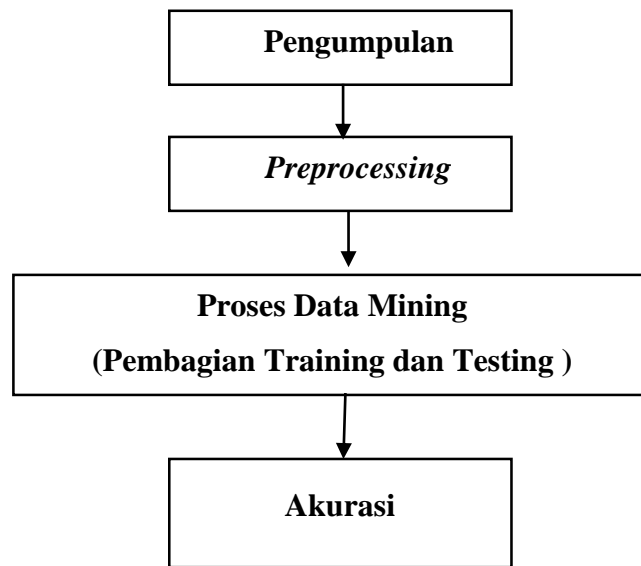


BAB III

METODOLOGI PENELITIAN

3.1 Alur Penelitian

Penelitian ini akan dilaksanakan berdasarkan alur penelitian seperti yang ditunjukkan pada Gambar 3.1.



Gambar 3.1 Alur Penelitian

3.2 Tempat dan Waktu Penelitian

Adapun tempat dan waktu penelitian yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Tempat

Penelitian akan dilakukan pada SMK Islam Al-Barokah Poncowarno.

2. Waktu

Dalam melaksanakan tahapan penelitian, peneliti merencanakan waktu penelitian dari bulan November 2020 sampai Januari 2021.

3. Data

Data yang akan digunakan dalam penelitian ini adalah data Alumni SMK Islam Al-Barokah Poncowarno.

3.3 Alat dan Bahan

Adapun alat dan bahan yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Hardware

Kebutuhan perangkat keras (hardware) yang digunakan : Laptop Toshiba Satellite L645 Processor : Intel® CORE™ i3 CPU M370 @2.40 GHz (2.40 GHz). RAM : 2.00 GB. Harddisk dengan kapasitas 500 GB.

2. Software

Kebutuhan perangkat lunak (software) yang digunakan :

- a. Sistem Operasi Windows 10
- b. Aplikasi Rapidminer 5.3

4. Data

Data yang akan digunakan dalam penelitian ini adalah data siswa alumni SMK Islam Al-Barokah .

3.4 Teknik Pengumpulan Data

Didalam penelitian ini, peneliti menggunakan beberapa metode yang akan digunakan untuk melakukan penelitian yang berkaitan dengan pengumpulan data Berikut adalah beberapa metode yang digunakan;

1. Observasi

Teknik pengumpulan data dengan mengadakan pengamatan dan penelitian secara langsung (Observasi) terhadap obyek yang diteliti melalui wawancara dengan staf sekolahan untuk mendapatkan informasi berdasarkan data yang di butuhkan khususnya data siswa SMK Islam Al-Barokah Poncowarno. Selain itu untuk merekam berbagai fenomena (situasi dan kondisi) yang terjadi mengenai pendidikan di SMK Islam Al-Barokah Poncowarno.

2. Dokumentasi

Metode ini juga berfungsi sebagai pengumpulan data-data pendukung seperti gambar-gambar atau dokumen yang lainnya sebagai penjelasan tentang data-data yang sudah diperoleh dari hasil pengamatan dan penelitian.

3. Tinjauan Pustaka (*Research Library*)

Tinjauan pustaka dilakukan dengan cara membaca, mengutip dan membuat catatan yang bersumber pada bahan-bahan pustaka yang mendukung dan berkaitan dengan penelitian dalam hal ini mengenai data mining Naïve Bayes, K-NN.

3.5 Teknik Analisis Data

Penelitian ini menggunakan model standarisasi data mining yaitu *Cross Industry Standart Process for Data Mining* (CRISP-DM) yang digambarkan pada skema tahap penelitian di atas dengan langkah-langkah sebagai berikut:

1. Fase Pemahaman Bisnis (*Bussiness Understanding Phase*)

Pada tahap ini berfokus pada tujuan penelitian yaitu untuk mengetahui algoritma terbaik untuk klasifikasi waktu tunggu alumni mendapatkan pekerjaan, dengan menerjemahkan data siswa alumni sehingga didapatkan model yang terbaik untuk memenuhi dari tujuan penelitian. Pada penelitian ini model yang digunakan adalah Naïve bayes dan KNN.

2. Fase Pemahaman Data (*Data Understanding Phase*)

Data diperoleh dari SMK Islam Al-Barokah dengan jumlah data 400 alumni, data berasal dari *tracer* alumni pada SMK Islam Al-Barokah yang didapat dari data arsip sekolah, setelah dilakukan analisis ternyata data *tracer* study yang ada belum tersinkron dengan data dapodik, data yang di dapat dari arsip sekolah sangat sederhana, sehingga atribut yang digunakan pada penelitian ini dapat dilihat pada tabel 3.1.

Tabel 3.1 Atribut Penelitian

Kriteria	Keterangan
Jurusan	Jurusan Siswa
Ekskul	Organisasi yang diikuti oleh siswa
Nilai Rata-rata	Nilai Rata-rata Siswa
Jenis Kelamin	Jenis Kelamin
Waktu	Lama mendapat pekerjaan (Bulan)

Atribut penelitian tersebut diperoleh berdasarkan pada data tracer alumni pada SMK Islam Al-Barokah . berikut adalah contoh dari data *tracer* alumni pada SMK Islam Al-Barokah dapat dilihat pada tabel 3.2.

Tabel 3.2 Contoh Data Hasil *Tracer* Alumni SMK Islam Al-Barokah

NIS	Nama	Jurusan	Jk	Ekskul	Nilai	Waktu Tunggu (Bulan)
4464	Nanda Khoirul Rizki	TKJ	L	Ikut	80	19
4571	Bagus Lutfi Maulana Aziz	AP	L	Ikut	76	6
4574	Lusiana	TKJ	P	Tidak Ikut	74	18
4576	Said Maulana	TKJ	L	Tidak Ikut	74	17
4577	Ani Wijaya	TKJ	P	Tidak Ikut	73	19
4578	Diah Anjarwati	AP	P	Tidak Ikut	76	11
4583	Intan Dewita Rani	TKJ	P	Ikut	74	21
4584	Listianingsih	AP	P	Ikut	73	7
4460	Naila Tuzulfa	TKJ	P	Ikut	77	15
4469	Nazwa Ayu Gunawan	TKJ	P	Ikut	78	18

Pada tabel 3.2 merupakan sampel data alumni, untuk data lengkap dapat dilihat di lampiran, data alumni yang digunakan adalah data 5 tahun terakhir dimana data yang terkumpul berjumlah 400 data *tracer* alumni.

3. Data Intergration dan Data Transformasi

Untuk meningkatkan dan memudahkan dalam proses analisis maka dari 5 Kriteria yang diperoleh akan dipilih beberapa kriteria inti yang digunakan dan menjadi variabel pada penelitian ini. Penelitian ini terdiri dari variabel terikat (Y) dengan tiga kategori dan variabel bebas (X) dengan pengelompokan variabelnya berdasarkan data yang diperoleh dari *tracer* alumni SMK Islam Al-Barokah dapat dilihat pada tabel 3.3 berikut ini.

Tabel 3.3 Variabel Penelitian

Variabel Penelitian	Keterangan	Kategori
Jurusan (X_1)	Jurusan yang asal alumni SMK Islam Al-Barokah	<ul style="list-style-type: none"> - 1=Teknik Komputer Jaringan (TKJ) - 2=Administrasi Perkantoran (AP)
Ekstrakurikuler (X_3)	Status alumni pernah mengikuti kegiatan Ekstrakurikuler	<ul style="list-style-type: none"> - 1= Mengikuti Ekskul - 2= Tidak Mengikuti Ekskul
Nilai (X_4)	Nilai rata-rata pada ijazah alumni SMK Islam Al-Barokah	<ul style="list-style-type: none"> - 1= ≤ 70(Cukup) - 2= $> 70 \& < 85$(Memuaskan) - 3= ≥ 85 (Dengan Pujian)
Jenis Kelamin (X_2)	Jenis Kelamin alumni SMK Islam Al-Barokah	<ul style="list-style-type: none"> - 1=Laki-laki - 2=Perempuan
Lama Mendapat Pekerjaan (Y)	Status Lama mendapatkan pekerjaan alumni SMK Islam Al-Barokah	<ul style="list-style-type: none"> - 1= (< 12 bulan) (Cepat) - 2= (≥ 12 bulan atau ≤ 18 bulan) (Sedang) - 3= > 18 bulan (Lama)

4. Pemodelan (Modeling Phase)

Algoritma yang digunakan dalam penelitian ini yaitu algoritma Naïve Bayes dan K-Nearest Neighbor (KNN), untuk mengklasifikasikan dalam memperkirakan waktu tunggu siswa memperoleh pekerjaan. Setelah diketahui hasil dari akurasi tiap algoritma kemudian dibandingkan untuk mencari algoritma terbaik.

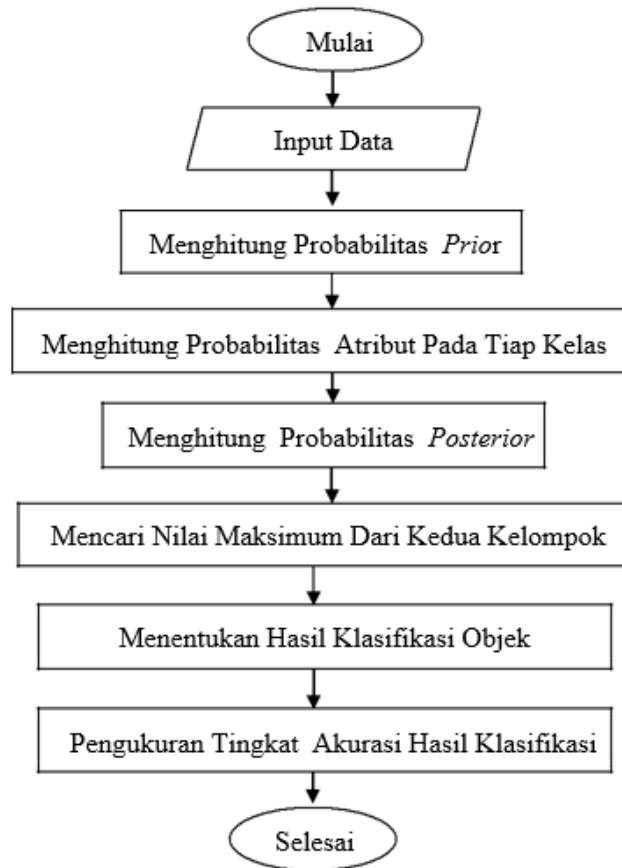
a. Metode Naïve Bayes

Langkah-langkah yang dilakukan dalam mengolah data pengklasifikasian dengan metode *naive* Bayes menggunakan bantuan software *Rapid miner* dalam penelitian ini adalah sebagai berikut:

1. Input Data.
2. Menghitung nilai probabilitas awal (prior) berdasarkan Persamaan (2.4).
3. Menghitung probabilitas setiap variabel bebas pada setiap kelompok berdasarkan Persamaan (2.6).
4. Menghitung probabilitas akhir (posterior) pada masing-masing kelas berdasarkan persamaan (2.7).
5. Mencari nilai maksimum pada kedua kelompok dan menentukan hasil klasifikasi obyek.
6. Mengevaluasi hasil klasifikasi dengan menghitung nilai *precision* dan *recall*.

Tahapan metode *naive* Bayes pada penelitian ini dimulai dari input data, lalu kemudian menghitung probabilitas *prior*, lalu kemudian probabilitas atribut setiap kelas, kemudian menghitung probabilitas *posterior*, lalu mencari nilai maksimum minimum dari prior dan posterior, dan akurasi, untuk lebih jelasnya dapat dilihat pada Gambar 3.2 berikut ini.

Berikut ini adalah Tahapan metode *naive* Bayes .



Gambar 3.2 Diagram Alir Metode *Naive* Bayes

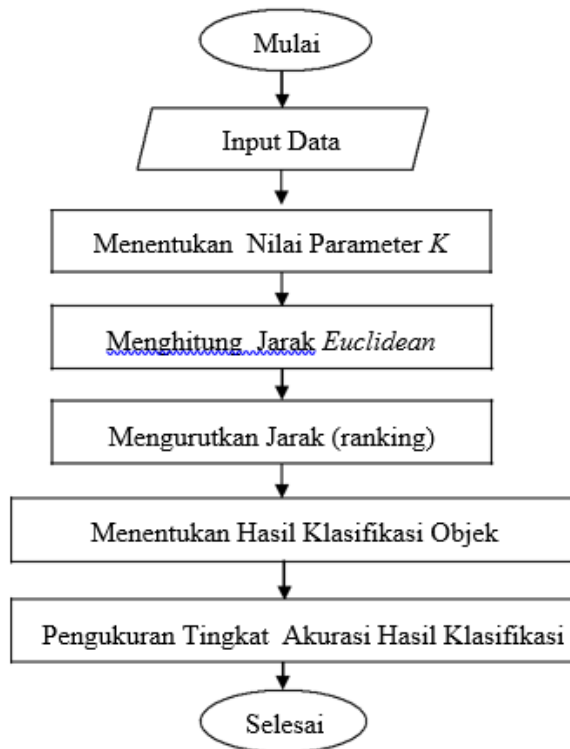
b. Metode *K-Nearest Neighbors* (K-NN)

Langkah-langkah yang dilakukan dalam mengolah data pengklasifikasian dengan metode K-NN adalah sebagai berikut:

1. Input Data.
2. Menentukan nilai parameter K .
3. Menghitung jarak Euclidean objek terhadap data training yang diberikan berdasarkan Persamaan (2.8)
4. Mengurutkan data yang mempunyai jarak terkecil hingga terbesar.

5. Menentukan hasil klasifikasi dengan menggunakan kategori *nearest neighbor* yang paling banyak.
6. Mengevaluasi hasil klasifikasi dengan menghitung nilai *precision* dan *recall*.

Langkah – langkah *K-Nearest Neighbors* (K-NN) dapat dilihat pada gambar 3.3.



Gambar 3.3 Diagram Alir Metode *K-NN*

5. Fase Evaluasi (*Evaluation Phase*)

Pada fase evaluasi antara algoritma Naïve Bayes dan K-NN digunakan metode evaluasi *precision recall* , dimana hasil akurasi dari setiap metode kemudian dibandingkan dan dicari hasil yang terbaik dengan menggunakan Persamaan 2.9.