

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1. Penelitian Terkait**

Penelitian terkait merujuk pada studi yang digunakan oleh penulis sebagai landasan acuan, sumber data pendukung, dan referensi untuk membandingkan persamaan serta perbedaan dengan penelitian yang akan dilakukan. Dalam bagian ini, akan diuraikan beberapa penelitian terkait yang memiliki keterkaitan dengan penelitian yang akan dilakukan. Penelitian sebelumnya, yang pertama dengan judul “Model Berbasis *K-Nearest Neighbor* Dalam Penentuan Penerima Beasiswa”, penelitian ini dilakukan di STIKes Aisyah Pringsewu dengan nama beasiswa BBM. Saat ini, seleksi penerima beasiswa BBM di perguruan tinggi STIKes Aisyah Pringsewu masih dilakukan secara manual, terutama berdasarkan nilai IPK saja. Dalam mengevaluasi jumlah mahasiswa yang besar merupakan tantangan bagi perguruan tinggi. Proses penilaian dan keputusan seringkali lebih bersifat subjektif, yang dapat menyebabkan kesalahan dalam menentukan mahasiswa yang layak mendapatkan beasiswa BBM. Studi ini membuktikan bahwa model yang diterapkan sangat efektif dalam membantu perguruan tinggi STIKes Aisyah Pringsewu dalam memilih mahasiswa penerima beasiswa BBM dengan tingkat keberhasilan mencapai 80% [9].

Kedua, penelitian dengan judul “Aplikasi Penentuan Penerima Beasiswa Menggunakan Algoritma C4.5”, dalam penelitian ini dilatar belakangi oleh Sistem penilaian beasiswa di Universitas Bumigora, seleksi penerima beasiswa bergantung pada empat kriteria pokok: jumlah sks, pendapatan orang tua, IPK, dan jumlah tanggungan keluarga. Tingginya jumlah pelamar beasiswa, seperti pada tahun 2019 dengan 110 pelamar, membuat proses pengambilan keputusan sulit dan memerlukan kecepatan serta ketepatan. Diperlukan suatu sistem yang bisa memprediksi calon penerima beasiswa guna mengurangi tingkat subjektivitas dalam pengambilan keputusan dan meningkatkan efisiensi dalam menentukan siapa yang layak mendapatkan beasiswa. Metode penelitian ini menggunakan algoritma C4.5. Berdasarkan

perencanaan, pembangunan, dan penerapan algoritma *C4.5* untuk menentukan penerima beasiswa, dapat disimpulkan bahwa akurasi yang dicapai melalui metode penanganan dengan menghilangkan nilai yang hilang secara keseluruhan (*Listwise Deletion*) dalam mengklasifikasikan penerima beasiswa mencapai 92%. *Spesifisitas* mencapai 92,3%, dan *sensitivitas* mencapai 91,6%. Sementara itu, menggunakan pendekatan penggantian nilai yang hilang dengan nilai rata-rata (*Mean Substitution*), akurasinya mencapai 88%. *Spesifisitas* mencapai 75%, dan *sensitivitas* mencapai 100%. Dalam kedua pendekatan penanganan, data *training* sebanyak 100 data (80%) dan data uji sebanyak 25 data (20%) digunakan. Berdasarkan hasil tersebut, metode *Listwise Deletion* terbukti lebih efektif untuk mengklasifikasikan penerima beasiswa. Algoritma *C4.5* berhasil diimplementasikan dengan sukses dalam proses menentukan calon penerima beasiswa[10].

Ketiga, penelitian dengan judul “Sistem Pendukung Keputusan Pemilihan Penerima Beasiswa Universitas Muhammadiyah Pontianak Menggunakan Metode *Fuzzy Mamdani*”. Studi ini menunjukkan bahwa proses seleksi penerima beasiswa PPA di Universitas Muhammadiyah Pontianak memakan waktu yang cukup lama karena jumlah pelamar yang tinggi. Panitia seleksi mengalami tantangan karena jumlah pelamar yang banyak, sedangkan kuota penerimaan terbatas. Disamping itu, metode penilaian yang digunakan sulit untuk memberikan hasil yang mendalam dan tepat dari setiap mahasiswa yang mengajukan permohonan. Studi ini berfokus pada pengembangan prototipe aplikasi berbasis *website* menggunakan metode *fuzzy mamdani* untuk proses pemilihan penerima beasiswa PPA di Universitas Muhammadiyah Pontianak. Diharapkan sistem yang dikembangkan dapat membantu panitia seleksi mengurangi kesalahan analisis dan penilaian, serta memastikan bahwa mahasiswa yang berhak menerima beasiswa PPA dapat dipilih dengan tepat. Hasil pengujian sistem menunjukkan bahwa aplikasi ini dapat digunakan sesuai dengan kebutuhan kampus dalam seleksi penerimaan beasiswa PPA. Proses ini melibatkan empat tahapan yang telah ditentukan dalam *logika fuzzy mamdani* dan memberikan perhitungan yang akurat. Pengujian dengan *Root Mean Square Error* menunjukkan nilai *error* sebesar 0,315, sementara

pengujian dengan *Mean Absolute Percentage Error* menunjukkan nilai *error* sebesar 1,06% [11].

Keempat, penelitian yang berjudul “Implementasi Algoritma *K-Means* dalam Pengklasteran untuk Rekomendasi Penerima Beasiswa PPA di UDINUS”. Pada penelitian ini dilatar belakangi oleh Masalah yang muncul alam seleksi penerima beasiswa PPA di UDINUS, salah satu kendala utamanya adalah kesulitan mengelola data dan atribut yang besar untuk menentukan penerima beasiswa. Akibatnya, beasiswa tersebut sering kali tidak diberikan kepada calon yang seharusnya memenuhi syarat. Dengan kata lain, mahasiswa yang seharusnya berhak menerima beasiswa tidak mendapatkannya, sementara mahasiswa yang seharusnya tidak memenuhi kriteria justru mendapatkan beasiswa. Tentu saja, Udinus berharap agar proses seleksi penerimaan beasiswa PPA dapat berjalan tanpa hambatan dan sesuai dengan aturan yang telah ditetapkan. Hasil penelitian menunjukkan bahwa dengan menggunakan metode *clustering* dan algoritma *K-Means* yang melibatkan 7 atribut dan *dataset* yang cukup besar, sistem ini mampu memberikan rekomendasi calon penerima Beasiswa Peningkatan Prestasi Akademik (PPA). Pentingnya menentukan *centroid* (titik pusat) terlihat dalam hasil pengujian, di mana hasil *cluster* berbeda ketika menggunakan data uji dan keseluruhan data. Saat diuji dengan 441 data, hasilnya menunjukkan bahwa dari 154 mahasiswa direkomendasikan sebagai calon penerima Beasiswa Peningkatan Prestasi Akademik (PPA), sementara 287 mahasiswa lainnya tidak memenuhi syarat. Dari rekomendasi tersebut, 113 mahasiswa yang diterima pada tahun 2016 sudah memperoleh beasiswa dan evaluasi oleh biro kemahasiswaan menunjukkan bahwa rekomendasi tersebut memang tepat. Dengan menggunakan metode *cluster*, proses seleksi data dapat disempurnakan dan dipercepat, memudahkan serta mempercepat proses seleksi penerima beasiswa [12].

Kelima, penelitian dengan judul “Klasifikasi Penerima Beasiswa Aceh Carong (Aceh Pintar) Di Universitas Malikussaleh Menggunakan Algoritma Knn (*K-Nearest Neighbors*). Penelitian ini dilakukan di Universitas Malikussaleh. Saat ini, pemilihan penerima Beasiswa Aceh Carong di

Universitas Malikussaleh masih dilaksanakan secara manual, mengakibatkan lamanya proses tersebut dan menurunkan tingkat efisiensi.. Penelitian ini menghasilkan bahwa penggunaan teknik *K-Nearest Neighbor* untuk mengenali calon penerima Beasiswa Aceh Carong memberikan hasil yang memuaskan. Algoritma ini terbukti efektif dan efisien, dengan tingkat akurasi mencapai 82,00%. Hal ini menandakan bahwa algoritma *K-Nearest Neighbor* mampu mengklasifikasikan penerima beasiswa Aceh Carong dengan baik[12].

Perbandingan penelitian terdahulu disajikan dalam Tabel 2.1

Tabel 2.1 Pemetaan Penelitian Terdahulu

No .	Nama Peneliti	Judul	Metode Penelitian	Hasil Penelitian	Perbandingan
1.	Zulkifli, dan Ilham Ubaidillah ( 2019 )	Model Berbasis <i>K-Nearest Neighbor</i> Dalam Penentuan Penerima Beasiswa	Algoritma <i>K-Nearest Neighbor</i>	Studi ini membuktikan bahwa model yang telah dirancang berhasil membantu STIKes Aisyah Pringsewu dalam menyeleksi mahasiswa penerima beasiswa BBM. Tingkat ketepatan model ini mencapai 80%, menegaskan efektivitasnya dalam proses seleksi penerima beasiswa.	Persamaan penelitian ini menggunakan algoritma <i>K-Nearest Neighbor</i> sedangkan perbedaannya adalah penulis melakukan penelitian di SMK swasta guna memprediksi penerima beasiswa yayasan
2.	Abdurraghib Segaf Suweleh,	Aplikasi Penentuan Penerima	Algoritma C4.5	Penggunaan Algoritma C4.5 untuk	Persamaan penelitian ini yaitu meneliti

	Dyah Susilowati, dan Hairani (2020)	Beasiswa Menggunakan Algoritma C4.5 (2020)		mengklasifikasi calon penerima beasiswa telah sukses, dengan tingkat keakuratan mencapai 92% dan hasil uji sistem yang dapat dipercaya.	calon penerima beasiswa. Penelitian ini menggunakan Algoritma C4.5, sedangkan penulis menggunakan algoritma <i>K-Nearest Neighbor</i> .
3.	Abu Salam, Diyan Adiatma, dan Junta Zeniarja (2020)	Implementasi Algoritma <i>K-Means</i> dalam Pengklasteran untuk Rekomendasi Penerima Beasiswa PPA di UDINUS	Algoritma <i>K-Means</i>	Dengan menggunakan seleksi atribut, algoritma <i>k-means</i> menghitung dan menempatkan setiap data ke dalam <i>cluster</i> yang telah ditetapkan. Hasilnya, sebanyak 154 mahasiswa direkomendasikan untuk mendapatkan beasiswa PPA, sementara 287 mahasiswa lainnya tidak memenuhi kriteria	Persamaan penelitian ini meneliti beasiswa. Penelitian ini menggunakan Algoritma <i>K-Means</i> , sedangkan penulis menggunakan algoritma <i>K-Nearest Neighbor</i> .

				untuk mendapatkannya.	
4.	Ilham Saputra, Syarifah Putri Agustini Alkadri, dan Rachmat Wahid Saleh Insani (2021)	Sistem Pendukung Keputusan Pemilihan Penerima Beasiswa Universitas Muhammadiyah Pontianak Menggunakan Metode Fuzzy Mamdani	Algoritma Fuzzy Mamdani	Hasil uji coba sistem menunjukkan bahwa sistem ini dapat diimplementasikan sesuai dengan kebutuhan universitas dalam seleksi penerima beasiswa PPA terpenuhi melalui penggunaan sistem yang melibatkan empat tahap logika fuzzy Mamdani, yang memberikan perhitungan yang sangat akurat. Keakuratan sistem ini dibuktikan melalui pengujian <i>Root Mean Square Error</i> dengan nilai kesalahan sebesar 0,315. serta pengujian <i>Mean Absolute</i>	Persamaan penelitian ini adalah sama sama meneliti calon penerima beasiswa di instansi pendidikan. Sedangkan perbedaanya adalah penulis menggunakan algoritma Algoritma <i>K-Nearest Neighbor</i>

				<i>Percentage Error</i> dengan tingkat kesalahan sekitar 1,06%.	
5.	Ar Razi (2022 )	Klasifikasi Penerima Beasiswa Aceh Carong (Aceh Pintar) Di Universitas Malikussaleh Menggunakan Algoritma Knn ( <i>K-Nearest Neighbors</i> )	Algoritma <i>K-Nearest Neighbor</i>	Dari penelitian ini, ditemukan bahwa penerapan algoritma <i>KNN</i> dalam mengklasifikasi penerima Beasiswa Aceh Carong memberikan tingkat keakuratan sekitar 82,00%.	Persamaan penelitian ini menggunakan algoritma <i>K-Nearest Neighbor</i> sedangkan perbedaannya adalah penulis melakukan penelitian di SMK swasta guna memprediksi penerima beasiswa yayasan

## 2.2. Landasan Teori

### 2.2.1. Algoritma

Algoritma merujuk pada pendekatan sistematis yang mengikuti langkah-langkah tertentu untuk menyelesaikan masalah dan membuat keputusan. Ada beberapa pertimbangan yang harus diambil saat menggunakan algoritma. Pertama, algoritma harus akurat, artinya output yang dihasilkan harus sesuai dengan instruksi yang dimasukkan. Jika ada kesalahan dalam algoritma, hasilnya juga akan keliru. Kedua, penting untuk mempertimbangkan tingkat ketepatan algoritma, terutama dalam mencocokkan hasilnya dengan nilai aslinya. Ketiga, efisiensi algoritma juga perlu diperhatikan, baik dari segi kapasitas penyimpanan maupun waktu eksekusi.[13].

### 2.2.2. Data Mining

Data *Mining* adalah proses pemanfaatan satu atau lebih teknik pembelajaran mesin untuk secara otomatis menganalisis dan mengekstraksi

pengetahuan. Ini melibatkan pendekatan iteratif dan interaktif untuk menemukan pola atau model yang valid, berguna, dan dapat dimengerti dalam basis data yang sangat besar[14].

### 2.2.3. Teknik Data Mining

Ada berbagai teknik dalam data *mining*, dan masing-masing teknik memiliki fungsi dan karakteristiknya sendiri. Beberapa teknik ini meliputi klasifikasi, pengelompokan (*clustering*), penemuan aturan asosiasi (*association rule discovery*), dan regresi, serta deteksi *deviasi* (*deviation detection*). Metode yang termasuk dalam teknik klasifikasi meliputi *decision tree*, *K-Nearest Neighbor*, dan *Naïve Bayes*. Klasifikasi data melibatkan pencocokan data baru (data uji) dengan data latihan, dan hasil pencocokan ini akan menghasilkan keputusan yang sesuai dengan target yang telah diatur dalam data latihan.[2]. *K-Nearest Neighbor* banyak digunakan untuk klasifikasi data input dengan menyimpan hanya sebagian dari data pelatihan[15].

### 2.2.4. Algoritma *K-Nearest Neighbor*

*K-Nearest Neighbor* adalah suatu metode dalam data mining yang digunakan untuk mengklasifikasikan sekelompok data berdasarkan label yang diberikan pada data pembelajaran. *KNN* termasuk dalam kategori (*supervised learning*) yang memungkinkan klasifikasi objek baru berdasarkan objek terdekatnya. Instansi pertanyaan yang baru akan diklasifikasikan berdasarkan mayoritas kategori yang dimiliki oleh tetangganya pada *KNN*. Dalam hal ini, kelas yang paling sering muncul dianggap sebagai hasil klasifikasi. *KNN* juga merupakan pendekatan yang digunakan untuk mengukur sejauh mana kedekatan antara kasus-kasus lama dengan kasus baru. Yang membedakan *KNN* adalah pengklasifikasian tidak memerlukan pembentukan model melainkan hanya bergantung pada memori. Dengan menggunakan nilai mayoritas dari tetangga terdekat, *KNN* dapat meramalkan nilai sampel uji yang baru. [16]. Metode *K-Nearest Neighbor* adalah teknik yang digunakan untuk mengelompokkan objek baru berdasarkan jumlah tetangganya (*K*) yang terdekat. *KNN* termasuk dalam kategori algoritma (*supervised learning*), di mana hasil dari *instance* baru yang diajukan akan diklasifikasikan berdasarkan mayoritas kategori di antara tetangganya. Kelas yang paling sering muncul akan menjadi hasil klasifikasi.



Dalam menghitung algoritma *KNN*, terdapat lima langkah untuk menemukan tetangga terdekat, yaitu sebagai berikut:

- 1) Rumus *Hamming*
- 2) Rumus *Cosine*
- 3) Rumus *Manhattan*
- 4) Rumus *Euclidean*
- 5) Rumus *Correlation*

Dalam penelitian ini, penulis hanya mengaplikasikan metode perhitungan jarak menggunakan rumus *Euclidean*, yang dirumuskan sebagai berikut:

$$\sqrt{\sum_{i=1}^K (X_i - Y_i)^2} \quad (2.1)$$

$X_i$  adalah nilai yang terdapat dalam data pelatihan, sementara  $Y_i$  merupakan nilai dari data uji. Parameter  $K$  menunjukkan jumlah dimensi atribut yang dipertimbangkan dalam algoritma *K-NN*. Proses perhitungan algoritma *K-NN* melibatkan langkah-langkah berikut:

1. Memilih nilai  $k$  yang sesuai..
2. Melakukan perhitungan jarak euclidean kuadrat dari setiap objek dalam data latih terhadap instansi pertanyaan (*query instance*).
3. Selanjutnya, mengurutkan objek-objek tersebut ke dalam kelompok dengan jarak euclidean terkecil.
4. Menghimpun label kelas  $Y$  dari tetangga terdekat (*Nearest Neighbor*).
5. Dengan memanfaatkan mayoritas kategori tetangga terdekat, nilai dari *query instance* pertanyaan yang telah dihitung dapat diprediksi. yang telah dihitung[17].

### 2.2.5. Confusion Matrix

Dalam penelitian ini, evaluasi kinerja algoritma dilakukan menggunakan metode *Confusion matrix*, sebuah teknik yang umumnya digunakan untuk mengukur akurasi dalam konteks data *mining*. Rumus ini menghasilkan empat metrik, yaitu: *recall*, *precision*, *accuracy*, dan *error rate*. *Recall* mengukur proporsi kasus positif yang berhasil diidentifikasi dengan benar, sedangkan *precision* mengukur proporsi kasus positif yang diidentifikasi dengan tepat. *Accuracy* mencerminkan perbandingan kasus yang diidentifikasi dengan benar

terhadap total kasus, sementara *Error Rate* mengukur jumlah kasus yang salah diidentifikasi dibandingkan dengan total kasus. *Confusion matrix* membantu menganalisis sejauh mana *Classifier* dapat mengenali tupel dari kelas yang berbeda. Rumus yang digunakan untuk menghitung akurasi adalah sebagai berikut:

$$\text{Akurasi} = \frac{\text{true positif} + \text{true negatif}}{\text{true positif} + \text{true negatif} + \text{false positif} + \text{false negatif}} \times 100\% \quad (2,2)$$

### 2.2.6. Prediksi

Prediksi adalah upaya untuk memproyeksikan situasi di masa depan dengan menggunakan informasi dari periode sebelumnya [18]. Prediksi adalah evaluasi yang dilakukan secara obyektif dengan memanfaatkan informasi dari kejadian masa lalu untuk meramalkan kejadian di masa mendatang. [19].

### 2.2.7. Beasiswa

Beasiswa merupakan bentuk pengakuan untuk individu yang masih berada di tingkat pendidikan dasar atau perguruan tinggi, memungkinkan mereka melanjutkan pendidikan ke level yang lebih tinggi. Umumnya, beasiswa diberikan oleh yayasan atau pemerintah sebagai bentuk penghargaan terhadap pencapaian siswa atau mahasiswa, meskipun memiliki keterbatasan sumber daya finansial. [20]. Beasiswa adalah sumbangan keuangan yang diberikan kepada individu, mahasiswa, atau pelajar untuk mendukung kelancaran pendidikan mereka. Bantuan ini bisa disediakan oleh entitas seperti pemerintah, perusahaan, yayasan, atau lembaga lainnya. Menurut Kamus Besar Bahasa Indonesia, beasiswa adalah bantuan finansial yang diberikan kepada siswa atau mahasiswa untuk membantu membiayai pendidikan mereka. Pemerintah kota Sabang memberikan beasiswa tahunan secara teratur setiap tahunnya, bertujuan untuk memberi semangat kepada mahasiswa yang sedang menyelesaikan pendidikannya, baik itu di dalam atau di luar kota [21].

### 2.2.8. Anaconda

*Anaconda* adalah paket perangkat lunak gratis yang mencakup kotak peralatan yang dirancang khusus untuk penelitian dan sains. Saat menginstal Anaconda, memiliki akses ke berbagai lingkungan yang memungkinkan untuk

mengembangkan dengan *Python* atau R. Lingkungan ini, sering dikenal sebagai lingkungan pengembangan terintegrasi (IDE), adalah platform atau program yang membuat penulisan kode lebih mudah[22].

### **2.2.9. Python**

*Python* adalah bahasa pemrograman yang tergolong menengah dan berbasis objek, yang tidak hanya mudah dipelajari dan dimengerti, tetapi juga cukup fleksibel untuk menangani berbagai macam tugas. [22]. *Python* adalah salah satu bahasa pemrograman yang populer di kalangan para pengembang. Bahasa ini dikenal karena kemudahannya dalam penggunaan, yang membuatnya banyak digunakan oleh *programmer* berbeda dari bahasa pemrograman lain yang lebih kompleks. *Python* dikenal di kalangan pengembang karena keunggulannya yang simpel dibandingkan dengan bahasa pemrograman lainnya. Bahasa ini memiliki struktur *sintak* yang teratur dan mudah dipahami oleh pembuat program. Kelebihan *Python* terletak pada kemudahan penulisan pernyataan (*statements*) atau perintah, yang merupakan instruksi yang akan dieksekusi oleh komputer. Selain itu, *Python* juga memiliki aturan yang jelas dalam penulisan *string* yang merupakan kumpulan karakter atau huruf. Selain itu, dalam *Python* terdapat aturan yang jelas mengenai *variabel* dan tipe data. *Variabel* adalah wadah untuk menyimpan informasi dalam program, dan tipe data merujuk pada jenis data yang dapat disimpan dalam *variabel* tersebut. *Variabel* dalam *Python* bersifat *mutable* yang berarti nilainya dapat diubah selama program berjalan[23].

### **2.2.10. Library**

*Library Python* adalah sekumpulan modul yang terkait, berisi kode-kode yang dapat digunakan berkali-kali dalam berbagai program. Keberadaan *library* membuat proses pemrograman *Python* menjadi lebih mudah dan efisien bagi para pengembang, karena mereka tidak perlu menulis ulang kode yang serupa untuk setiap program yang mereka buat. Pertama dalam proses analisis menggunakan *Python*, langkah awal adalah menyiapkan perpustakaan (*library*) yang akan digunakan. Beberapa perpustakaan yang digunakan meliputi *numpy*, *pandas*, *matplotlib*, *seaborn*, dan *sklearn*. *Numpy* digunakan untuk alat analisis data, sementara *matplotlib* dan *seaborn* digunakan untuk visualisasi data. Selain

itu, *scikit-Learn* digunakan untuk implementasi *machine learning* dalam analisis ini.[24]. Sebelum mulai menulis kode *Python*, langkah pertama yang harus diambil adalah mengimpor *library* yang dibutuhkan. Berikut adalah contoh penggunaannya:

1. *Pylab* merupakan bagian dari *matplotlib* yang biasanya terinstal secara otomatis saat menginstal *matplotlib*.
2. *Seaborn* adalah perangkat pustaka *Python* yang bersifat *open-source* dan bermanfaat untuk membuat visualisasi data statistik dengan kualitas tampilan yang sangat baik.
3. *Pandas* merupakan perpustakaan *Python* yang terutama digunakan untuk analisis data, mencakup manipulasi, persiapan, dan pengelolaan data.
4. *Matplotlib.pyplot* merupakan bagian dari modul *matplotlib* yang memungkinkan modifikasi gambar, termasuk pembuatan grafik dan *plot*. *Matplotlib* adalah pustaka yang menyediakan fungsi-fungsi untuk melakukan modifikasi pada visualisasi gambar.
5. *Numpy*, yang merupakan singkatan dari *Numerical Python*, adalah perpustakaan yang menawarkan antarmuka efisien untuk menyimpan dan menjalankan operasi pada data yang padat melalui struktur *buffer*.
6. *Scikit-learn*, biasa disebut *sklearn*, adalah perpustakaan *Python* yang menyajikan pilihan *estimator* yang tepat untuk berbagai masalah. Estimator yang dipilih digunakan untuk membentuk model *linier setelahnya*. [24].

### 2.2.11. Jupyter Notebook

*Jupyter Notebook* telah populer sebagai alat untuk melakukan manipulasi data dalam bahasa pemrograman *Python*. Dalam *Jupyter Notebook*, Anda dapat berinteraksi antara kode dan hasilnya pada sebuah dokumen.. *Jupyter Notebook*, yang memiliki ekstensi *file ipynb*, merupakan dokumen yang diciptakan menggunakan *Jupyter Notebook*. Dokumen ini mencampurkan kode komputer dengan elemen-elemen teks seperti paragraf, rumus matematika, gambar, dan tautan. Alat ini, yang sebelumnya dikenal sebagai *IPython Notebook* dan berkembang menjadi *Jupyter Lab*.[24]. *Jupyter Notebook* adalah program pengembangan interaktif berbasis *web* yang berjalan di *browser default* . Setiap

blok kode dapat dilakukan secara independen, membuatnya sangat *fleksibel* dan mudah dimainkan. Ini memungkinkan penggunaan berbagai jenis teks dalam *Notebook* yang sama. Akibatnya, keluaran kode, visualisasi, persamaan, dan teks biasa semuanya dapat diakses di satu lokasi. Ini memudahkan untuk membuat dan mendistribusikan dokumen, serta menunjukkan kode dan temuan dengan cara yang logis dan menarik. Ini sangat ideal untuk kolaborasi karena berbasis web dan memudahkan untuk berbagi *Notebook* dengan orang lain. *JupyterLab* adalah *plugin Notebook* *Jupyter* dengan lebih banyak fungsi. *Notebook Jupyter* digabungkan dengan aplikasi tambahan di *JupyterLab*, seperti Terminal baris perintah, Konsol kode, dan *Editor Teks*[21]. Adapun tujuan penggunaan *Jupyter Notebook* adalah: memberikan kemudahan kepada peneliti dalam mengakses *Software Data Mining*.

#### **2.2.12. PHP**

*PHP* merupakan bahasa pemrograman standar yang sering digunakan dalam pengembangan situs *web*. Ini merupakan bahasa *skrip* yang dijalankan di *server web*[25].

#### **2.2.13. HTML**

*Hyper Text Markup Language* (HTML) dalam bidang komputer adalah bahasa format teks yang digunakan untuk dokumen di jaringan komputer yang dikenal sebagai *World Wide Web* (atau biasa disebut *web*)[25].

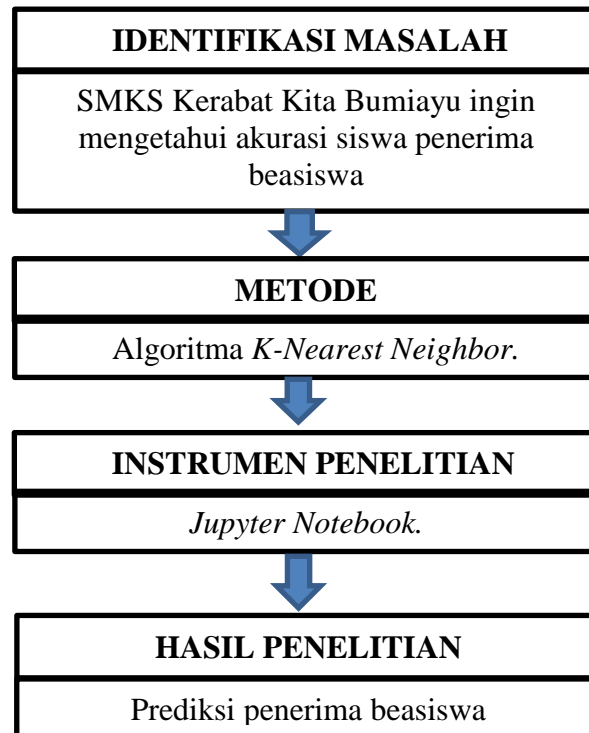
#### **2.2.14. MYSQL**

*MySQL* adalah sistem manajemen basis data (*DBMS*) berbasis *SQL* yang mendukung multi *thread* dan multi *user*. Dengan sekitar enam juta instalasi di seluruh dunia, *MySQL* digunakan secara luas sebagai perangkat lunak manajemen basis data. [25]. *MySQL* menyediakan antarmuka pengguna (*client interface*) yang memudahkan akses ke basis data dengan menggunakan kata sandi untuk memberikan izin untuk tindakan-tindakan yang biasanya diizinkan. Sementara itu, *phpMyAdmin* adalah sebuah perangkat lunak yang dirancang dalam bentuk halaman web, yang dapat diakses melalui *server web*. Fungsinya adalah sebagai pengendali basis data *MySQL*, memungkinkan pengguna *MySQL* mengelola basis data tanpa harus menguasai perintah-perintah *SQL* secara mendalam[26].

### 2.3. Kerangka Pemikiran

Kerangka Kerangka pemikiran merupakan dasar pemikiran yang menjadi garis besar alur logika hubungan antar variabel dalam sebuah penelitian. Kerangka pemikiran adalah dasar pemikiran yang menjadi garis besar alur logika hubungan antar variabel dalam sebuah penelitian

Kerangka pemikiran dapat dilihat pada Gambar 2.4 berikut:



Gambar 2.4 Kerangka Pemikiran

Berikut adalah penjelasan dari kerangka penelitian pada Gambar 2.4:

#### 1. Identifikasi masalah

Permasalahan yang ada di SMKS Kerabat Kita Bumiayu adalah belum adanya suatu sistem untuk memprediksi penerima beasiswa untuk memudahkan sekolah menyalurkan beasiswa tepat sasaran.

#### 2. Metode

Pada studi ini algoritma yang digunakan adalah Algoritma *K-Nearest Neighbor* untuk memprediksi penerima beasiswa di SMKS Kerabat Kita Bumiayu.

### 3. Instrumen Penelitian

Instrumen yang digunakan pada penelitian ini adalah *software Jupyter Notebook* untuk membantu mengolah data penerima beasiswa di SMKS Kerabat Kita Bumiayu.

### 4. Hasil Penelitian

Tujuan dari hasil studi ini adalah sebuah informasi berupa prediksi siswa penerima beasiswa dimana dapat mempermudah sekolah untuk memilih siswa yang memenuhi syarat untuk mendapatkan beasiswa.





