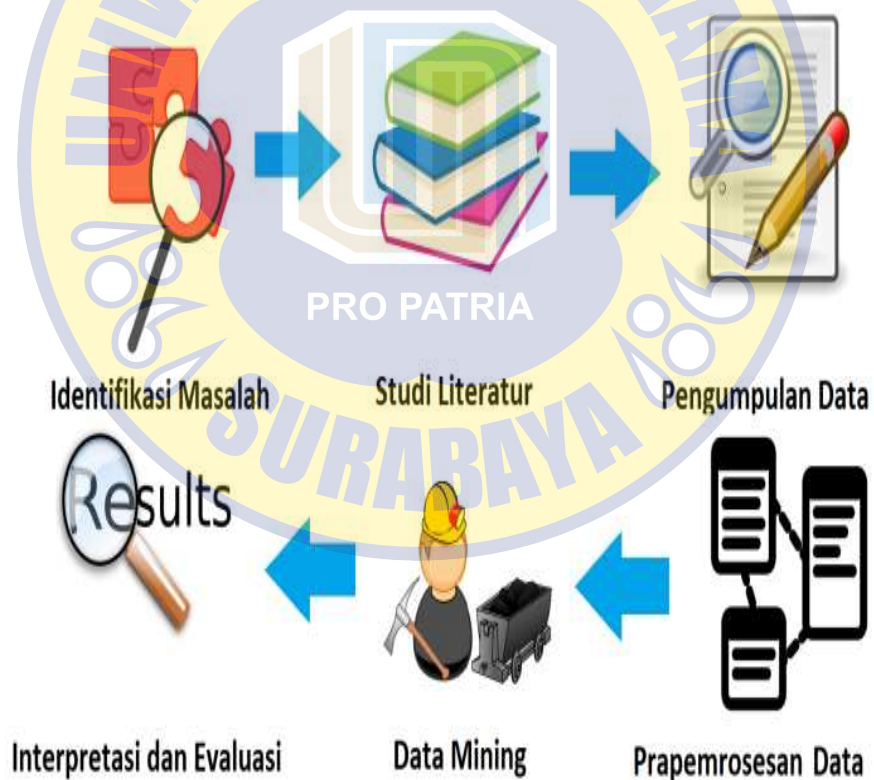


BAB III

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

merupakan gambar yang mewakili tahap-tahap sebuah proses secara terurut terpisah sehingga proses tersebut menjadi lebih sederhana agar mudah dipahami. Berikut ini adalah diagram alir yang digunakan pada penelitian ini.



Gambar 3.1 Tahapan penelitian

3.2 identifikasi masalah

Universitas Narotama merupakan salah satu perguruan tinggi yang ada di Surabaya. Tetapi banyak masyarakat umum yang tahu tentang dan dimana universitas tersebut berada, maka dari itu dibutuhkan strategi marketing yang baik agar masyarakat umum mengetahui prodi dan fakultas apa saja yang ada di Universitas Narotama Surabaya. , langkah awal yang dilakukan adalah melakukan studi literatur, observasi

3.3 Studi Literatur

Tahap ini mencari referensi teori yang relevan pada kasus pengendalian persediaan dari jurnal, artikel, dan buku. Tujuan dari referensi tersebut akan digunakan sebagai dasar teori untuk menyelesaikan permasalahan pada penelitian ini.

3.3 Pengumpulan Data

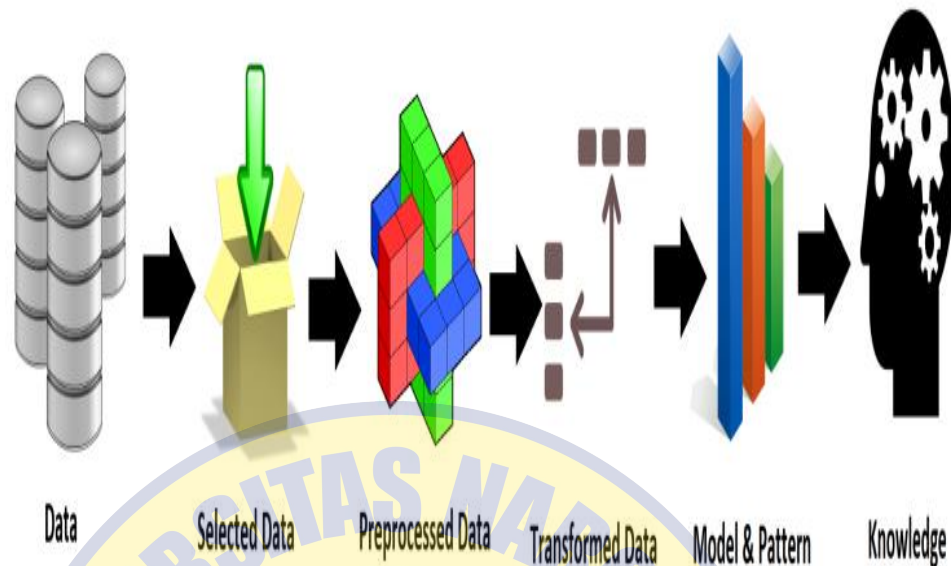
pengumpulan data berupa suatu peranan tentang sifat, keadaan, kegiatan tertentu dan sejenisnya. Pengumpulan data dilakukan untuk mendapatkan suatu informasi yang dibutuhkan untuk mencari tujuan penelitian. Dalam penyusunan tugas akhir ini penulis mengambil objek penelitian pada Bagian Adminitrsi dan Registrasi. Pengumpulan data dalam penelitian di Universitas Narotama. Berdasarkan dari pengamatan pada tahun 2015 sampai 2018 data pendaftaran mahasiswa baru di Universitas Narotama mencapai 2681 mahasiswa dari semua

prodi S1. Data yang berisi nama mahasiswa, nim, alamat rumah, dan alamat sekolah kemudian akan melalui pengolahan data hingga menghasilkan dataset yang siap diklasifikasi.

Adapun jenis data yang kami gunakan dalam penelitian ini adalah Data Primer yaitu data yang diperoleh langsung dari pihak kepala BAR. Dalam hal ini diperoleh langsung dari kepala BAR. Dalam hal ini diperoleh data pendaftaran mahasiswa baru dari tahun 2015 sampai 2018 dan menganalisa jumlah pendaftar.

3.4 Prapemrosesan

Pada tahap ini melakukan perubahan data dikarenakan data tersebut masih belum sempurna seperti data tidak lengkap dan tidak konsisten. Tujuan utama dalam tahap prapemrosesan data adalah melakukan pembersihan, penambahan, pengurangan dan penyusunan data menjadi terstruktur sesuai dengan aturan pada proses *mining* yang akan dijalankan.



Gambar 3.1 pemodelan

- 1) *Selected Data*: untuk meningkatkan akurasi dan kualitas hasil data mining, data diambil dari pendaftaran mahasiswa baru antara tahun 2015-2018 sebanyak 2681
- 2) *Preprocessed Data*: pada tahap ini melakukan pembersihan data karena data yang diperoleh dari perusahaan umumnya masih kotor seperti data yang kurang lengkap dan atribut yang tidak relevan, data yang masih kotor menyebabkan hasil yang kurang akurat dan terjadinya eror saat pemrosesan data sehingga pada proses ini perlu dilakukan pembersihan data. Pada proses ini peneliti menghapus data yang tidak lengkap yaitu produk yang tidak bersinambung dikarenakan produk tersebut sudah tidak di produksi dan varibel serta kolom yang tidak memiliki nilai.

3) *Transformation*: tahap ini melakukan perubahan data, data yang didapatkan dari perusahaan masih menggunakan format yang tidak sesuai, karena pada saat pemrosesan data hanya bisa menerima input data kategorikal dan tipe data numerik sehingga data yang digunakan harus berbentuk format excel dengan ekstensi CSV.

4) *Model & Patterns*: tahap ini merupakan proses pembentukan pola yaitu mengolah data menjadi kelompok dengan menggunakan metode klasterisasi. Proses ini bertujuan untuk menentukan kategori produk yang diminati dengan menggunakan algoritma *K-Means*[11].

Algoritma *K-Means* merupakan salah satu metode klasterisasi yang berbasis partisi yaitu metode yang digunakan untuk memisah data menjadi kelompok, kelompok yang memiliki kemiripan yang sama di masukan dalam kelompok yang sama dan kelompok yang berbeda di masukan dalam kelompok yang berbeda, untuk memisahkan kelompok tersebut metode *K-Means* membagi data dengan mengelompokkan objek berdasarkan nilai k yang telah ditentukan. Tujuan membagi data adalah untuk mengetahui kelompok produk yang diminati dan kurang diminati[12]. Beberapa tahapan metode *K-Means* sebagai berikut dalam bentuk *pseudocode*.

Input: D (Data), K (jumlah klaster)

- 1) Tentukan jumlah kelompok (K)
- 2) Inisialisasi nilai pusat secara acak

Process:

- 3) Hitung setiap titik data dalam D dengan nilai pusat terdekat. Data yang telah dihitung akan menjadi kelompok K.
- 4) Hitung ulang posisi nilai pusat.

Repeat:

Ulangi langkah *process* sampai tidak ada lagi perubahan keanggotaan titik data.

Output:

Data yang sudah dikelompokkan berdasarkan perhitungan jarak minimum. Untuk menghitung setiap objek data dengan nilai pusat menggunakan rumus Euclidean Distance sebagai berikut:

$$d(x, y) = \sqrt{\sum_{j=1}^n (x_j - y_j)^2} \quad (1)$$

d = jarak antara x dan y

x = nilai pusat

y = data pada atribut

j = setiap data

n = jumlah data

x_j = nilai pusat ke j

y_j = objek data ke j

- 5) *Knowledge:* Pada tahap ini mengetahui dari hasil penggalian data. Hasil tersebut digunakan sebagai sumber pengetahuan untuk digunakan pada tujuan

penelitian ini dan mempresentasikan pengetahuan dalam bentuk yang mudah dipahami.

Setelah data data berhasil dikumpulkan, langkah selanjutnya adalah proses pengolahan. Peneliti melakukan pemeriksaan terhadap jawaban – jawaban informan dan survei yang telah dilakukan dari data wawancara dan observasi yang telah dilakukan oleh peneliti. Tujuannya adalah untuk kerapian data dan selanjutnya melakukan perbaikan kata dan kalimat.

3.5 Transformasi data

Data yang berjenis nominal seperti kota asal dan program studi harus dilakukan proses inialisasi data terlebih dahulu ke dalam bentuk angka atau numerikal. Untuk melakukan inialisasi kota asal dapat dilakukan dengan:

- a. Pada kota asal mahasiswa terlebih dahulu dilakukan pembagian wilayah-wilayah menjadi beberapa bagian wilayah.

Tabel 3.2 Inialisasi wilayah

	Frekuensi	Inisial
Kota Asal		
Surabaya	1852	1
Sidoarjo	223	2
Gresik	73	3

Ponorogo	15	5
Trenggalek	8	6
Tulungagung	2	7
blitar	11	8
kediri	34	9
malang	11	10
Lumajang	3	11
Jember	5	12
Banyuwangi	6	13
Bondowoso	3	14
Situbondo	2	15
probolinggo	2	16
pasuruan	10	17
mojokerto	30	18
Jombang	27	19
Nganjuk	16	20
madiun	14	21

Magetan	7	22
Ngawi	4	23
Bojonegoro	30	24
Tuban	19	25
Lamongan	30	26
Bangkalan	17	27
Sampang	3	28
Pamekasan	2	29
Sumenep	1	30
jateng	29	31
jabar	7	32
jakarta	4	33
luar jawa	63	34
luar negri	121	35

- b. Kemudian pada program studi juga dilakukan inisialisasi data terlebih dahulu ke dalam bentuk angka atau numerikal.

Tabel 3.2 Inisialisasi Program Study

Program Studi	Frekuensi	Inisial
Akuntansi	335	1
Manajemen	614	2
Hukum	385	3
Teknik Sipil	563	4
Sistem Komputer	158	5
Sistem Informasi	327	6
Teknik Informatika	300	7

- c. Dan Yang Terakhir adalah Inisialisasi umur dan diubah kedalam bentuk angka atau numerikal.

Tabel 3.3 Inisialisasi Umur

Umur	Frekuensi	Inisial
1998-0000 (21-00)	670	1
1994-1997 (22-25)	1267	2
0000-1993 (00-26)	744	3

3.5 Evaluasi

Setelah melakukan pengujian data penulis melakukan evaluasi berdasarkan akurasi dan objektivitas informasi data segmentasi mahasiswa dengan unsupervised learning sebagai prediksi penerimaan mahasiswa.

3.6 Membuat Laporan

Setelah penelitian selesai dilakukan, tahap akhir adalah melaporkan hasil serta temuan yang di dapat dan mendokumentasikannya ke dalam laporan ilmiah. Penulisan dilakukan berdasarkan format penulisan yang telah ditentukan, serta menggunakan bahasa yang baik dan mudah dimengerti. Laporan hasil penelitian kelak dapat dimanfaatkan bagi penelitian selanjutnya sebagai referensi dan sumber informasi terkait dengan topik yang relevan.

