



## KLASIFIKASI DATA MINING UNTUK MEMPREDIKSI KELULUSAN MAHASISWA MENGUNAKAN METODE NAIVE BAYES

**Jefri<sup>a\*</sup>, Zaehol Fatah<sup>b</sup>,**

<sup>a</sup> Fakultas Sains & Teknologi / Sistem Informasi, [renorisma903@gmail.com](mailto:renorisma903@gmail.com), Universitas Ibrahimy,  
Situbondo Jawa Timur

<sup>b</sup> Fakultas Sains & Teknologi / Sistem Informasi, [zaeholfatah@gmail.com](mailto:zaeholfatah@gmail.com), Universitas Ibrahimy,  
Situbondo Jawa Timur

\* korespondensi

### ABSTRACT

*Data mining helps provide precise and careful decisions. Student graduation on time is one of the assessment points in the higher education accreditation process. However, student graduation cannot always be detected quickly, which can reduce the assessment of a university in the accreditation process. This problem arises to find out whether students will be able to graduate on time or not Classification method for predicting student graduates using the Naïve Bayes algorithm. Whether a student graduates on time or not, it is hoped that the results will provide information and input for the university in making future policies. From the results of this test, it was found that by applying the Naïve Bayes algorithm the system can predict student graduation in a timely manner. After comparing several literatures, it can be concluded that this method can be used for this prediction with an accuracy rate of 90%. This literature review is important as a supporting factor for research.*

**Keywords:** *Data Mining, Graduation Prediction, Naive Bayes, Prediction, Student Graduation*

### Abstrak

Data mining membantu memberikan keputusan yang tepat dan cermat. Kelulusan mahasiswa tepat waktu merupakan salah satu point penilaian dalam proses akreditasi perguruan tinggi. Namun kelulusan mahasiswa tidak selalu dapat dideteksi secara cepat sehingga dapat mengurangi penilaian suatu perguruan tinggi dalam proses akreditasi. Permasalahan inilah yang muncul untuk mengetahui mahasiswa nantinya bisa lulus tepat waktu atau tidak. metode klasifikasi untuk prediksi lulusan mahasiswa menggunakan algoritma *Naive Bayes*. kelulusan mahasiswa tepat waktu atau tidak, yang diharapkan hasilnya dapat memberikan informasi dan masukan bagi pihak perguruan tinggi dalam membuat kebijakan kedepannya. Dari hasil pengujian ini didapatkan hasil bahwa dengan menerapkan algoritma *Naive Bayes* sistem dapat memprediksi kelulusan mahasiswa dengan tepat waktu, Setelah membandingkan beberapa literatur, dapat disimpulkan bahwa metode ini dapat digunakan untuk prediksi tersebut dengan tingkat keakuratan 90%. Kajian literatur ini penting sebagai faktor pendukung bagi penelitian.

**Kata Kunci:** *Data Mining, Prediksi Kelulusan, Naive Bayes, Prediksi, Kelulusan Mahasiswa*

### 1. PENDAHULUAN

Perguruan tinggi merupakan lembaga pendidikan yang menyelenggarakan layanan pembelajaran tingkat tinggi yang merupakan tahap akhir pilihan dalam pendidikan formal. Perguruan tinggi tersebut umumnya berbentuk universitas, akademi, institut atau sekolah menengah atas. Jenis pendidikan tinggi meliputi vokasi, akademik, dan profesional. Berdasarkan jenjangnya, pendidikan tinggi menyelenggarakan program diploma, sarjana, magister, spesialis, dan doktor[1].

Perguruan tinggi diharapkan dapat memberikan pendidikan yang berkualitas bagi mahasiswanya sehingga menghasilkan mahasiswa yang memiliki kompetensi di bidangnya. Di Indonesia, mutu suatu institusi

pendidikan tinggi ditentukan oleh nilai akreditasi yang dikeluarkan oleh Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT). Ada banyak aspek yang bisa dijadikan tolak ukur mutu pendidikan tinggi. Salah satunya terlihat dari banyaknya mahasiswa yang dapat menyelesaikan studinya dalam waktu yang ditentukan atau dengan kata lain mahasiswa lulus tepat waktu. Semakin banyak mahasiswa yang lulus tepat waktu maka semakin baik pula kinerja perguruan tinggi tersebut, sehingga tingkat kelulusan mahasiswa tepat waktu menjadi salah satu kriteria penilaian. akreditasi suatu universitas.

Banyak faktor yang dapat mendukung peningkatan jumlah mahasiswa yang lulus tepat waktu, antara lain:

- a. Manajemen Akademik yang Efektif
- b. Kurikulum yang Relevan
- c. Dukungan Bimbingan Akademik

Dengan menerapkan manajemen akademik yang efektif, kurikulum yang relevan, dan dukungan bimbingan akademik yang kuat, universitas dapat meningkatkan jumlah mahasiswanya yang lulus tepat waktu secara signifikan. Hal ini tidak hanya menguntungkan peserta didik, namun juga meningkatkan reputasi dan kredibilitas lembaga pendidikan tersebut[2]. Data kelulusan dapat dianalisis dan diolah secara detail dengan menggunakan teknik data mining.

Penambangan data adalah proses mengekstraksi pengetahuan atau informasi berharga dari kumpulan data yang besar dan kompleks. Tujuan utama dari data mining adalah untuk mengidentifikasi pola, hubungan, atau informasi yang mungkin tidak terlihat secara langsung pada data, sehingga dapat memberikan informasi yang lebih dalam dan valid[3]. Data mining dalam arti lain adalah istilah yang digunakan untuk menggambarkan penemuan pengetahuan dalam database, yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstraksi dan mengidentifikasi informasi dan pengetahuan berguna yang dikumpulkan dari berbagai database besar[4].

Tingkat kelulusan siswa diprediksi menggunakan suatu sistem. Namun beberapa perguruan tinggi belum memiliki sistem untuk memprediksi keterlambatan kelulusan mahasiswa sehingga perguruan tinggi tidak dapat mencegah hal tersebut.

Bagaimana algoritma Naïve Bayes diterapkan untuk memprediksi kelulusan siswa tepat waktu? Algoritma Nave Bayes adalah salah satu algoritma pembelajaran induktif yang paling efektif dan efisien untuk pembelajaran mesin dan penambangan data. Kinerja Nave Bayes kompetitif dalam proses klasifikasi meskipun menggunakan asumsi independensi atribut (tidak ada hubungan antar atribut). Naive Bayes digunakan untuk teknik klasifikasi data dengan menggunakan metode probabilitas dan statistik yang memprediksi peluang masa depan berdasarkan pengalaman sebelumnya, sehingga dikenal dengan Teorema Bayes. Teorema ini digabungkan dengan Naive dimana diasumsikan bahwa kondisi antara atribut yang satu dengan atribut yang lain saling independen. Klasifikasi Naive Bayes berasumsi bahwa ada atau tidaknya ciri-ciri tertentu pada suatu kelas tidak ada hubungannya dengan ciri-ciri kelas lainnya[5].

Dari beberapa penjelasan mengenai prediksi lulusan terdapat beberapa perbedaan antara lain menggunakan algoritma Naive Bayes dengan metode klasifikasi menggunakan atribut nama mahasiswa, NPM, jalur masuk, IPK mahasiswa dan jenis kelamin. Dengan data atribut dan kriteria tersebut, algoritma dapat memprediksi tingkat kelulusan siswa. Sehingga dengan adanya sistem prediksi kelulusan mahasiswa ini dapat membantu untuk mengetahui apakah mahasiswa lulus tepat waktu atau tidak, dengan harapan hasilnya dapat memberikan informasi dan masukan bagi perguruan tinggi dalam mengambil kebijakan kedepannya.

Berdasarkan permasalahan yang ada, maka perlu adanya suatu sistem untuk memprediksi tingkat kelulusan siswa berdasarkan variabel-variabel yang ada. Selain itu juga diperlukan algoritma yang sesuai agar dapat menghasilkan nilai akurasi yang baik. Dengan sistem yang dibuat, diharapkan perguruan tinggi dapat mengambil kebijakan agar mahasiswanya dapat lulus tepat waktu.

## 2. METODOLOGI PENELITIAN

Penelitian ini dimulai dengan perumusan masalah dan studi literatur. Diperlukan sejumlah data yang diperoleh melalui metode observasi dan dokumentasi untuk dapat memenuhi permasalahannya. Berikutnya pengumpulan data mahasiswa yang dibutuhkan, dengan perolehan sejumlah 146 data set mahasiswa angkatan 2018 yang sudah selesai studinya dalam 15 atribut. Kemudian tahapan kedua

dilaksanakan pra-pemrosesan data untuk mendapatkan data yang baik sebelum data diolah menggunakan naïve bayes. Jika pra-pemrosesan data sudah selesai dilakukan, akan digunakan sejumlah 146 data mahasiswa dalam 15 atribut untuk melangsungkan proses mining. Ketiga, proses mining dengan menerapkan algoritma naïve bayes Tahap keempat eksperimen dan uji model. Tahap kelima evaluasi dan validasi hasil pengujian. hasil nilai akurasi yang didapat dari penerapan algoritma Naive Bayes. Untuk menghitung akurasi dari pola yang didapat, pengujian metode akurasi dalam konsep data mining. Perhitungan akurasi dapat dibantu dengan penerapan rumus perhitungan Naive Bayes . Adapun hasil penerapan rumus Naive Bayes menciptakan akurasi yang sangat baik berdasarkan perhitungan penerapan model naïve bayes[6].

## **2.1 Pengumpulan Data**

Adapun untuk teknik yang digunakan untuk mengumpulkan data penelitiannya terdiri dari :

### **2.1.1. Metode observasi**

Upaya dalam memperoleh atau mengumpulkan data secara langsung dengan mengamati di lapangan, untuk kemudian akan berguna dalam perhitungan ketika memprediksikan kinerja sistem yang berjalana pada perusahaan[7].

### **2.1.2. Metode kepustakaan**

Mengumpulkan data dengan cara mempelajari buku, jurnal ilmiah, serta informasi dari internet yang mendukung penelitian[8].

## **2.2 Analisis Data**

Data mahasiswa pada penelitian ini adalah data dari kaggle berupa jenis kelamin, status mahasiswa, status menikah, umur, Indeks Prestasi Semester (IPS) dari semester 1 sampai dengan semester 8, dan Indeks Prestasi Kumulatif (IPK). Dengan kelas target klasifikasinya yaitu status kelulusan yang mencakup apakah tepat atau terlambat[9].

## **2.3 Algoritma Naïve Bayes**

Naive Bayes adalah bentuk paling sederhana dari pengklasifikasi jaringan Bayesian. Dalam naïf Bayes, setiap node fitur memiliki simpul kelas sebagai induknya, tetapi tidak memiliki orangtua dari node fitur lainnya. Namun, sejumlah besar pekerjaan dalam pembelajaran terbimbing telah menunjukkan bahwa pengelompokan naïve bayes yang begitu sederhana dapat bersaing dengan pengklasifikasi canggih seperti C4.5 dan masih merupakan salah satu dari 10 algoritma penambahan data teratas di dunia[10]

## **2.4 Rapidminer**

Rapidminer adalah suatu perangkat lunak untuk Sains data yang dibentuk oleh Rapidminer, Inc yang menyajikan terintegrasi data, text mining, machine learning, dan analisis prediktif. Serta membantu langkah dalam suatu proses pembelajaran mesin seperti hasil visualisasi, optimasi, persiapan data dan validasi model. Rapidminer membawa kecerdasan buatan ke perusahaan melalui platform sains data yang terbuka dan dapat diperluas. Dibangun untuk tim analitik, Rapidminer menyatukan seluruh siklus ilmu Sains data dari persiapan data hingga machine learning hingga penyebaran prediktif (Hofmann & Klinkenberg, 2016)[9].

Secara sederhana Rapidminer adalah sebuah aplikasi yang digunakan untuk mengolah sebuah data dengan berbagai teknik dan metode dalam data mining, sehingga dapat menjadi informasi yang berguna. Rapidminer merupakan salah satu perangkat lunak yang bersifat open-source yang berguna untuk pengolahan Suatu data mining. Rapidminer mengolah dengan cara mengekstrak pola dari dataset dan mengkombinasikannya dengan metode statistika, kecerdasan buatan, dan Database. Tujuan dari penggunaan Rapidminer adalah untuk mendapatkan informasi bermutu tinggi dari teks yang diolah. Rapidminer digunakan pada berbagai penelitian untuk pengujian data yang di teliti[10].

## **3. HASIL DAN PEMBAHASAN**

Pada bagian ini berisi hasil dan pembahasan dari topik penelitian, yang bisa di buat terlebih dahulu metodologi penelitian. Bagian ini juga merepresentasikan penjelasan yang berupa penjelasan, gambar, tabel dan lainnya. Banyaknya kata pada bagian ini berkisar.

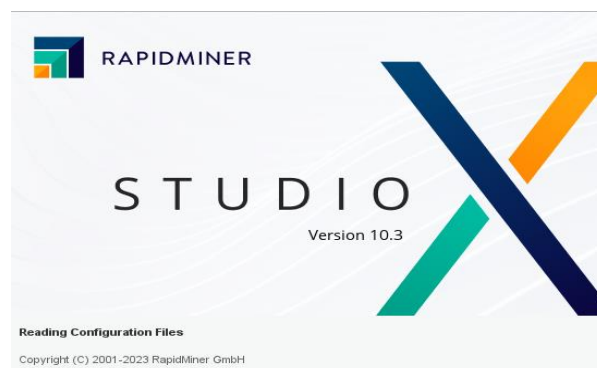
**Pengumpulan Data**

Proses pengumpulan data dilakukan dengan mengambil dari Kaggle. Data yang digunakan berbentuk file Microsoft Excel dengan format .xlsx. Data tersebut berasal dari mahasiswa angkatan masuk tahun 2018 yang telah menyelesaikan studinya, dengan total 146 data dan 14 atribut. Adapun contoh data yang digunakan dapat dilihat dalam gambar berikut. Data ini diambil dengan tujuan untuk dianalisis lebih lanjut dalam proyek penelitian, khususnya untuk keperluan klasifikasi Naive Bayes menggunakan RapidMiner. Adapun contoh data yang digunakan terdapat dalam gambar.

NAMA	JENIS KELAMIN	STATUS MAHASISWA	UMUR	STATUS NIKAH	IPS 1	IPS 2	IPS 3	IPS 4	IPS 5	IPS 6	IPS 7	IPS 8	IPK	STATUS KELULUSAN	
UNAMA	LAKI - LAKI	MAHASISWA	24	BELUM MENIKAH	3.17	2.7	3.23	2.41	3	2.47	1.75	0	2.75	TEPAT	
LEILA TRIYANA PRATIWI	PEREMPUAN	MAHASISWA	26	BELUM MENIKAH	3.6	3.5	3.42	2.85	3.31	2.95	2.18			3.39	TEPAT
VERIS SOFYANA PRAYOGA	LAKI - LAKI	MAHASISWA	29	BELUM MENIKAH	2.67	2.66	2.93	3.14	2.92	2.64	2.88	0.5	2.81	TEPAT	
ADITYA AKBAR NUGRAHA	LAKI - LAKI	MAHASISWA	27	BELUM MENIKAH	2.48	2.86	2.09	2.55	2.55	2.43	2.55	2.17	2.82	TEPAT	
ERNA EKA RIYANTI	PEREMPUAN	MAHASISWA	25	BELUM MENIKAH	3.19	3.08	3.31	2.83	3.36	2.73	3.06	0	3.09	TEPAT	
FARID DWI NORYANTO	LAKI - LAKI	MAHASISWA	24	BELUM MENIKAH	3.1	2.98	3.17	3.25	3.41	3.08	3.43	3	3.23	TEPAT	
DAFIK HADI WINOTO	LAKI - LAKI	MAHASISWA	24	BELUM MENIKAH	2.98	2.68	2.23	2.86	2.25	2.64	1.52	2.1	2.54	TEPAT	
WAHYU FITRIYANTO	LAKI - LAKI	MAHASISWA	24	BELUM MENIKAH	3.45	3.15	3.54	3.78	3.42	3.88	2.5	4	3.56	TEPAT	
IMAM SURYO SUSANTO	LAKI - LAKI	MAHASISWA	26	BELUM MENIKAH	3.31	3.02	3.48	3.7	3.19	3.21	2.58	4	3.4	TEPAT	
AHMAD SUTOPD	LAKI - LAKI	MAHASISWA	24	BELUM MENIKAH	2.62	2.73	2.11	3.36	2.79	3.17	2.45	0	2.97	TEPAT	
HERI SULFAT	LAKI - LAKI	MAHASISWA	24	BELUM MENIKAH	3.24	3.06	2.79	2.73	3.02	2.59	3.08	0	3.08	TEPAT	
MURYA AMIEN NUR PRABOWO	LAKI - LAKI	MAHASISWA	26	BELUM MENIKAH	2.83	3.02	3.02	3.63	3.21	3.09	2.57	3	3.33	TEPAT	
EDY PURNOMO	LAKI - LAKI	MAHASISWA	24	BELUM MENIKAH	2.83	3	2.54	3.36	2.79	3.28	3.28	3	3.15	TEPAT	
DIDIK KURNIAWAN	LAKI - LAKI	MAHASISWA	24	BELUM MENIKAH	3	3.08	3.02	3.07	2.88	2.76	3	2.5	2.92	TEPAT	
LYDIA EVITA SANDRA DEWI	PEREMPUAN	MAHASISWA	25	BELUM MENIKAH	3.52	3.52	3.31	3.33	3.3	2.94	3.46	0	3.26	TEPAT	
DAVID KURNIAWAN	LAKI - LAKI	MAHASISWA	36	BELUM MENIKAH	2.29	0.5	0.16	0.32	1.13	0.37	0.28	0.11	1.02	TEPAT	
DEVI KISTIANI	PEREMPUAN	MAHASISWA	25	BELUM MENIKAH	3.52	3.46	3.69	3.41	3.4	0	0		3.43	TEPAT	
ADI PURWADI	LAKI - LAKI	MAHASISWA	27	BELUM MENIKAH	2.6	2.91	2.66	2.91	3.25	2.88	2.83	0	2.86	TEPAT	
ARIF WIBOWO	LAKI - LAKI	MAHASISWA	30	BELUM MENIKAH	2.52	2.59	2.11	3	2.73	2.5	2.48	0	2.68	TEPAT	
NOVAZEN DWI SAPUTRO	LAKI - LAKI	MAHASISWA	28	BELUM MENIKAH	2.79	3.05	3.21	2.92	3.41	2.86	2.5	0	3.04	TEPAT	
MOHAMAD ADE FADHORI	LAKI - LAKI	MAHASISWA	30	BELUM MENIKAH	2.93	2.52	1.95	2.37	3.07	2.91	1.73	1.6	2.81	TEPAT	
SUBAEDAH	PEREMPUAN	MAHASISWA	23	BELUM MENIKAH	3.52	3.48	3.71	3.83	3.25	3.65	3.33	3.5	3.56	TEPAT	
DANI PRIMAYANTI	PEREMPUAN	MAHASISWA	24	BELUM MENIKAH	3.48	3.73	3.6	3.52	2.9	3.32	3.21	0	3.45	TEPAT	
DERA BAHTIAR WIDIYANTO	LAKI - LAKI	MAHASISWA	23	BELUM MENIKAH	2.52	3.3	3.79	3.71	3.65	2.1	3.59	4	3.52	TEPAT	
MUKHAMAD YAHYA WICKAKSON	LAKI - LAKI	MAHASISWA	24	BELUM MENIKAH	2.45	2.66	2.93	3.08	2.29	3.18	2.1	3	3.1	TEPAT	
WAHYU DWI UTOHJO	LAKI - LAKI	MAHASISWA	28	BELUM MENIKAH	2.52	1.75	1.95	2.17	1.95	2.18	1.45	1.08	2.34	TEPAT	
AHMAD FIRDAUS MABRURI	LAKI - LAKI	MAHASISWA	25	BELUM MENIKAH	2.33	3	2.19	0.44	1.17	0.59	0.26		1.72	TEPAT	
HARYOKO ABDUL HAMID	LAKI - LAKI	MAHASISWA	26	BELUM MENIKAH	3.24	3.46	3.56	3.83	3.73	3.74	2.33	3.5	3.58	TEPAT	
ERSA GILANG ARDHANI	LAKI - LAKI	MAHASISWA	25	BELUM MENIKAH	2.95	2.91	3.44	3.38	3.33	3.36	2.08	0	3.1	TEPAT	
NEORITA NUR FATIMAH	PEREMPUAN	MAHASISWA	25	BELUM MENIKAH	3.48	3.5	2.77	3.59	3.5	3.35	3.27	0	3.31	TEPAT	
ULUL ALBAB	LAKI - LAKI	MAHASISWA	23	BELUM MENIKAH	2.93	3.18	3.31	3.25	3.44	3.22	3.07	3.5	3.33	TEPAT	
DEWI KHOIRUN NISA'	PEREMPUAN	MAHASISWA	27	BELUM MENIKAH	3	3.33	3.27	2.96	3.16	3.12	2.61	0	3.08	TEPAT	
AGUS NUGROHO	LAKI - LAKI	MAHASISWA	23	BELUM MENIKAH	2.83	3.11	3.1	2.96	3.2	3.53	2.48	3	3.27	TEPAT	
MUKHAMAD LIQMAN HABIBI	LAKI - LAKI	MAHASISWA	25	BELUM MENIKAH	2.88	3.25	3.42	3.29	3.33	2.75	3	0	3.13	TEPAT	
AMBAR SETIYANI	PEREMPUAN	MAHASISWA	25	BELUM MENIKAH	3.71	3.79	3.96	3.91	3.75	3.94	2.58	4	3.85	TEPAT	
AVINA SEPTIANA	PEREMPUAN	MAHASISWA	27	BELUM MENIKAH	3.07	2.85	3.2	3	3.21	2.47	2.55	2.75	3.1	TEPAT	
MARIA ULFA	PEREMPUAN	MAHASISWA	24	BELUM MENIKAH	3.19	2.9	2.59	2.77	3.09	2.83	3.11	2.9	3.08	TEPAT	
IVANDHY SETYARACHMAN	LAKI - LAKI	MAHASISWA	31	BELUM MENIKAH	2.86	2.98	3.39	3.27	2.98	2.85	2.12		3.21	TEPAT	
CHASAN MUBAROK	LAKI - LAKI	MAHASISWA	25	BELUM MENIKAH	2.98	3.18	3.73	3.42	3.5	3.47	2.58	3.5	3.52	TEPAT	

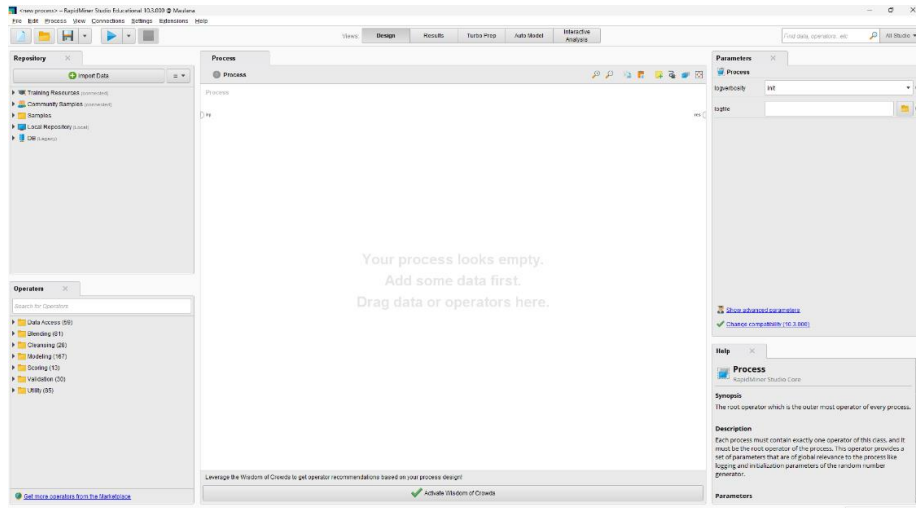
Gambar 1 Cobtoh Data

RapidMiner Studio 10.3 menyediakan lingkungan yang efisien dan efektif untuk melakukan pengujian data dan analisis prediktif. Proses pengujian yang terstruktur dan fitur-fitur canggih yang ditawarkan memastikan bahwa pengguna dapat memperoleh hasil yang akurat dan dapat diandalkan Berikut adalah proses pengujian data dan tampilan RapidMiner Studio 10.3



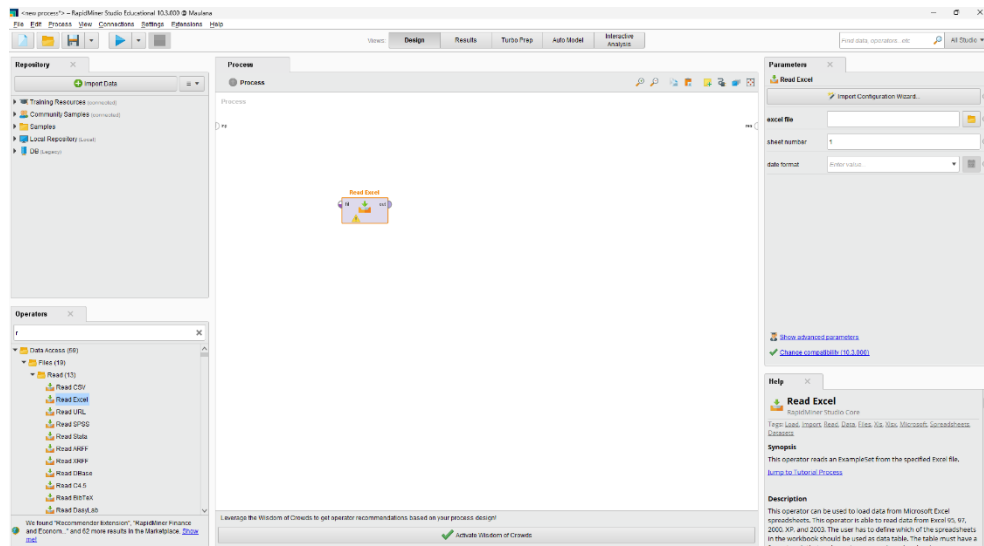
Gambar 2. 2 Tampilan Rapidminer Studio 10.3

Buka aplikasi Rapidminer, setelah muncul seperti maka akan muncul tampilan start with kemudian pilih blank process untuk membuka lembar kerja baru/ membuka file yang telah tersimpan seperti Gambar 2.3



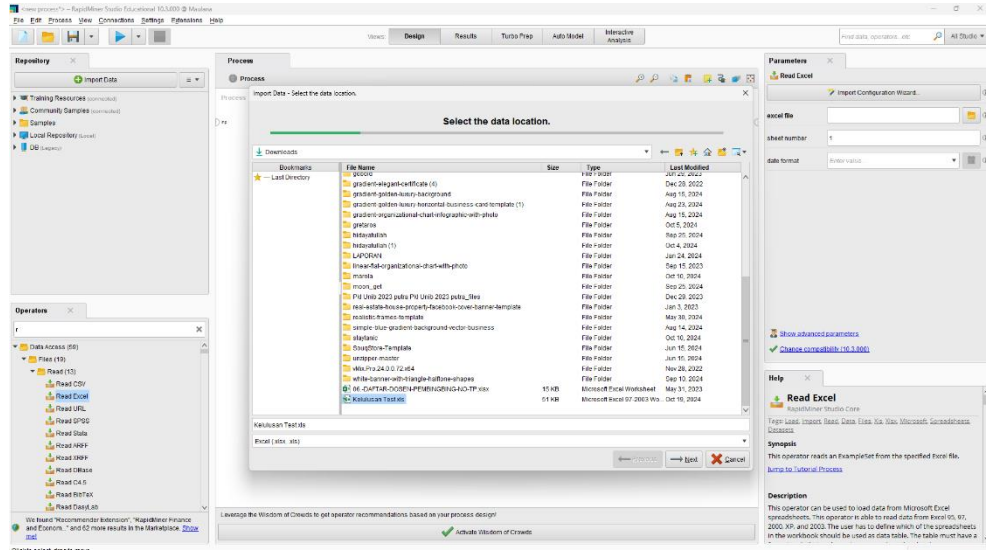
**Gambar 2**Tampilan Lembaran Kerja

Selanjutnya proses input data dan membutuhkan operator read excel dengan drag anddrop kedalam lembar kerja seperti Proses pengambilan data dengan menggunakan double click pada operator read excel



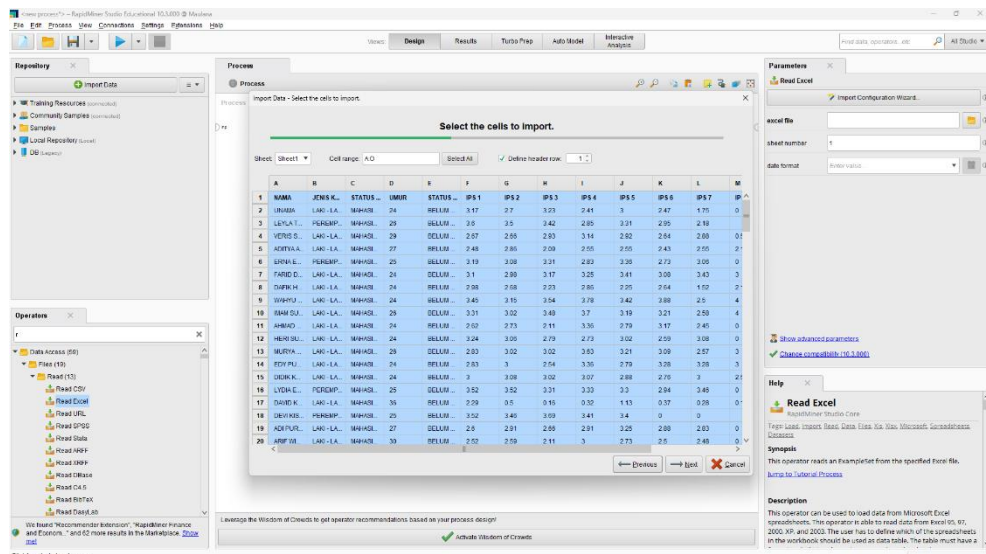
**Gambar 3** Operator Read Excel

Klik pada modul Read Excel untuk membuka jendela pemilihan file. Pada jendela ini, Anda akan diberikan opsi untuk mencari dan memilih file data yang ingin diimpor.gunakan tombol 'Browse' atau 'Pilih File' untuk membuka dialog pemilihan file. Navigasikan ke direktori tempat file Excel yang ingin Anda gunakan disimpan, pilih file data yang sesuai dengan kebutuhan analisis Anda. Pastikan bahwa file yang dipilih memiliki format yang benar dan data yang diperlukan sudah lengkap.



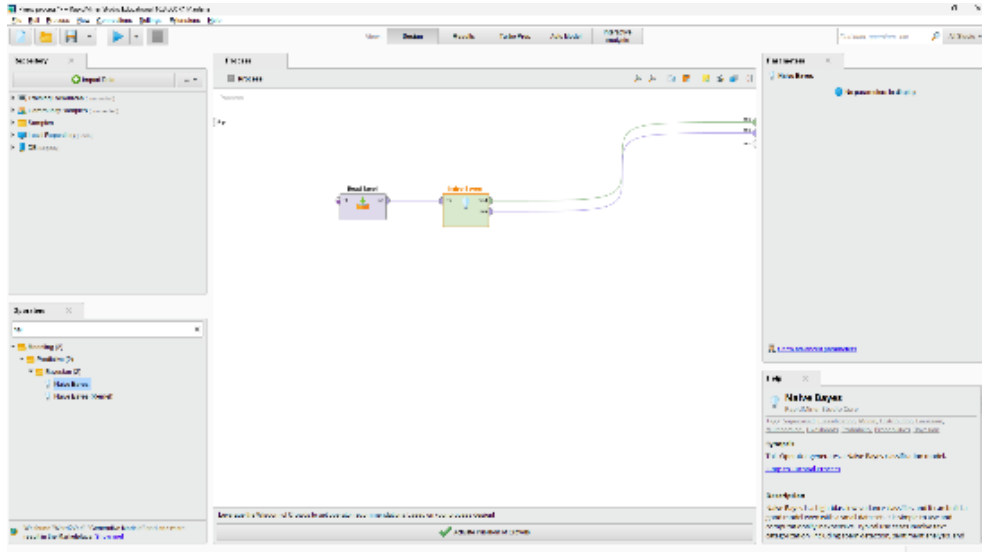
Gambar 4 Inport Data Mahasiswa

Setelah memilih file, langkah selanjutnya adalah memverifikasi bahwa file tersebut telah dimuat dengan benar ke dalam RapidMiner. Modul Read Excel akan menampilkan pratinjau data dari file yang dipilih, periksa header kolom, format data, dan pastikan tidak ada kesalahan dalam data yang diimpor. Jika diperlukan, lakukan pengaturan tambahan seperti mengatur delimiter, sheet yang akan dibaca, dan rentang sel.



Gambar 5 ferivikasi File

Selanjutnya klik next pada data yang dipilih seperti Setelah itu tentukan tipe data lalu pada kelas data diberi atribut label dengan pilihan change role dan klik finish maka tampilannya seperti Gambar 7 dan 8



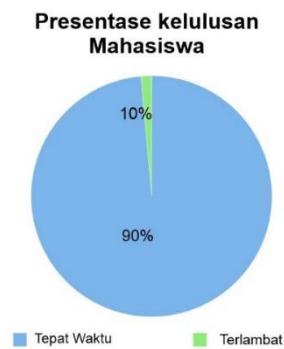
Gambar 6 Inport data

The image displays a data table with 146 rows and 14 columns. The columns include 'Row No.', 'STATUS KEL.', 'NAMA', 'JENIS KEL.', 'STATUS MA.', 'UMUR', 'STATUS MK.', 'IPS 1', 'IPS 2', 'IPS 3', 'IPS 4', 'IPS 5', 'IPS 6', and 'IPS 7'. The 'STATUS KEL.' column contains values like 'TEPAT' and 'TERLAMBAT'. The 'STATUS MA.' column contains 'MHAJIRIN' and 'BELUM MHAJIRIN'.

Row No.	STATUS KEL.	NAMA	JENIS KEL.	STATUS MA.	UMUR	STATUS MK.	IPS 1	IPS 2	IPS 3	IPS 4	IPS 5	IPS 6	IPS 7
39	TEPAT	CHAGAN MU.	LAKI-LAKI	MHAJIRIN	25	BELUM MHAJIRIN	2.980	3.180	3.730	3.420	3.550	3.470	2.580
112	TEPAT	BAMBANG ZA.	LAKI-LAKI	MHAJIRIN	28	BELUM MHAJIRIN	3.240	3.270	3.500	3.220	2.770	2.940	2.250
49	TEPAT	AUSTRIKA JA.	LAKI-LAKI	MHAJIRIN	23	BELUM MHAJIRIN	2.740	2.850	2.810	2.850	2.880	2.680	3.180
36	TEPAT	AJMA SEPTI.	PEREMPUAN	MHAJIRIN	27	BELUM MHAJIRIN	3.070	2.850	3.200	3	3.210	2.470	2.550
85	TEPAT	AFIF RICHAL	LAKI-LAKI	MHAJIRIN	40	BELUM MHAJIRIN	2.820	2.540	2.890	3.140	2.850	3.420	3.030
81	TEPAT	ALFAL WEDAD	LAKI-LAKI	MHAJIRIN	23	BELUM MHAJIRIN	2.620	2.380	2.250	2.270	2.210	2.160	2.860
135	TEPAT	ARSOFFA	PEREMPUAN	MHAJIRIN	23	BELUM MHAJIRIN	2.640	2.590	3.080	3.140	2.540	3.310	1.630
141	TEPAT	ARY AELI SE.	LAKI-LAKI	MHAJIRIN	23	BELUM MHAJIRIN	1.980	2.900	2.140	2.770	2.610	2.830	2.820
93	TEPAT	ARIS HEDIKAT	LAKI-LAKI	MHAJIRIN	24	BELUM MHAJIRIN	2.710	3.050	3.040	3.140	2.830	3	3.630
19	TEPAT	ARIF WEDVIO	LAKI-LAKI	MHAJIRIN	30	BELUM MHAJIRIN	2.520	2.590	2.110	3	2.730	2.500	2.480
83	TEPAT	ARIS CHATA.	PEREMPUAN	MHAJIRIN	24	BELUM MHAJIRIN	2.170	2.950	2.890	2.890	2.640	2.830	1.570
135	TEPAT	ANIS WATUL.	PEREMPUAN	MHAJIRIN	24	BELUM MHAJIRIN	2.880	2.770	2.610	3.230	2.710	3.450	3.110
126	TEPAT	ANIK FARVA.	PEREMPUAN	MHAJIRIN	25	BELUM MHAJIRIN	2.900	3.070	2.440	2.410	2.400	2.750	2.980
75	TEPAT	ANIF FARIZ.	LAKI-LAKI	MHAJIRIN	25	BELUM MHAJIRIN	2.740	2.640	2.390	2.820	2.790	3.950	2.700
60	TEPAT	ANIND PRANEE.	LAKI-LAKI	MHAJIRIN	23	BELUM MHAJIRIN	2.710	2.410	2.800	2.850	3.180	3.100	2.500
132	TERLAMBAT	ANIND ALFAN.	LAKI-LAKI	BEKERJA	23	BELUM MHAJIRIN	2.740	2.680	2.360	2.550	2.110	2.500	1.350
35	TEPAT	ANBAR SETI.	PEREMPUAN	MHAJIRIN	25	BELUM MHAJIRIN	3.710	3.790	3.950	3.910	3.750	3.940	2.580
76	TEPAT	ALJ HURSID	LAKI-LAKI	MHAJIRIN	22	BELUM MHAJIRIN	2.690	2.200	2.270	2.550	2.450	2.390	2.550
122	TEPAT	ALJ ANWAR D.	LAKI-LAKI	MHAJIRIN	25	BELUM MHAJIRIN	2.880	2.660	2.700	2.730	2.800	3.330	2.430
139	TEPAT	AJ FRASELYO	LAKI-LAKI	MHAJIRIN	23	BELUM MHAJIRIN	3.050	3.100	3.330	3.170	3.170	3.290	2.330
10	TEPAT	AHMAD SUT.	LAKI-LAKI	MHAJIRIN	24	BELUM MHAJIRIN	2.820	2.730	2.110	3.360	2.790	3.170	2.450
27	TEPAT	AHMAD FIRDI.	LAKI-LAKI	MHAJIRIN	25	BELUM MHAJIRIN	2.330	3	2.190	0.440	2.170	0.590	0.290
86	TEPAT	AHMAD FAZIN.	LAKI-LAKI	MHAJIRIN	23	BELUM MHAJIRIN	3	2.890	3.250	2.830	2.860	2.870	1.310
73	TEPAT	AGUSTRIANS	LAKI-LAKI	MHAJIRIN	23	BELUM MHAJIRIN	2.980	2.200	2.160	1.800	2.110	1	1.190
129	TEPAT	AGUS SUPRI.	LAKI-LAKI	MHAJIRIN	24	BELUM MHAJIRIN	3.050	3.230	3.420	3.420	3.250	3.990	2.420
134	TEPAT	AGUS BETRI.	LAKI-LAKI	MHAJIRIN	24	BELUM MHAJIRIN	3.070	3.210	2.840	3.390	3.150	2.770	3.470

Gambar 7 hasil Prediksi Pengujian Data Mahasiswa

Jumlah data testing yang digunakan sebanyak 146 mahasiswa dengan menggunakan metode Naive Bayes didapatkan hasil prediksi bahwa mahasiswa lulus tepat waktu sebanyak 144 mahasiswa



Gambar 8 Garfik Presentase Kelulusan Mahasiswa

Hasil prediksi pengujian data Jumlah data testing yang digunakan sebanyak 146 mahasiswa dengan menggunakan metode Naïve Bayes didapatkan hasil prediksi bahwa mahasiswa lulus tepat waktu sebanyak 144 mahasiswa dibersihkan kemudian digunakan. Adapun atribut yang digunakan yaitu jenis kelamin, Status Mahasiswa, Umur, indeks prestasi kumulatif (IPK). Data yang diambil dibagi menjadi dua kelompok, Yang memiliki dua kelas data yaitu tepat waktu dan terlambat. Selanjutnya pengujian dengan RapidMiner Studio 10.3 sebagai alat bantu untuk memperoleh tingkat akurasi dari metode Naïve Bayes. Berikut ini adalah perhitungan untuk mendapatkan nilai akurasi dengan persamaan berikut:

$$\begin{aligned} \text{Akurasi} &= \frac{\text{TP}+\text{TN}}{\text{TP}+\text{FN}+\text{FP}+\text{TN}} = \text{X}100\% \\ &= \frac{50+40}{50+5+5+40} \times 100\% \\ &= 90\% \end{aligned}$$

Prediksi dilakukan dengan data training menggunakan Naïve Bayes mendapatkan hasil sebanyak 144 mahasiswa diprediksi tepat waktu, 2 mahasiswa diprediksi terlambat.

#### 4. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian menunjukkan bahwa metode data mining Naïve Bayes dapat digunakan untuk memprediksi kelulusan siswa tepat waktu dengan tingkat akurasi yang baik. Dari hasil tes terhadap 146 siswa, prediksi menunjukkan bahwa 144 siswa lulus tepat waktu dan 2 siswa terlambat. Hal ini menunjukkan bahwa Naïve Bayes merupakan algoritma yang efektif dan efisien dalam mengklasifikasikan kelulusan siswa meskipun menggunakan asumsi independensi atribut.

Penerapan atribut seperti nama mahasiswa, NPM, jalur masuk, IPK dan jenis kelamin pada metode ini terbukti memberikan informasi akurat mengenai kelulusan mahasiswa. Dengan sistem prediksi kelulusan ini, perguruan tinggi dapat memperoleh informasi berharga untuk membuat kebijakan yang dapat meningkatkan jumlah mahasiswa yang lulus tepat waktu. Oleh karena itu, metode Naïve Bayes sangat cocok digunakan dalam sistem prediksi kelulusan mahasiswa untuk membantu perguruan tinggi meningkatkan kualitas pendidikan dan kinerja institusinya.

#### Ucapan Terima Kasih

Terima kasih kepada semua pihak atas dukungan dan bantuan dalam penyelesaian penelitian ini. dengan penuh rasa syukur, kami mengucapkan terima kasih kepada semua pihak yang telah mendukung dan membantu terlaksananya penelitian ini. Pertama-tama, terima kasih kepada Tuhan Yang Maha Esa atas rahmat-Nya. Terima kasih kepada keluarga kami atas dukungan moral dan motivasi; kepada dosen dan pembimbing di Universitas Ibrahimy yang memberikan bimbingan dan pengetahuan berharga; serta rekan-rekan mahasiswa atas bantuan dan kolaborasi mereka. Terima kasih juga kepada seluruh staf akademik dan administratif di Fakultas Sains & Teknologi yang menyediakan fasilitas dan dukungan teknis. Semoga penelitian ini bermanfaat bagi perkembangan ilmu pengetahuan.

#### DAFTAR PUSTAKA

- [1] M. P. Dr. Achmad Fathoni Rodli and S. E. M. S. M. Nikma Yucha, *Manajemen Kinerja Institusi Perguruan Tinggi*. CV Rey Media Grafika, 2022. [Online]. Available: <https://books.google.co.id/books?id=S6p-EAAAQBAJ>
- [2] L. Setiyani, M. Wahidin, D. Awaludin, and S. Purwani, "Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes : Systematic Review," *Fakt. Exacta*, vol. 13, no. 1, p. 35, 2020, doi: 10.30998/faktorexacta.v13i1.5548.
- [3] A. Wasik *et al.*, "Implementasi data mining untuk memprediksi penjualan accessoris handphone dan handphone terlaris menggunakan metode k-nearest neighbor (k-nn) 1," vol. 1, no. 2, pp. 469–479, 2024.
- [4] E. luthfi and U. Amikom, *Algoritma Data Mining*. Penerbit Andi, 2009. [Online]. Available: <https://books.google.co.id/books?id=-Ojclag73O8C>

- [5] A. S. Huda, R. M. Awangga, and R. N. S. Fathonah, *Prediksi Penerimaan Pegawai Baru Dengan Metode Naive Bayes*. in *Data Science*. Kreatif, 2020. [Online]. Available: <https://books.google.co.id/books?id=aawGEAAAQBAJ>
- [6] W. I. Rahayu, M. H. K. Saputra, R. M. Awangga, and R. Habibi, *Penerapan Metode Naive Bayes dan Skala Likert Pada Aplikasi Prediksi Kelulusan Mahasiswa*. Kreatif, 2020. [Online]. Available: <https://books.google.co.id/books?id=CZv9DwAAQBAJ>
- [7] S. P. M. S. Ni'matuzahroh and M. P. Susanti Prasetyaningrum, *OBSERVASI: TEORI DAN APLIKASI DALAM PSIKOLOGI*. in 1. UMMPress, 2018. [Online]. Available: <https://books.google.co.id/books?id=CMh9DwAAQBAJ>
- [8] M. Zed, *Metode Penelitian Kepustakaan*. Yayasan Pustaka Obor Indonesia, 2008. [Online]. Available: <https://books.google.co.id/books?id=zG9sDAAAQBAJ>
- [9] L. J. Moleong and T. Surjaman, *Metodologi penelitian kualitatif*. Remadja Karya, 1989. [Online]. Available: <https://books.google.co.id/books?id=YXsknQEACAAJ>
- [10] M. Hasanah, N. H. Harani, N. Riza, R. M. Awangga, and R. Habibi, *Implementasi Barcode Dan Algoritma Regresi Linear Untuk Memprediksi Data Persediaan Barang*. Kreatif. [Online]. Available: <https://books.google.co.id/books?id=db38DwAAQBAJ>