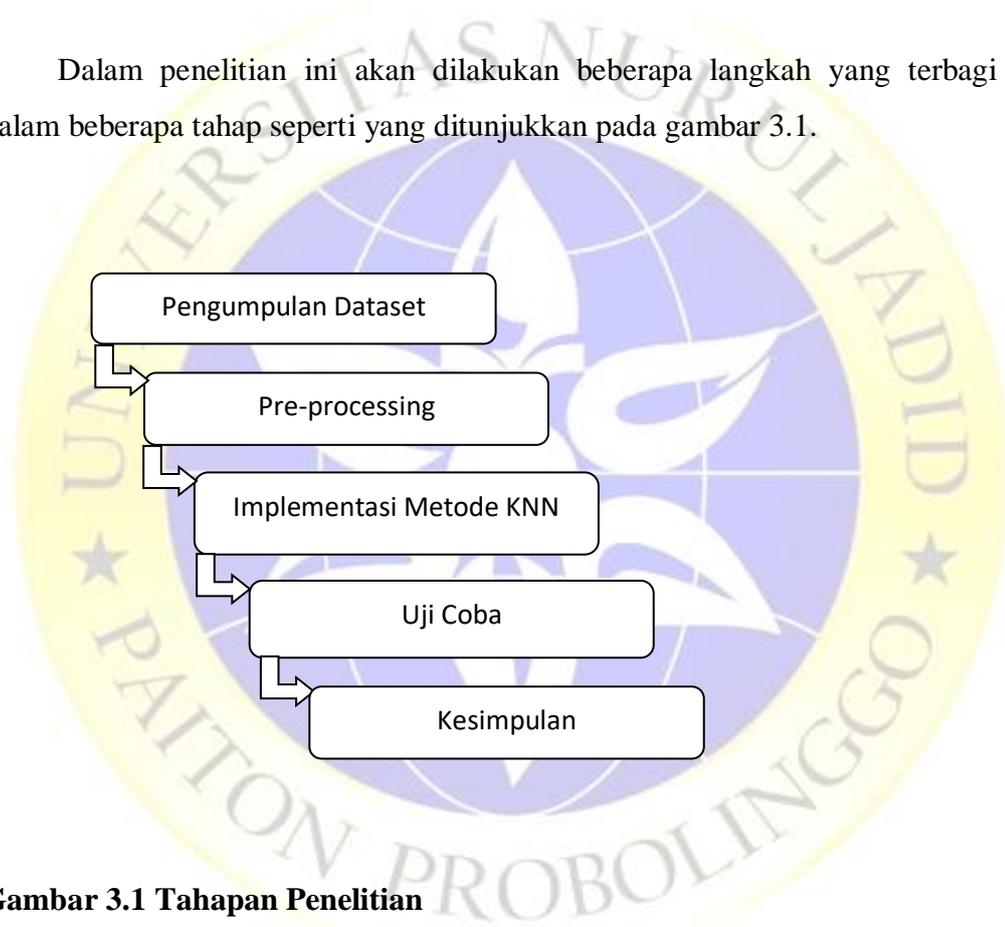


BAB III METODE PENELITIAN

3.1. Kerangka Penelitian

Penelitian ini menggunakan pendekatan kuantitatif karena dalam penelitian ini dilakukan proses perhitungan matematis untuk mendapatkan hasil dari sebuah penelitian.

Dalam penelitian ini akan dilakukan beberapa langkah yang terbagi ke dalam beberapa tahap seperti yang ditunjukkan pada gambar 3.1.



Gambar 3.1 Tahapan Penelitian

3.2. Metode Analisis

Pada penelitian ini, kami menggunakan metode penelitian dengan algoritma *K-Nearest Neighbor* (KNN), bahasa pemrograman, tools dan software yang digunakan dalam penelitian ini antara lain:

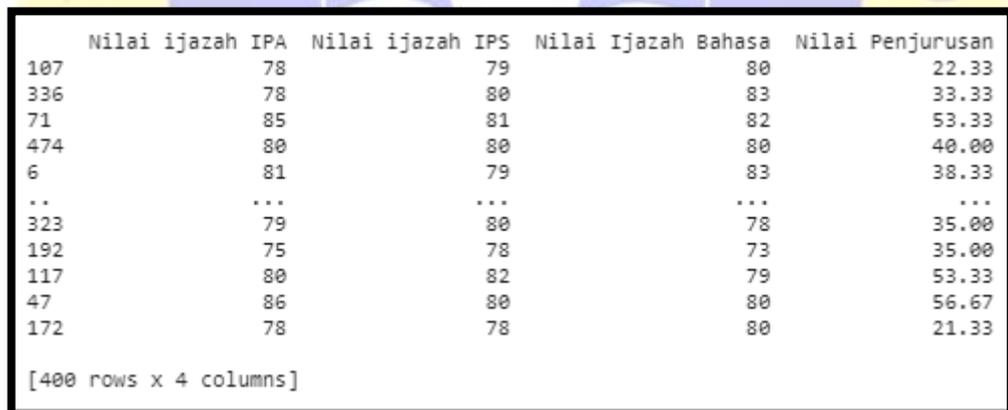
1. Python versi 3
2. Google Colab
3. Browser Google Chrome

Pengumpulan Data

Teknik pengumpulan data merupakan faktor penting demi keberhasilan penelitian. Data yang digunakan dalam penelitian ini diperoleh dari website sekolah Man 1 Probolinggo. Data siswa yang diambil mulai dari angkatan tahun 2018 sampai 2020.

3.3. Pre Processing

Tahap preprocessing dapat digunakan untuk menyiapkan data jurusan Man 1 Probolinggo agar dapat digunakan secara optimal. Sedangkan hasil dari preprocessing ini dihasilkan 500 data yang dapat diproses lebih lanjut dengan berbagai atribut yaitu Nilai Ijazah IPA, Nilai Ijazah IPS, Nilai Ijazah Bahasa, Nilai Penjurusan, dan Penjurusan. Dimana untuk data trainingnya ada 400, sedangkan untuk data testingnya 100. Di gambar 3.2 ialah data training.



	Nilai ijazah IPA	Nilai ijazah IPS	Nilai Ijazah Bahasa	Nilai Penjurusan
107	78	79	80	22.33
336	78	80	83	33.33
71	85	81	82	53.33
474	80	80	80	40.00
6	81	79	83	38.33
..
323	79	80	78	35.00
192	75	78	73	35.00
117	80	82	79	53.33
47	86	80	80	56.67
172	78	78	80	21.33

[400 rows x 4 columns]

Gambar 3.2 Data Training

Pada gambar 3.3 ialah banyaknya data testing.

	Nilai ijazah IPA	Nilai ijazah IPS	Nilai Ijazah Bahasa	Nilai Penjurusan
90	85	80	80	51.33
254	80	80	83	50.00
283	80	79	80	51.67
445	80	81	79	30.00
461	85	80	83	51.67
..
372	80	80	80	28.00
56	80	81	80	41.67
440	82	79	80	40.00
60	84	80	82	63.33
208	81	81	83	56.67

[100 rows x 4 columns]

Gambar 3.3 Data Testing

Keterangan pada variable Y ialah label atau nominal, dimana 1 adalah jurusan Bahasa, 2 adalah jurusan IPA, dan 3 adalah jurusan IPS. Lebih jelasnya pada tabel 3.1.

Tabel 3.1 Variabel yang digunakan

Nama Atribut	Variabel X	Keterangan
P/L	X1	String
Nilai Ijazah IPA	X2	Numerik
Nilai Ijazah IPS	X3	Numerik
Nilai Ijazah Bahasa	X4	Numerik
Nilai Penjurusan	X5	Numerik
Jurusan/Label	Y	Nominal(1, 2, 3)

3.4. Klasifikasi KNN

K-Nearest Neighbor (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran (neighbor) yang jaraknya paling dekat dengan objek tersebut. Setelah data terkumpul diperlukan sistem klasifikasi seperti suatu sistem yang dapat menemukan informasi. Metode KNN dibagi menjadi dua fase yaitu pembelajaran (praktik) dan klasifikasi atau tes (uji). Selama fase pelatihan, algoritma ini hanya menyimpan vektor fitur dan mengklasifikasikan data pelatihan. Pada tahap klasifikasi, dilakukan perhitungan kesamaan karakteristik data yang diuji (klasifikasi tidak diketahui). Jarak dari vektor baru ini ke semua vektor data pelatihan dihitung dan k tetangga terdekat diperoleh.

3.5. Pengujian

Setelah mendapatkan model dari proses pelatihan, maka akan melakukan pengujian model untuk mengevaluasi data yang dibuat untuk mengidentifikasi kumpulan data. Proses pengujian model ini adalah proses penginputan data siswa di Man 1 Probolinggo dari hasil pembelajaran(training) selanjutnya dilakukan pengujian(testing) menggunakan metode K-Nearest Neighbor(KNN).