

## BAB III METODOLOGI PENELITIAN

### 3.1 Gambaran Umum Objek Penelitian

Studi kasus ini memprediksi kelulusan siswa STABN Sriwijaya dengan Memakai metode klasifikasi data mining. Atribut yang digunakan antara lain Jurusan, IPS (*Indeks Prestasi Semester*), IPK (*Indeks Prestasi Kumulatif*), yang mempengaruhi kelulusan. Proses klasifikasi dilakukan dalam Memakai algoritma *Naives Bayes* dan C4.5 dan diuji pada data baru yang belum pernah diamati sebelum ini. Tujuan dari penelitian ini adalah untuk memberikan informasi dan rekomendasi kepada STABN Sriwijaya untuk meningkatkan tingkat kelulusan mahasiswa. Data yang diolah untuk penelitian ini terdiri dari dataset berikut:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	Nama	Jurusan	Jenis Kelamin	Angkatan	IPS 1	IPS 2	IPS 3	IPS 4	IPS 5	IPS 6	IPS 7	IPK	Status Kelulusan		
1	Dyah Pittaloka Puspita Rini	Pendidikan Agama Buddha	Perempuan	2016	2.93	2.87	3.18	3.23	3.22	3.53	3.2	3.165714286	Tepat Waktu		
2	Ageng Seger Wibowo	Pendidikan Agama Buddha	Laki-Laki	2016	2.2	2.15	2.12	1.92	1.89	2.35	1.5	2.018571429	Tepat Waktu		
3	Cici Dewi Paramita	Pendidikan Agama Buddha	Laki-Laki	2016	3.58	3.4	3.05	3.73	3.28	3.58	3.67		3.47 Tepat Waktu		
4	Darma Setiawan	Pendidikan Agama Buddha	Perempuan	2016	3.75	3.55	3.79	3.84	3.78	3.84	3.8	3.764285714	Tepat Waktu		
5	Duti Purvosari Ningsih	Pendidikan Agama Buddha	Perempuan	2016	3.28	3.37	3.51	3.57	3.62	3.67	3.43	3.492857143	Tepat Waktu		
6	Gunawan	Pendidikan Agama Buddha	Perempuan	2016	3.22	3.09	3.12	3.28	3.33	3.46	3.57	3.295714286	Tepat Waktu		
7	Herni Erliawati	Pendidikan Agama Buddha	Perempuan	2016	2.88	2.94	2.94	2.97	3.08	3.44	3.43	3.097142857	Tepat Waktu		
8	Hendaryadi	Pendidikan Agama Buddha	Laki-Laki	2016	2.63	2.61	2.75	2.9	3.03	3.17	2.43	2.788571429	Tepat Waktu		
9	Ida Lestari	Pendidikan Agama Buddha	Perempuan	2016	2.97	2.99	2.9	3.37	3.3	3.5	3.33	3.194285714	Tepat Waktu		
10	Satya Tiratana	Pendidikan Agama Buddha	Laki-Laki	2016	2.55	2.24	2.15	2.38	2.28	2.7	2.4	2.385714286	Terlambat		
11	Santi Kristina	Pendidikan Agama Buddha	Perempuan	2016	2.91	3.15	3.48	3.57	3.57	3.68	3.67	3.432857143	Tepat Waktu		
12	Santi	Pendidikan Agama Buddha	Laki-Laki	2016	3.49	3.25	3.39	3.67	3.53	3.61	3.53	3.495714286	Tepat Waktu		
13	Suyanti	Pendidikan Agama Buddha	Laki-Laki	2016	2.31	2.3	2.37	2.38	2.5	2.91	1.94	2.387142857	Terlambat		
14	Sauudin	Kepenyuluhan Buddha	Laki-Laki	2016	3.13	2.68	2.47	2.85	3.43	3.24	3.26	3.008571429	Tepat Waktu		
15	Saril Ependi	Kepenyuluhan Buddha	Laki-Laki	2016	2.3	2.17	2.35	2.64	2.72	3.07	2.64	2.555714286	Terlambat		
16	Anah Susilawati	Kepenyuluhan Buddha	Perempuan	2016	2.7	2.85	2.74	3.05	3.12	3.14	3.17	2.967142857	Tepat Waktu		
17	Anes Dwi Prasetya	Kepenyuluhan Buddha	Laki-Laki	2016	2.76	2.34	2.39	2.25	2.85	2.73	3.12	2.634285714	Terlambat		
18	Ahyar	Kepenyuluhan Buddha	Perempuan	2016	2.69	3.09	2.93	3.02	3.28	3.36	3.17	3.077142857	Tepat Waktu		
19	Adhita Meiliana Sari	Kepenyuluhan Buddha	Laki-Laki	2016	2.2	2.16	2.25	2.51	2.79	2.81	2.47	2.455714286	Terlambat		
20	Dariyanto	Pendidikan Agama Buddha	Laki-Laki	2016	3.1	3.15	3.21	3.25	3.54	3.65	0	2.842857143	Terlambat		

Gambar 3. 1 Data Latih

Gambar 3.1 merupakan data pelatihan dimana data pelatihan ini digunakan menjadi salahsatu dataset dalam penelitian ini pada data latih ini terdapat beberapa atribut yang digunakan seperti Nama, Jurusan, Jenis Kelamin, Angkatan, IPS 1, IPS 2, IPS 3, IPS 4, IPS 5, IPS 6, dan IPS 7 IPK dan status kelulusan, data yang diperoleh adalah data dari Angkatan 2016 hingga 2022 data ini di dapatkan dari bidang kemahasiswaan STABN Sriwijaya, untuk status kelulusan sendiri berdasarkan hasil wawancara dengan pihak kampus jika range IPK mahasiswa 2.75 – 4.00 bisa diprediksi jika mahasiswa tersebut

akan lulus tepat waktu sedangkan jika ipk dibawah 2.75 biasanya mahasiswa tersebut lulus terlambat.

	A	B	C	D	E	F	G	H	I	J	K
	Nama	Jurusan	Jenis Kelamin	Angkatan	IPS 1	IPS 2	IPS 3	IPK	Status Kelulusan		
1	Dyah Pitaloka Puspita Rini	Pendidikan Agama Buddha	Perempuan	2016	2.93	2.87	3.18	2.993333333	Tepat Waktu		
2	Ageng Seger Wibowo	Pendidikan Agama Buddha	Laki-Laki	2016	2.2	2.15	2.12	2.156666667	Terlambat		
3	Cici Dewi Paramita	Pendidikan Agama Buddha	Laki-Laki	2016	3.58	3.4	3.05	3.343333333	Tepat Waktu		
4	Darma Setiawan	Pendidikan Agama Buddha	Perempuan	2016	3.75	3.55	3.79	3.696666667	Tepat Waktu		
5	Duti Purwosari Ningsih	Pendidikan Agama Buddha	Perempuan	2016	3.28	3.37	3.51	3.386666667	Tepat Waktu		
6	Gunawan	Pendidikan Agama Buddha	Perempuan	2016	3.22	3.09	3.12	3.143333333	Tepat Waktu		
7	Herni Erlawati	Pendidikan Agama Buddha	Perempuan	2016	2.88	2.94	2.94	2.92	Tepat Waktu		
8	Hendaryadi	Pendidikan Agama Buddha	Laki-Laki	2016	2.63	2.61	2.75	2.663333333	Terlambat		
9	Ida Lestari	Pendidikan Agama Buddha	Perempuan	2016	2.97	2.99	2.9	2.953333333	Tepat Waktu		
10	Setya Tiratana	Pendidikan Agama Buddha	Laki-Laki	2016	2.55	2.24	2.15	2.313333333	Terlambat		
11	Santi Kristina	Pendidikan Agama Buddha	Perempuan	2016	2.91	3.15	3.48	3.18	Tepat Waktu		
12	Santri	Pendidikan Agama Buddha	Laki-Laki	2016	3.49	3.25	3.39	3.376666667	Tepat Waktu		
13	Suyanti	Pendidikan Agama Buddha	Laki-Laki	2016	2.31	2.3	2.37	2.326666667	Terlambat		
14	Sanudin	Kepenyulahan Buddha	Laki-Laki	2016	3.13	2.68	2.47	2.76	Tepat Waktu		
15	Saril Ependi	Kepenyulahan Buddha	Laki-Laki	2016	2.3	2.17	2.35	2.273333333	Terlambat		
16	Anah Susilavati	Kepenyulahan Buddha	Perempuan	2016	2.7	2.85	2.74	2.763333333	Tepat Waktu		
17	Anes Dwi Prasetya	Kepenyulahan Buddha	Laki-Laki	2016	2.76	2.34	2.39	2.496666667	Terlambat		
18	Ahyar	Kepenyulahan Buddha	Perempuan	2016	2.69	3.09	2.93	2.903333333	Terlambat		
19	Adhita Meilliana Sari	Kepenyulahan Buddha	Laki-Laki	2016	2.2	2.16	2.25	2.203333333	Terlambat		
20	Dariyanto	Pendidikan Agama Buddha	Laki-Laki	2016	3.1	3.15	3.21	3.153333333	Tepat Waktu		
21											

Gambar 3. 2 Data Testing

Gambar 3.2 merupakan data set yang digunakan pada penelitian kali ini dimana pada gambar diatas terdapat beberapa atribut yang digunakan seperti nama, jurusan, jenis kelamin, Angkatan, IPS 1, IPS 2, IPS 3, IPK serta status kelulusan, data yang diperoleh yaitu Angkatan 2016 sampai dengan Angkatan 2022, data ini di dapatkan dari bidang kemahasiswaan STABN Sriwijaya, untuk status kelulusan sendiri berdasarkan hasil wawancara dengan pihak kampus jika range IPK mahasiswa 2.75 – 4.00 bisa diprediksi jika mahasiswa tersebut akan lulus tepat waktu sedangkan jika range ipk dibawah 2.75 biasanya mahasiswa tersebut lulus terlambat.

### 3.1.1 Profile Kampus STABN Sriwijaya

Pada tahun 1934, Dengan kehadiran Bodhisattva Narada Maha Thera dari Sri Lanka, umat Buddha dibangun untuk menegakkan prinsip dharma yang tinggi. Pada saat itu, keadaan pelayanan internal umat Buddha sangat memprihatinkan. Ini termasuk kurangnya guru agama Buddha di semua tingkat sekolah, kurangnya tenaga kerja untuk melayani keagamaan Buddha, dan kurangnya perhatian tokoh-

tokoh umat Buddha terhadap pendidikan agama Buddha. Akibatnya, para pendiri Yayasan Sriwijaya memutuskan untuk mengambil kelas Buddha.

Sekolah Tinggi Agama Buddha (STAB), sebuah lembaga privat yang telah menghasilkan guru dan pelayan keagamaan Buddha, tidak memenuhi semua persyaratan kualitas dan kuantitas umat Buddha. Akibatnya, ada upaya untuk membangun Sekolah Tinggi Agama Buddha Sriwijaya

Atas arahan Dr. I Wayan Suarjaya, M.Si., Dirjen Bimas Hindu dan Budha, Yayasan Sriwijaya didirikan pada tanggal 5 September 2001, Dr. Joko Wuryanto, S.Sos., dan Sapardi, S.Ag., bersama dengan Cornelis Wowor, M.A. (Direktur Urusan Agama Buddha), Dr. Budi Setiawan, M.Sc., dan Pdt. T. Harmanto. Sekolah Tinggi Agama Buddha Sriwijaya didirikan oleh Yayasan Sriwijaya dengan SK Nomor 13/YYS-SWJ/V/2002 pada 5 Mei 2002. Beberapa jurusan STAB adalah Dharmaduta, Dharmacarya, Filsafat dan Sastra Buddha, Ekonomi dan Wirausaha Buddha, dan program pascasarjana.

### 3.1.2 Visi dan Misi Kampus STABN Sriwijaya

**b. Visi**

Menjadi perguruan tinggi Buddha terkemuka dengan ciri-ciri nasional.

**c. Misi**

Menciptakan lulusan yang berkualitas tinggi dan berkarakter melalui:

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

1. Pendidikan dan pengajaran yang berkualitas dan bersaing;
2. Penelitian yang berbasis kenusantara di bidang pendidikan, agama, dan keagamaan;
3. Pengabdian kepada masyarakat berdasarkan nilai-nilai ajaran Buddha; dan
4. Kerjasama dengan lembaga pendidikan dan non-pendidikan.

### 3.2 Metode Penelitian

Penelitian ini Memakai data mining dengan dua metode klasifikasi., yaitu *Naive Bayes* dan C4.5. Algoritma C4.5 dan *Naive Bayes* dipilih dalam penelitian ini karena kemudahan implementasi, hasil lebih bagus dalam hal klasifikasi, dan kelebihan dari setiap algoritma (C4.5 untuk mengatasi atribut diskrit dan numerik, *Naive Bayes* dengan sekali pemindaian data pelatihan). Untuk Kedua algoritma ini dapat memberikan prediksi yang akurat dan efisien terkait kelulusan mahasiswa tepat waktu. Perbandingan kedua metode tersebut dapat dilihat pada tabel berikut:

Tabel 3. 1 Perbandingan Dua Metode

Perbedaan	Algoritma	
	Naïve Bayes	C4.5
<i>Output</i>	Simple Distribusi	Pohon Keputusan
Persentase pengguna pada RapidMiner Community	12.6%	19.3%
Kelebihan	Updatable Mampu handle data paling banyak Dapat handle missing value	Dapat handle missing value Paling hemat waktu dalam pembuatan mode
Kekurangan	Tidak dapat menangani data yang tidak memenuhi asumsi indepedensi	Dapat menjadi kompleks dan sulit untuk di presentasikan

Berdasarkan tabel 3.2 diatas dapat dilihat perbedaan antara kedua metode perbedaan pertama adalah dari output dimana output untuk metode naïve bayes adalah simple distribusi, sedangkan untuk algoritma C4.5 adalah pohon keputusan, persentase pengguna metode algoritma lebih tinggi daripada metode *Naïve Bayes*, metode naïve bayes memiliki kelebihan updatable, mampu handle data paling banyak, dan dapat handle missing value atau data yang hilang, sedangkan kelebihan metode algoritma C4.5 adalah dapat handle data yang hilang, waktu pembuatan mode lebih hemat, untuk kekurangan dari metode naïve bayes adalah tidak dapat menangani data yang tidak memenuhi asumsi independensi, sedangkan untuk metode C4.5 dapat menjadi kompleks dan sulit di presentasikan.

#### Perbandingan Dengan Tools WEKA

Penggunaan tool rapid miner tidak terlepas dari beberapa penelitian yang menggunakan tools weka dimana, tools weka memiliki nilai akurasi yang kurang besar jika dibandingkan dengan tools rapidminer, berikut hasil dari penggunaan tools weka.

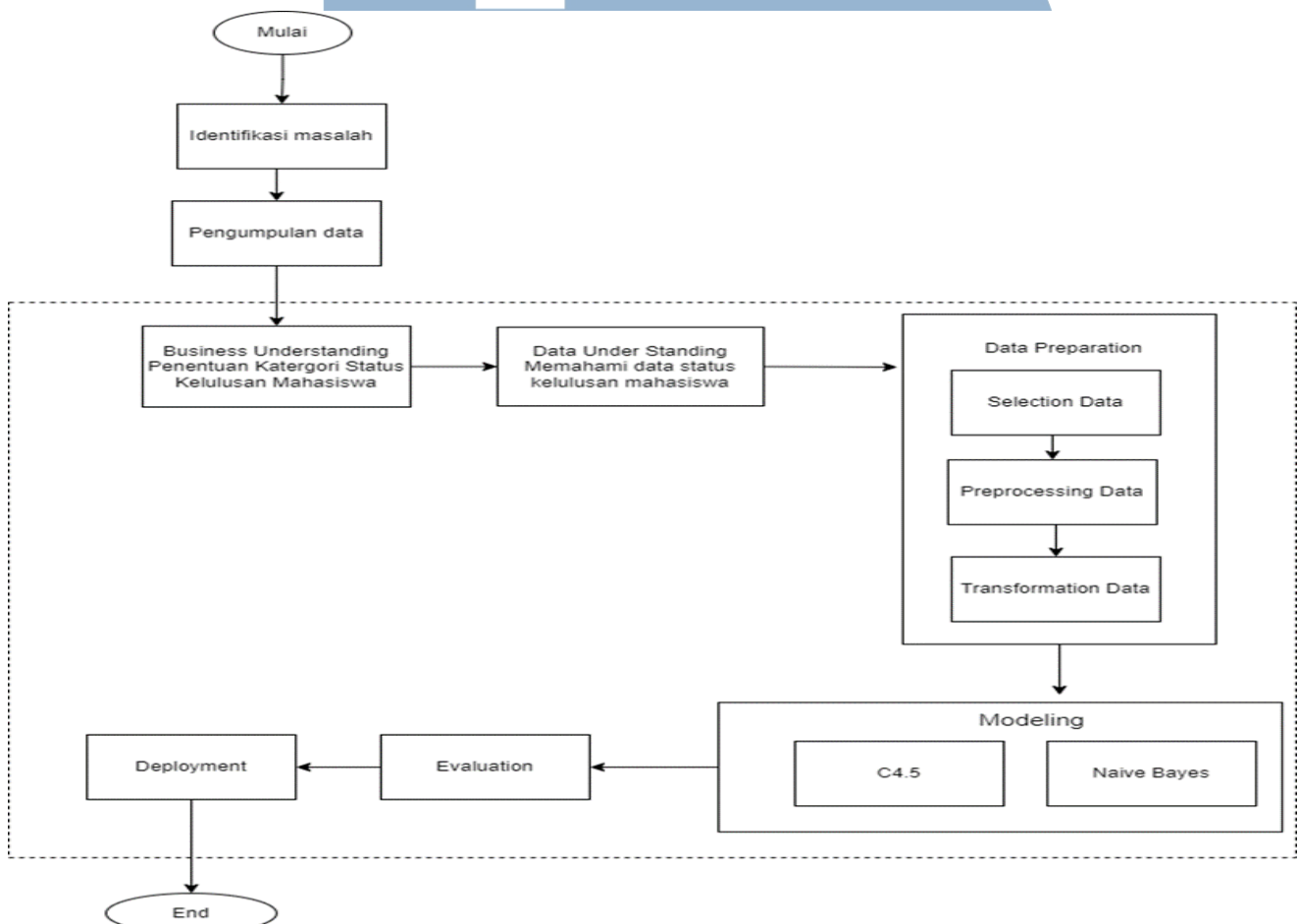
Tabel 3. 2 Hasil Tools Weka

Performance	RapidMiner	WEKA
Access	Free Community Edition, Commercial Enterprise Edition	Open Source
Programming Language	Java	Java
Launch Date	2001	2002
Development Team	Rapid-I Foundation	University Of Waikato
Kelebihan	<ul style="list-style-type: none"> <li>-Mempunyai statistikal dan prediktif analisis yang mudah diimplementasi pada sistem</li> <li>-Mempunyai kapabilitas algoritma terbanyak</li> <li>-Mempunyai user interface yang menarik dan lebih grafis (GUI / graphic user</li> </ul>	<ul style="list-style-type: none"> <li>-Software WEKA berlisensi open source, sehingga tidak berbayar - Dapat berjalan diberbagai macam platform, sehingga bersifat portable</li> </ul>

	<p>interface). -Mampu melakukan operasi parameter machine learning / metode statistic -Mampu validasi model dengan cross-validation dengan independent validation set</p>	<p>-Memiliki GUI yang mudah dipahami oleh pengguna awam</p> <p>-Mampu diimplementasikan menggunakan bahasa pemrograman Java</p> <p>-Tingkatan pengguna WEKA bisa untuk pemula maupun expert dikarenakan banyaknya fitur built-in</p> <p>-Less programming / coding knowledge.</p>
Kekurangan	<p>-Diharuskan mempunyai license RapidMiner community untuk mengakses aplikasi.</p> <p>-More programming / coding knowledge.</p>	<p>-Tidak dapat memasukan data yang banyak untuk diproses (overloading) -Tidak dapat menyimpan parameter untuk di apply pada future dataset. - Tidak otomatis dalam melakukan operasi parameter machine learning / metode statistik -Tidak bisa menyimpan proses saat validasi model dengan cross validation, maka diharuskan rebuild model.</p>

Berdasarkan tabel 3.3 perbandingan tools data mining, penelitian ini menggunakan tools Rapiminer karena memiliki beberapa keunggulan dibandingkan dengan tools WEKA. Pemilihan tools Rapidminer juga didukung dengan penelitian yang dilakukan pada tahun 2019, dimana jika tools rapidminer memiliki kecepatan yang lebih baik dibandingkan dengan tools weka [31].

### 3.2.1 Alur Penelitian



Gambar 3. 3 Alur Penelitian

Gambar 3.3 merupakan alur penelitian dalam pemodelan pada penelitian dari laporan yang berjudul Pemilihan Algoritma Data Mining Terbaik Untuk Prediksi Kelulusan Mahasiswa Studi Kasus STABN Sriwijaya, berikut ini merupakan penjelasan dari alur penelitian yang dibuat :

#### 1. Identifikasi Masalah

STABN Sriwijaya merupakan salah satu Sekolah Tinggi Agama Buddha Negeri Sriwijaya adalah institusi pendidikan keagamaan Buddha Negeri yang menciptakan tenaga pendidik maupun tenaga pelayan agama buddha untuk memenuhi segala kebutuhan umat Buddha, dalam segi kuantitas dan kualitas. Juga seperti hal yang lainnya, dosen STABN Sriwijaya juga membimbing mahasiswa selama kurang lebih 4 tahun lamanya agar mahasiswa bisa lulus tepat waktu dan tidak memiliki kendala, Kampus STABN Sriwijaya juga memiliki tujuan yaitu menghasilkan mahasiswa dengan lulusan yang berkompeten dan berkarakter. Kampus STABN Sriwijaya pada setiap tahunnya melakukan wisuda kelulusan mahasiswa. Namun, berdasarkan hasil dari observasi dengan pegawai kampus di STABN Sriwijaya masih terdapat masalah yaitu mahasiswa yang lulusnya terlambat.hal tersebut diperkuat dengan masih banyak Angkatan 2016-2018 yang masih aktif dan belum lulus, Untuk memprediksi status kelulusan mahasiswa, perlu dilakukan analisis tentang prediksi kelulusan mahasiswa menggunakan metode data mining dengan algoritma C4.5 dan Naïve Bayes.

## **2. Pengumpulan Data**

Pengumpulan data pada penelitian dengan cara observasi untuk memperoleh data mahasiswa dari tahun 2016 sampai dengan 2022.

## **3. *Business Understanding***

Tujuan penelitian ini adalah untuk mengumpulkan data mahasiswa yang diperlukan untuk memprediksi status kelulusan mahasiswa di STABN Sriwijaya Kampus.

## **4. *Data Understanding***

Untuk mengelompokan data status kelulusan mahasiswa berdasarkan segmentasi siswa, penetapan dan validasi data diperlukan pada tahap pemahaman data.

## **5. *Data Preparation***

Tahapan atau proses data persiapan, di mana kumpulan data disesuaikan untuk memenuhi persyaratan pada tahap pemodelan. Pengelompokan data,



atau transformasi data, yang diselesaikan untuk mengkonversi data asli menjadi data yang siap digunakan.

## 6. *Modeling*

Pada tahap ini, model dibuat menggunakan teknik Metode pengolahan data yang digunakan. Langkah-langkah algoritma C4.5 diimplementasikan dalam visualisasi dengan alat Rapidminer dan Microsoft Excel:

- a) Import dataset yang digunakan ke dalam local repository.
- b) Memasukan operator dataset dan operator rule serta operator decision tree.
- c) Menekan tombol plan untuk mengeksekusi proses yang sudah dilakukan.

Sedangkan untuk langkah – langkah pada metode naïve bayes adalah sebagai berikut:

- a) Import dataset yang digunakan dan simpan ke loca repository
- b) Memasukan operator dataset dan operator naïve bayes
- c) Menekan tombol play untuk mengeksekusi metode Naïve Bayes

## 7. *Evaluation*

Tahap ini dimana tahapan menganalisis hasil pengukuran yang telah dilakukan dengan menggunakan kedua metode yang digunakan dimana kedua hasil tersebut kemudian disempurnakan dengan mengukur tingkat akurasi, precision, recall dengan menggunakan confusion matrix.

## 8. *Deployment*

Tahapan ini merupakan tahapan dimana hasil prediksi dari metode naïve bayes serta algoritma C4.5 digunakan di dalam hal prediksi tingkat kelulusan pada kampus STABN Sriwijaya.

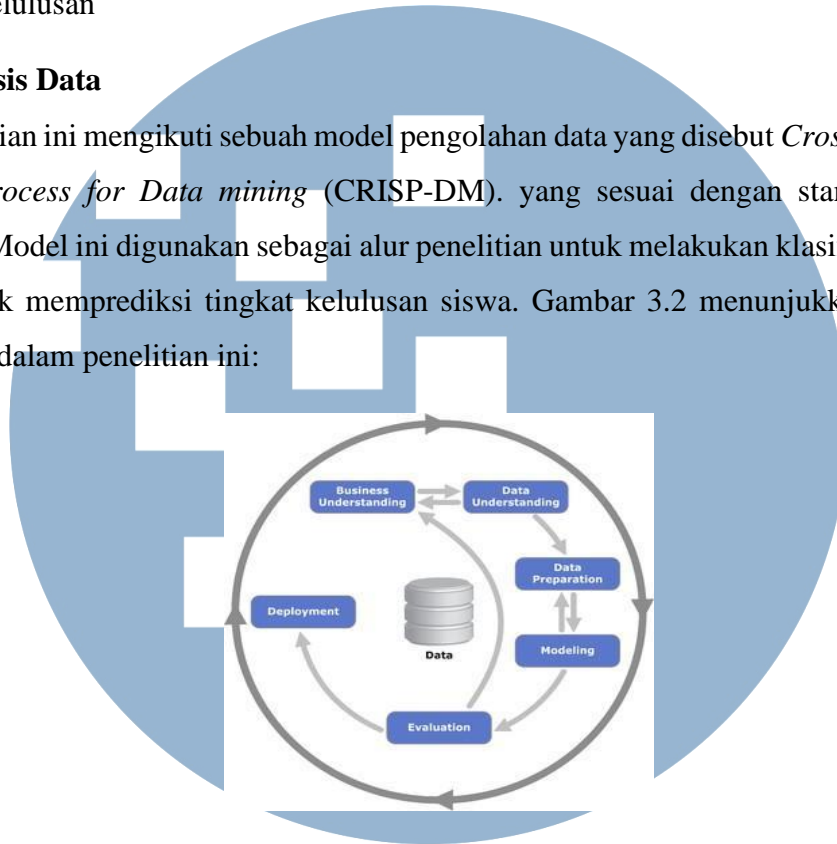
### 3.4 Teknik Pengumpulan Data

Penelitian ini Memakai dataset berasal dari STABN Sriwijaya. untuk mendapatkan data dengan melakukan observasi datang ke STABN Sriwijaya dataset yang didapat yaitu

berisi data dari akademik mahasiswa, seperti Nama, Jurusan, Jenis Kelamin, IPS, IPK dan status kelulusan

### 3.5 Teknik Analisis Data

Penelitian ini mengikuti sebuah model pengolahan data yang disebut *Cross Industry Standard Process for Data mining* (CRISP-DM). yang sesuai dengan standar yang ditetapkan, Model ini digunakan sebagai alur penelitian untuk melakukan klasifikasi data mining untuk memprediksi tingkat kelulusan siswa. Gambar 3.2 menunjukkan proses CRISP-DM dalam penelitian ini:



Gambar 3. 4 CRISP-DM

#### 3.5.1 *Business Understanding*

Dalam ini, memahami tujuan bisnis dan mengidentifikasi masalah yang akan diselesaikan Memakai data mining. Tujuan adalah untuk memprediksi kelulusan mahasiswa apakah akan lulus pada waktu yang tepat atau terlambat.

#### 3.5.2 *Data Understanding*

Dalam ini, data yang diperlukan dikumpulkan dan dipahami, termasuk jenis data, ukuran, dan nilai-nilai yang hilang atau tidak valid. Data yang dibutuhkan antara lain data mahasiswa dari Angkatan 2016 - 2022, Jurusan, IPS, IPK, dan status kelulusan.

#### 3.5.3 *Data Preparation*

Tahap ini adalah untuk menyiapkan data supaya dapat dipakai untuk analisis data. Hal ini meliputi pembersihan data, penghapusan data yang

tidak relevan, penggabungan data, dan pemilihan atribut. Misalnya, data yang tidak lengkap atau tidak valid akan dihapus atau diisi dengan nilai yang tepat, dan atribut yang tidak relevan akan dihapus.

#### **3.5.4 Modeling**

Tahap ini adalah untuk memilih teknik pemodelan yang sesuai dengan tujuan bisnis dan data yang telah dipersiapkan. Dalam hal ini, teknik pemodelan yang tepat adalah algoritma klasifikasi seperti *Naïve bayes* dan C4.5.

#### **3.5.5 Evaluation**

Dalam ini, model yang dibangun pada tahap sebelumnya akan dievaluasi dan mengukur keberhasilan model dalam memecahkan masalah. Model akan diuji memakai data data uji dan diukur akurasinya, *precision* dan *recall*. Untuk mendapatkan nilai akurasi yang dimaksud yaitu menggunakan confusion matrix.

#### **3.5.6 Deployment**

Tahap ini adalah untuk mengimplementasikan model yang telah dibangun dan dievaluasi pada tahap sebelumnya ke dalam lingkungan produksi sehingga dapat memberikan nilai bisnis yang nyata. Model dapat digunakan untuk memprediksi mahasiswa yang berisiko putus kuliah dan memberikan tindakan yang tepat untuk mengurangi risiko tersebut dengan menggunakan Rapidminer.

U M M N  
U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A