKUNTANSI	70	13,8	12	95,8	LULUS
70	ALMA SHARIKA	AKUNTANSI	70	14,25	13	97,25	LULUS
71	AYU DWI CAHYANI	AKUNTANSI	70	14,7	12	96,7	LULUS
72	DIAN NATULHIKMAH	AKUNTANSI	61,25	13,65	10	84,9	LULUS
73	DIAN SUNARSIH	AKUNTANSI	0	0	0	0	TIDAK LULUS
74	DWI YANA RAVIKA	MANAJEMEN	52,5	13,5	13	79	LULUS
75	EFA ROSIFA	MANAJEMEN	17,5	7,5	13	38	TIDAK LULUS
76	EKA RAHMATULYANI	MANAJEMEN	43,75	12,75	13	69,5	TIDAK LULUS
77	EKA SAPUTRIANI	MANAJEMEN	70	9	13	92	LULUS
78	ELI ERMAWATI	MANAJEMEN	70	12,75	13	95,75	LULUS
79	ANITA SAFITRI	ILKOM	61,25	12	10	83,25	LULUS
80	ASTRI FITRIANTI	ILKOM	43,75	12,45	10	66,2	TIDAK LULUS
81	DESTIAN CAHYANI	ILKOM	70	14,25	11	95,25	LULUS
82	BUNGA AYU NINGTYAS	THP	70	13,5	10	93,5	LULUS
83	DINDA SARI	THP	70	11,25	9	90,25	LULUS
84	RANI HUSDIANA	TIP	35	0,75	5	40,75	TIDAK LULUS
85	NILASARI	TIP	0	0	0	0	TIDAK LULUS
86	NUR BARIA	TIP	52,5	1,5	10	64	TIDAK LULUS

87	NUR SYAFITRI	TIP	70	0,75	5	75,75	LULUS
88	PATRIANA	TIP	70	0,75	5	75,75	LULUS
89	PUTRI LESTARI	TIP	70	0,15	1	71,15	LULUS
90	ALYA SAYYIDAH ZN	TEKNOBIOLOGI	0	0	0	0	TIDAK LULUS
91	ANDINI CINDY ROHAYU	TEKNOBIOLOGI	70	13,5	12	95,5	LULUS
92	SISKA PUTRI ADEYANTI	TEKNOBIOLOGI	70	13,8	10	93,8	LULUS
93	SRI HASDAYANTI	TEKNOBIOLOGI	70	13,5	10	93,5	LULUS
94	SUCI RAMDANI	TEKNOBIOLOGI	70	14,55	10	94,55	LULUS
95	WULAN WILA KANTARI	TEKNOBIOLOGI	70	13,95	12	95,95	LULUS
96	AMBAR WATI	TI	70	12,45	10	92,45	LULUS
97	ANDINI WULANDARI	TI	70	13,95	11	94,95	LULUS
98	ARMANITA ANING	TI	70	12	10	92	LULUS
99	ASTI OKTAVIANI	TI	70	12,75	9	91,75	LULUS
100	NELY HARYANINGSI	TI	70	10,5	11	91,5	LULUS
101	NOVITASARI	TI	70	14,25	13	97,25	LULUS
102	NURUL HIDAYAT	TI	70	13,5	9	92,5	LULUS
103	ALDA	TEKNIK SIPIL	70	13,95	12	95,95	LULUS
104	CICI PUTRI SAKINA	TEKNIK INDUSTRI	70	15	10	95	LULUS
105	FITRI HIDAYA	TEKNIK INDUSTRI	70	0	0	70	LULUS
106	MAWAR JULIANA	TEKNIK	70	9	9	88	LULUS
107	ZADDWI JIWA ILMIAH	TI	70	13,125	13	83,125	LULUS
108	WULANDARI	TI	70	12,825	8	82,825	LULUS
109	OKTA SUSANTI	TIP	70	13,05	10	83,05	LULUS
110	ADILAH JULIA SUMA	PSIKOLOGI	0	0	0	0	TIDAK LULUS

22xxxxxxxxxxxxxx

BAB IV

HASIL DAN PEMBAHASAN

4.1 Pengujian dan Validasi Hasil

4.1.1 Proses Mining

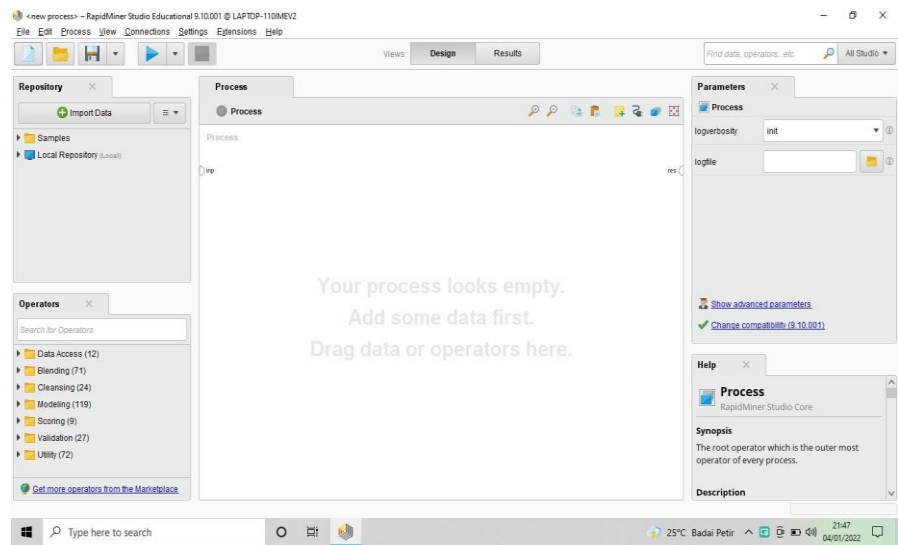
Pada tahap ini metode data mining untuk menemukan pengetahuan tersembunyi dan berharga dari data. Metode yang digunakan adalah klasifikasi dengan algoritma C4.5. Berikut penerapan algoritma C4.5 Memakai tool RapidMiner 9.10.001.

1. Langkah pertama adalah membuka aplikasi *RapidMiner*, maka yang akan muncul pertama kali di layar adalah tampilan loading seperti berikut:



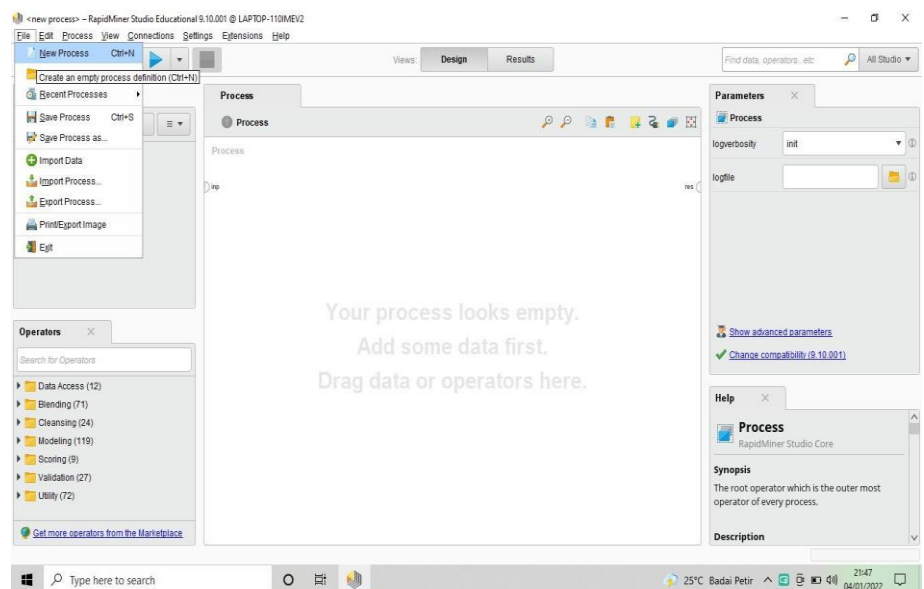
Gambar 1 tampilan awal RapidMiner

- Setelah loading selesai, selanjutnya akan muncul tampilan utama



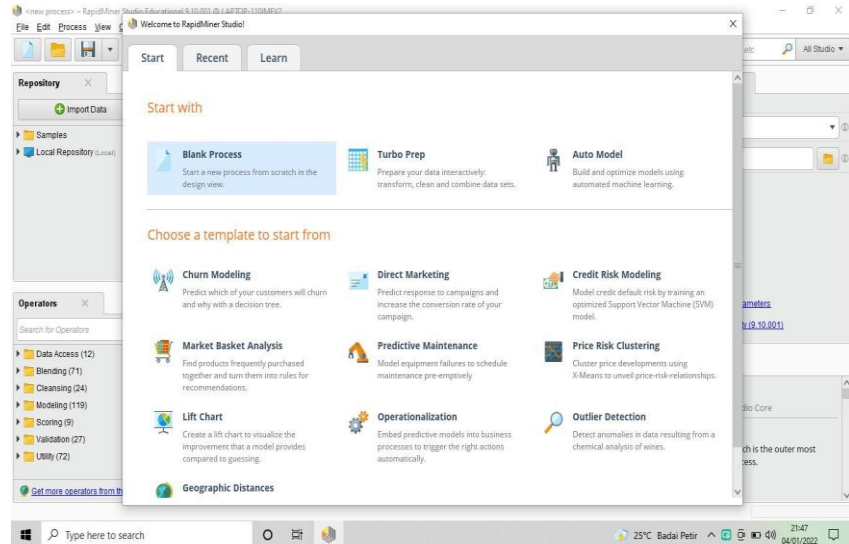
Gambar 2 tampilan menu proses RapidMiner

- Setelah muncul tampilan menu utama, tekan *New Process* untuk melanjutkan dalam proses pengolahan data.



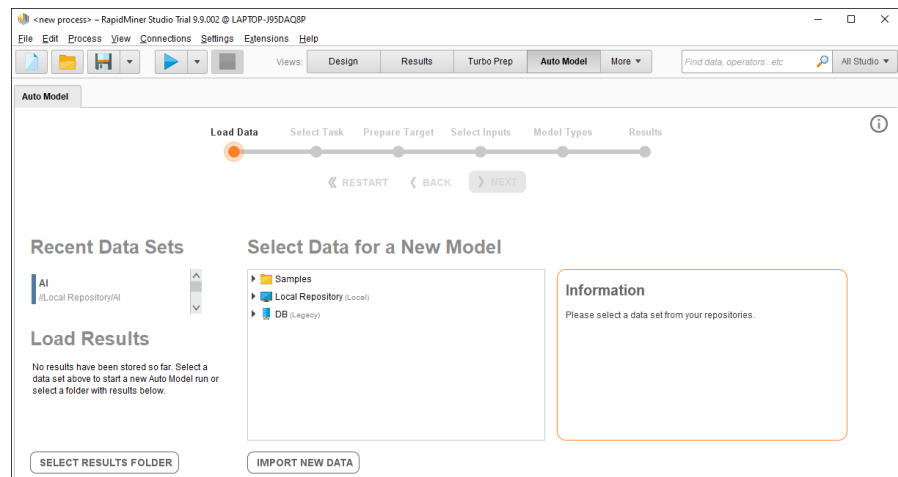
Gambar 3 tampilan new pros pada Rapidminer

4. Kemudian lanjutkan dan pilih *Blank Process*



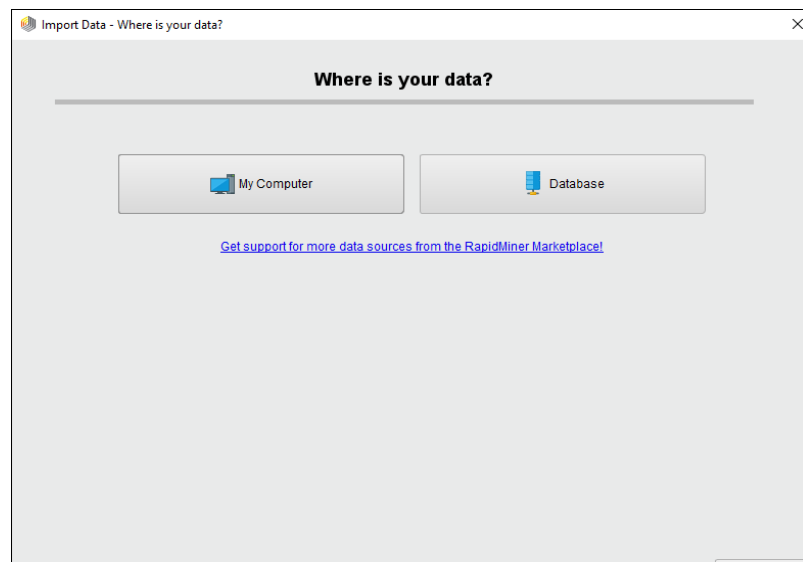
Gambar 4 Tampilan *Blank Process* pada RapidMiner

5. Kemudian lanjut untuk memasukkan data yang kita miliki untuk di proses, sebelum itu klik Auto Model dal lanjut klik lagi bagian informasi



Gambar 5 Tampilan Auto Model pada RapidMiner

6. Lanjut dan pilih My Comouter untuk melankutkan proses datanya.



Gambar 6 Tampilan informasi di My Computer pada RapidMiner

7. Maka akan muncul data yang telah kita masukkan untuk di proses.

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	NO <i>integer</i>	NAMA <i>polynomial</i>	PRODI <i>polynomial</i>	POIN KEHA... <i>polynomial</i>	POIN UJIAN <i>polynomial</i>	POINT TIL... <i>integer</i>
1	1	LENA LESTARI	EKBANG	52,5	15	13
2	2	LINDA AMANDA ...	EKBANG	0	0	0
3	3	NANDA PUTRI W...	EKBANG	35	13,5	11
4	4	NENDI ROHAYU ...	EKBANG	17,5	12,75	12
5	5	NUR HASANA	EKBANG	61,25	15	15
6	6	ANIRAWATI	EKBANG	70	15	10
7	7	SULIS YUNITA U...	EKBANG	0	12	9
8	8	DINDA MERLIANA	EKBANG	61,25	12	10
9	9	EMA SAHRITA	EKBANG	70	12,75	9
10	10	GINA HARLINA	EKBANG	61,25	12,75	9
11	11	ISNAINI	EKBANG	26,25	10,5	8
12	12	CINDYA RAMAYA	EKBANG	43,75	0	0

no problems.

Gambar 7 Tampilan data di Format Your Columns pada RapidMiner

8. Lanjutkan dan klik next untuk melanjutkan proses pada data

Import Data - Select the cells to import.

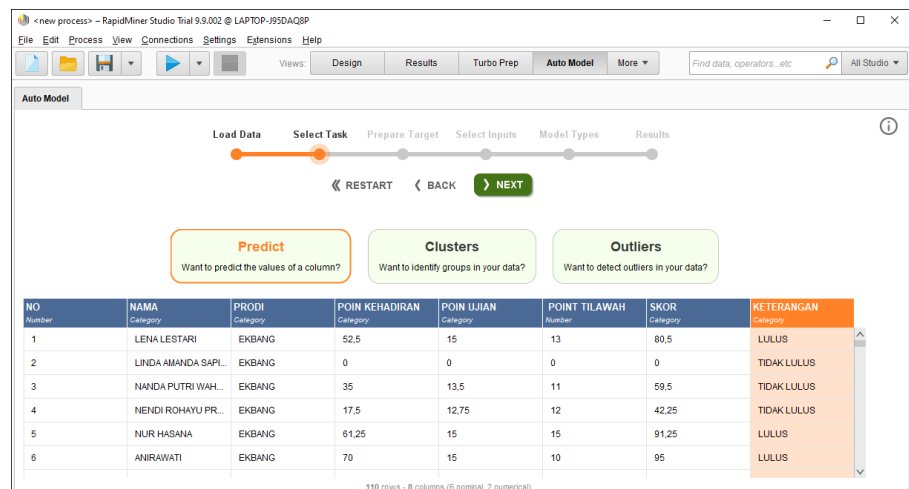
Select the cells to import.

Sheet: Sheet1 Cell range: A:H Select All ☒ Define header row: 1

	A	B	C	D	E	F	G	H
1	NO	NAMA	PRODI	POIN KEHA...	POIN UJIAN	POINT TIL...	SKOR	KETERANG...
2	1.000	LENA LEST...	EKBANG	52,5	15.000	13.000	80,5	LULUS
3	2.000	LINDA AMA...	EKBANG	0.000	0.000	0.000	0.000	TIDAK LUL...
4	3.000	NANDA PU...	EKBANG	35.000	13,5	11.000	59,5	TIDAK LUL...
5	4.000	NENDI RO...	EKBANG	17,5	12,75	12.000	42,25	TIDAK LUL...
6	5.000	NUR HASA...	EKBANG	61,25	15.000	15.000	91,25	LULUS
7	6.000	ANIRAWATI	EKBANG	70.000	15.000	10.000	95.000	LULUS
8	7.000	SULIS YUN...	EKBANG	0.000	12.000	9.000	21.000	TIDAK LUL...
9	8.000	DINDA ME...	EKBANG	61,25	12.000	10.000	83,25	LULUS
10	9.000	EMA SAHRI...	EKBANG	70.000	12,75	9.000	91,75	LULUS
11	10.000	GINA HARL...	EKBANG	61,25	12,75	9.000	83.000	LULUS
12	11.000	ISNAINI	EKBANG	26,25	10,5	8.000	44,75	TIDAK LUL...
13	12.000	CINDYA RA...	EKBANG	43,75	0.000	0.000	43,75	TIDAK LUL...
14	13.000	WITRI CAH...	EKBANG	52,5	0.000	0.000	52,5	TIDAK LUL...

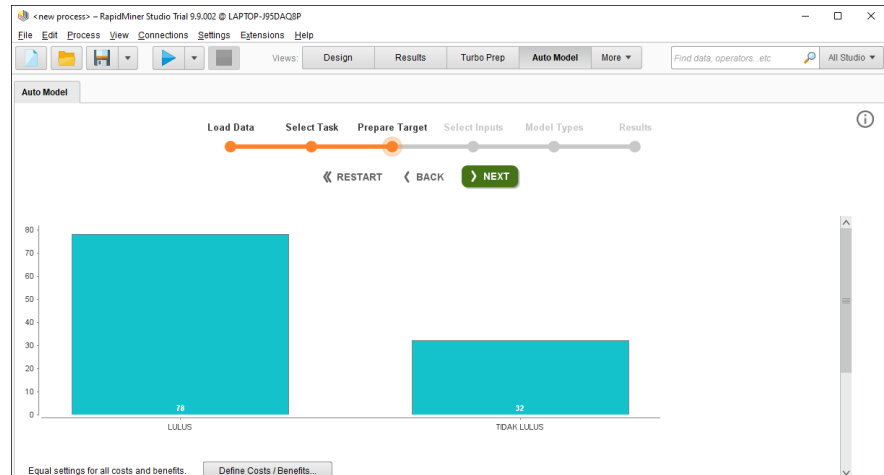
Gambar 8 Tampilan data di select the cell to import pada RapidMiner

9. Maka setelah proses pemasukan data selesai, maka akan kembali lagi pada bagian Auto Model dan selanjutnya masuk ke bagian select Task dan klik Predict untuk memprediksi data dengan dilanjutkan dengan klik bagian data keterangan Lulus dan Tidak Lulus untuk memprediksi data kelulusan mentoringnya dan lanjut klik next.



Gambar 9 Tampilan predict data di keterangan pada RapidMiner

10. Setelah itu akan muncul hasil prediksinya di bagian Prapare Target dengan presentasi Lulus 78 % dan Tidak Lulus 32% . Lanjut dengan klik next untuk melanjutkan proses data di bagian Select Inputs



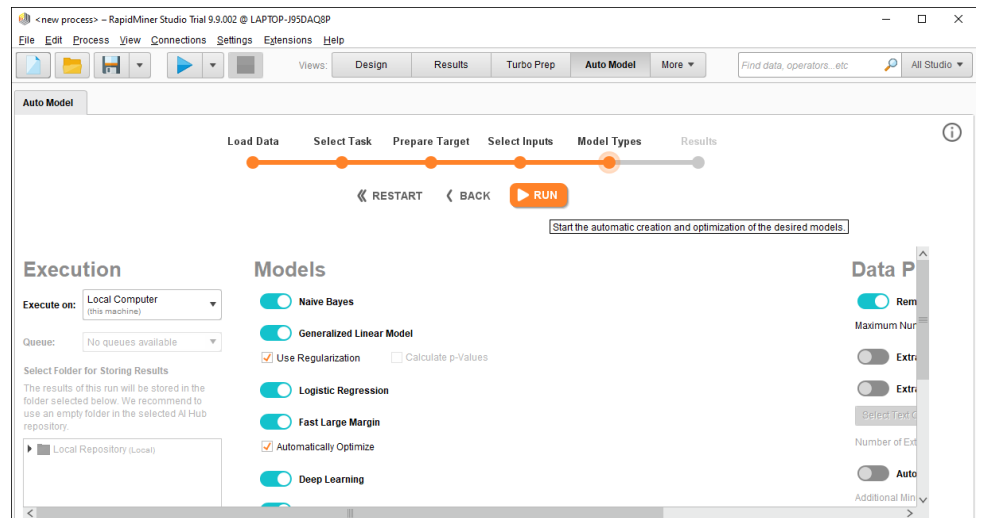
Gambar 10 Tampilan pada Prepare Target pada RapidMiner

11. Maka akan muncul tampilan data pada bagian Select Inputs dengan bagian-bagian yang sudah ada hasil persen (%) dari setiap atribut, kemudian lanjut lagi untuk proses data di bagian Model Types dengan klik next untuk melanjutkan

Selected	Status	Quality	Name	Correlation	ID-ness	Stability	Missing	Text-ness
<input type="checkbox"/>	Red	Bar chart	NO	2.77%	100.00%	0.91%	0.00%	0.00%
<input type="checkbox"/>	Red	Bar chart	NAMA	1.71%	74.55%	2.73%	0.00%	55.64%
<input checked="" type="checkbox"/>	Green	Bar chart	PRODI	8.34%	11.82%	23.64%	0.00%	9.35%

Gambar 11 Tampilan Select Inputs pada RapidMiner

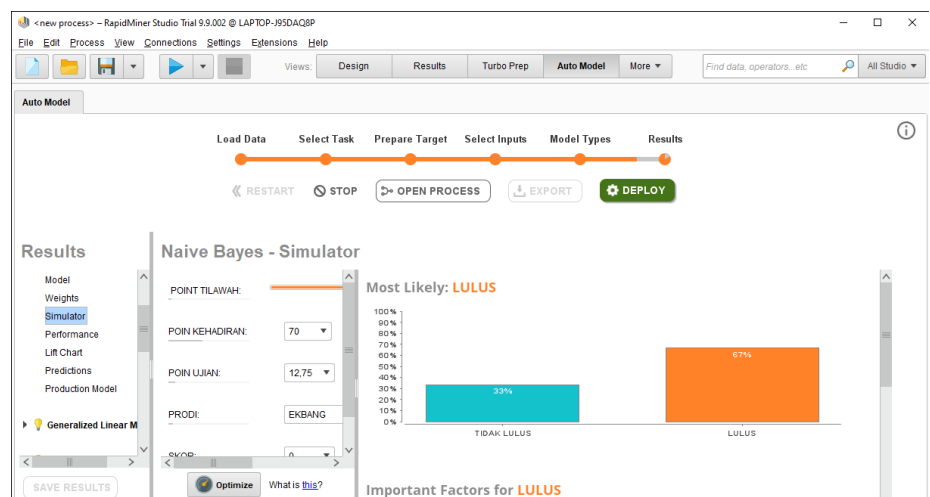
12. Selanjutnya hasil dari Model Types pada proses data tersebut dan dilanjutkan dengan klik Run untuk mengetahui hasil akhir dari prediksi data kelulusan mentoring tersebut



Gambar 12 Tampilan Model Types pada RapidMiner

4.1.2 Hasil Proses Validation

1. Hasil dari Run tersebut akan muncul dan bagian ini ada beberapa hasil yang ditampilkan dengan menggunakan RapidMiner ditentukan dengan Algoritma C4.5 yang terdapat pada data kelulusan mentoring



Gambar 13 Tampilan simulator di bagian lulus pada RapidMiner

2. Dari hasil data tersebut dapat dilihat bahwa kelulusan data mentoring dengan presentase dengan hasil prediksi Tidak Lulus 33% dan Lulus 67%.

The screenshot shows the 'Naive Bayes - Predictions' results table in RapidMiner Studio. The table has columns for Row No., KETERANGAN, prediction(K..., confidence..., cost, PRODI, POIN KEHAD..., POIN UJIAN, and SKOR. The data is as follows:

Row No.	KETERANGAN	prediction(K...	confidence...	confidence...	cost	PRODI	POIN KEHAD...	POIN UJIAN	SKOR
1	LULUS	LULUS	0.500	0.500	0.000	EXBANG	52.5	15	MIS
2	LULUS	LULUS	0.167	0.833	0.667	AKUNTANSI	78	14.25	97.5
3	LULUS	LULUS	0.167	0.833	0.667	AKUNTANSI	61.25	13.65	84.75
4	TIDAK LULUS	TIDAK LULUS	1.000	0.000	1.000	AKUNTANSI	9	0	0
5	TIDAK LULUS	TIDAK LULUS	0.842	0.158	0.883	MANAJEMEN	43.75	12.75	69.75
6	LULUS	LULUS	0.167	0.833	0.667	MANAJEMEN	70	9	92
7	LULUS	LULUS	0.333	0.667	0.333	MANAJEMEN	70	12.75	95.75
8	LULUS	LULUS	0.500	0.500	0.000	ILKOM	61.25	12	83.25

Gambar 14 Tampilan Hasil pada Predictions pada RepidMiner

3. Hasil penentuan confidention dari data dengan rules klasifikasi yang telah dibuat dari data.

The screenshot shows an Excel spreadsheet with the following data:

KETERANGAN	POIN KEHAD...	POIN UJIAN	PRODI	SKOR
1 LULUS	52.5	15	EXBANG	60.5
2 LULUS	78	14.25	AKUNTANSI	97.5
3 LULUS	61.25	13.65	AKUNTANSI	84.75
4 TIDAK LULUS	9	0	AKUNTANSI	0
5 TIDAK LULUS	43.75	12.75	MANAJEMEN	69.75
6 LULUS	70	9	MANAJEMEN	92
7 LULUS	70	12.75	MANAJEMEN	95.75
8 LULUS	61.25	12	ILKOM	83.25

Gambar 15 Tampilan hasil akhir dipindahkan ke Excel.

BAB V

KESIMPULAN

5.1 Kesimpulan

Dalam proses pembuatan data klasifikasi dapat menggunakan algoritma C4.5 pada RapidMiner. Hasil dari prediksi databaset tersebut yang terdapat pada data kelulusan mentoring Dari hasil data tersebut dapat dilihat bahwa kelulusan data mentoringdari berbagai prodi dengan presentase dengan hasil prediksi Tidak Lulus 33% dan Lulus 67%. Dari hasil predeksi tersebut menggunakan 3 bagian dari RapidMiner yaitu simulator, Predictions dan yang terakhir hasil akhir yang dipindahkan ke bentuk Excel.

5.2 SARAN

Agar penelitian ini terus berkembang, maka hasil dari penelitian ini dari rule yang terbentuk dapat dikembangkan menjadi pengetahuan untuk sistem pendukunng keputusan.

DAFTAR PUSTAKA

Link : <https://rapidminer.com>, <https://www.dicoding.com> ,

Diakses 15 Januari 2022, 20:00