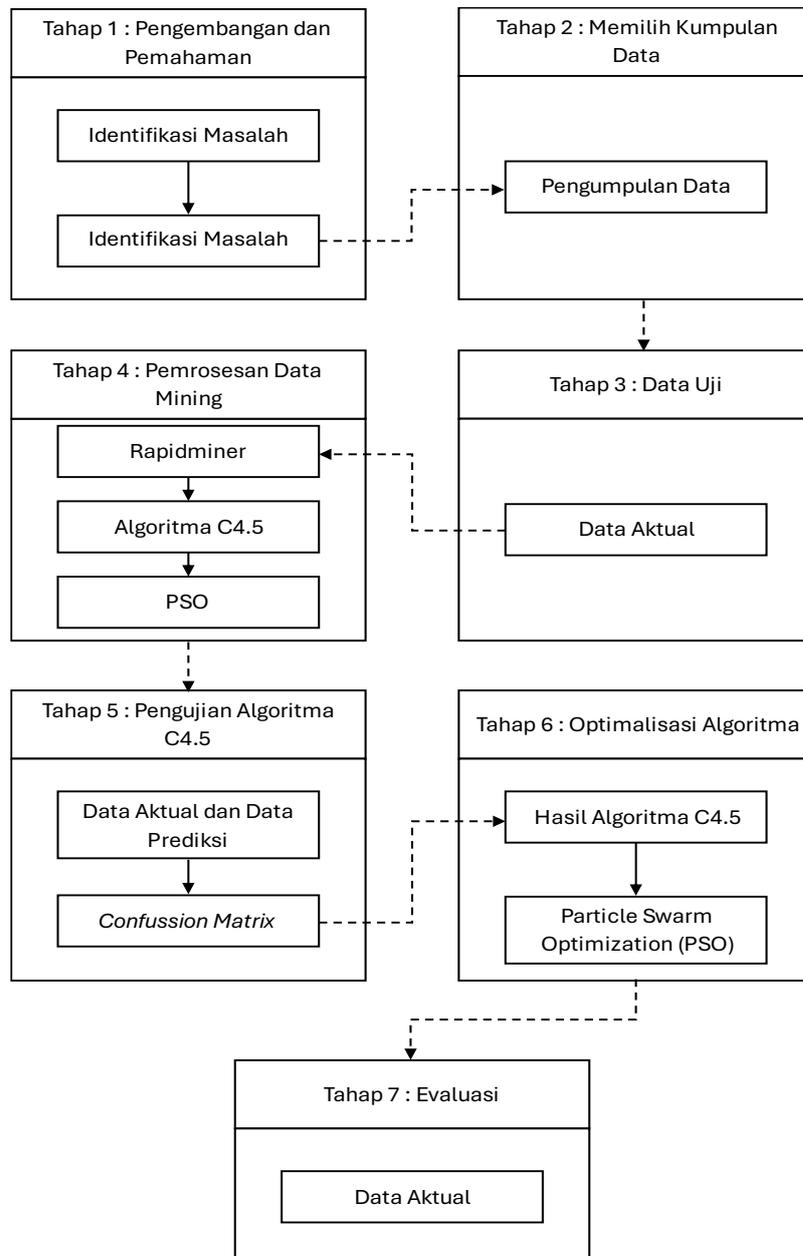


BAB III

METODE PENELITIAN

3.1 Tahapan Penelitian

Tahapan-tahapan yang dilakukan pada penelitian ini dapat dilihat pada gambar berikut :



Gambar 3.1 Alur Tahapan Penelitian

Penelitian ini dibagi menjadi 7 tahapan, diantaranya:

1. Mengidentifikasi masalah yang akan diteliti, mulai dari informasi terkait peluang masalah yang, dataset yang akan digunakan, hingga algoritma yang akan di gunakan untuk proses data mining. Kemudian pada tahap selanjutnya menentukan tujuan untuk penelitian ini, agar penelitian menjadi terstruktur dan terarah dalam proses penelitian yang menjadi hasil akhir
2. Pengumpulan data dengan mencari dataset siswa kelas X SMA Negeri 2 Garut tahun pelajaran 2022/2023 yang akan digunakan untuk proses penelitian.
3. Pengelompokan data aktual yang diperoleh dari dataset. Data aktual merupakan data yang akan diprediksi dalam proses *mining*. Pemrosesan *dataset* menggunakan Tools RapidMiner 9.10.
4. Pengujian dan analisis pada data siswa diuji menggunakan algoritma C4.5 untuk menentukan hasil kelayakan siswa dalam menerima bantuan dana pendidikan atau beasiswa.
5. Penerapan Optimalisasi algoritma untuk meningkatkan tingkat akurasi dari hasil algoritma C4.5.
6. Hasil akhir diketahui dari tingkat akurasi yang sudah di optimalisasi.
7. Evaluasi tingkat akurasi dari hasil yang diproses dengan metode algoritma C4.5 dan *Particle Swarm Optimization* (PSO).

3.3.1 Pra Penelitian

a. Studi Pustaka

Studi Pustaka dilakukan untuk membuka wawasan guna mendapatkan peluan dalam mengangkat tema yang akan diteliti, kemudian studi pustaka juga dilakukan untuk mencari referensi baik itu bersumber dari jurnal, buku ataupun media lain dalam mendapatkan teori-teori dalam menjawab permasalahan yang sudah ditetapkan.

b. Pengumpulan Data

Data yang digunakan pada proses prediksi kelayakan siswa dalam menerima bantuan dana pendidikan menggunakan data primer yang bersumber dari pihak sekolah SMA Negeri 2 Garut. Pada penelitian ini, *dataset* yang digunakan adalah data Siswa kelas X (sepuluh) pertahun pelajaran 2022/2023.

3.3.2 Data

Data yang digunakan dalam penelitian ini merupakan data siswa, dimana data yang digunakan ini dalam pengambilannya bersifat primer. Data siswa ini didapatkan dari Wakil Kepala Sekolah bidang Kesiswaan. *Dataset* yang digunakan dalam penelitian ini merupakan data siswa kelas X (sepuluh) tahun pelajaran 2022/2023 untuk menguji prediksi kelayakan siswa dalam menerima bantuan dana pendidikan dari Algoritma C4.5 yang sudah diuji.

3.3.3 Tools

Dalam penelitian pengolahan *data mining* umumnya digunakan *software* sebagai alat bantu. Rapidminer memiliki kemudahan dalam mengumpulkan data dari semua sumber seperti basis data, *cloud*, dokumen, media sosial dan aplikasi

bisnis. Selain itu dapat mengeksplorasi dan memvisualisasi data secara statistik. Maka dari itu dengan segala kelebihan dan kemudahannya dalam penggunaan mengenai penelitian *data mining*, Rapidminer dipilih sebagai *tools* untuk dijadikan *software* dalam penelitian ini.

3.3.4 Pengujian dan Analisis

Dalam menguji prediksi akurasi metode dengan *dataset* siswa kelas X, tahap pengujian dan analisis menjadi tahapan yang sangat berperan dalam alur penelitian ini. Proses dalam penelitian ini, *dataset* terlebih dahulu dibagi menjadi dua kelompok yaitu 80% sebagai data latih, dan 20% data uji. Kemudian data di uji oleh Algoritma C4.5 sehingga menghasilkan tingkat kesalahan yang kecil.

3.3.5 Optimalisasi Algoritma

Pada penelitian ini mengusulkan metode optimalisasi algoritma berbasis *Particle Swarm Optimization* (PSO) untuk mengatasi permasalahan pada algoritma C4.5 dalam menentukan tingkat akurasi yang nantinya dihubungkan dengan *decision tree*.

3.3.6 Hasil dan Evaluasi

Dari hasil pengujian implementasi algoritma C4.5 pada dataset siswa akan menghasilkan performansi dalam bentuk akurasi dan *confusion matrix*. Pada tahap ini dilakukan untuk mengukur sejauh mana model yang dipilih memenuhi tujuan dari penelitian dan jika demikian, apakah perlu lebih banyak model untuk di buat lagi.

Setelah dataset diuji oleh algoritma C4.5 dan *Decision Tree* maka akan menghasilkan suatu informasi mengenai tingkat akurasi prediksi dengan optimalisasi algoritma berbasis *Particle Swarm Optimization* (PSO).