

BAB II KERANGKA TEORITIS

A. Tinjauan Pustaka

Penelitian rujukan merupakan acuan yang dibutuhkan seorang peneliti untuk melakukan penelitian. Penelitian rujukan pada penelitian ini diambil berdasarkan kesamaan metode yaitu dengan Naive Bayes, banyak penelitian yang menggunakan metode ini dalam berbagai kasus. Antara lain adalah :

1. **Evicienna, Hilda Amalia (2013) didalam penelitiannya yang berjudul “Algoritma C4.5 Untuk Prediksi Hasil Pemilihan Legislatif DPRD DKI JAKARTA”.** Mengemukakan bahwa akan memprediksi hasil pemilu dengan menggunakan algoritma dengan menganalisis sejumlah atribut yang menjadi parameter untuk prediksi hasil pemilu DPRD DKI Jakarta, diantaranya: nama partai, no urut partai, suara sah partai, nama caleg, kota administrasi, jenis kelamin, suara sah caleg, no urut caleg, jumlah perolehan kursi, dan daerah pemilihan. Hasil eksperimen dan evaluasi penelitian prediksi hasil pemilihan legislatif DPRD DKI Jakarta menggunakan algoritma C4.5 terbukti akurat, terlihat dari hasil yang didapat yaitu sebesar nilai akurasi sebesar 97.84% dan nilai AUC sebesar 0.970 dengan tingkat diagnosa Excellent Classification.
2. **Muhammad Fauzul Arifin , Devi Fitriannah (2017) di dalam penelitiannya yang berjudul “Penerapan Algoritma Klasifikasi C4.5 dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada”.** Mengemukakan bahwa penerapan algoritma klasifikasi C4.5 akan dapat diimplementasikan pada rekomendasi penerimaan mitra penjualan di PT. Atria Artha Persada dengan variabel No. Tlp, alamat, status tempat, tempo, cara pembayaran, omset, dan jumlah pembelian, dilihat dari tingkat accuracy yang mencapai 96.26 % dan recall 71.43%., yang menyatakan bahwa perhitungan yang dilakukan akan mampu memprediksi dan me- rekomendasikan penerimaan mitra penjualan dengan baik.
3. **Fina Nasari (2014) di dalam penelitiannya yang berjudul “Penerapan Algoritma C4.5 Dalam Pemilihan Bidang Peminatan Program Studi Sistem Informasi Di STMIK Potensi Utama Medan”.** Mengemukakan bahwa Berdasarkan perhitungan menggunakan algoritma C4.5 diperoleh factor dominan seseorang memilih bidang peminatan adalah berdasarkan nilai JK dengan tingkat kecocokan data hingga 80.14%. 2. Variable penelitian ini masih melihat data nilai dan jenis kelamin, untuk pengembangannya perlu dilihat juga minat dan bakat dari mahasiswa yang akan memilih peminatan, sehingga pemilihan peminatan akan lebih tepat.

4. **Rizky Haqmanullah Pambudi, Budi Darma Setiawan, Indriati (2018) di dalam penelitiannya yang berjudul “Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal”**. Mengemukakan bahwa pendidikan dalam kehidupan suatu negara memegang peranan yang sangat penting untuk menjamin kelangsungan hidup negara dan bangsa. Statistik menunjukkan bahwa tingkat pendidikan Portugal berada di peringkat terbawah yang disebabkan banyak siswa yang putus sekolah. Faktor eksternal berpengaruh pada kegagalan siswa dalam menyelesaikan bidang studi khususnya bidang studi matematika. Algoritma C4.5 merupakan salah satu metode data mining untuk memprediksi kemampuan siswa dalam menyelesaikan bidang studi dilihat dari faktor eksternal siswa. Algoritma C4.5 digunakan untuk mengetahui tingkat akurasi prediksi kemampuan siswa sekolah menengah. Parameter pemilihan fitur adalah faktor-faktor yang mempengaruhi kemampuan siswa sekolah menengah dalam bidang studi matematika. Hasil pengujian dan analisis menunjukkan bahwa Algoritma Decision Tree C4.5 akurat diterapkan untuk prediksi nilai akhir siswa sekolah menengah dengan tingkat akurasi 60%.
5. **Juna Eska (2016) di dalam penelitiannya yang berjudul “Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4.5”**. Mengemukakan bahwa pembelian wallpaper dengan menggunakan metode Data Mining khususnya Algoritma C4.5 akan bermanfaat sekali dalam proses pengambilan keputusan dalam pembelian wallpaper. Yang menjadi faktor tertinggi yang mempengaruhi penjualan adalah faktor jumlah motif wallpaper. Faktor Harga, Ukuran, Kualitas Bahan, dan Warna tidak mempengaruhi pembelian karena wallpaper dengan harga mahal, ukuran yang kecil, kualitas bahan yang baik, dan warna yang sedikit ternyata masih diminati oleh pelanggan.
6. **Lingga Hidayat, Dodi Siregar, Ilham Faisal (2016) di dalam penelitiannya yang berjudul “Sistem Monitoring dan Prediksi Prestasi Akademik Mahasiswa Menggunakan Algoritma Decision Tree C4.5”**. Mengemukakan bahwa informasi yang telah dihasilkan dapat melakukan proses monitoring prestasi untuk setiap mahasiswa, Proses pengecekan nilai maupun pencarian data bisa dilakukan dengan cepat dan tepat. Keunggulan Sistem mampu menghasilkan informasi hasil prediksi untuk setiap mahasiswa. Dengan Menerapkan metode klasifikasi C4.5 untuk melakukan prediksi prestasi akademik mahasiswa sudah tepat.

Tabel 2. 1 Ringkasan Tinjauan Pustaka

| No | Peneliti | Judul | Permasalahan | Kesimpulan |
|----|---|---|--|--|
| 1 | Evicienna, Hilda Amalia,(AMIK Bina Sarana Informatika Jakarta, 2013) | Algoritma C4.5 Untuk Prediksi Hasil Pemilihan Legislatif DPRD DKI JAKARTA | Setiap lima tahun tatacara perhitungan suara selalu berubah sesuai dengan peraturan perundang-undangan yang berlaku. | Pada penelitian ini akan memprediksi hasil pemilu dengan menggunakan algoritma dengan menganalisis sejumlah atribut yang menjadi parameter untuk prediksi hasil pemilu DPRD DKI Jakarta, diantaranya: nama partai, no urut partai, suara sah partai, nama caleg, kota administrasi, jenis kelamin, suara sah caleg, no urut caleg, jumlah perolehan kursi, dan daerah pemilihan. Hasil eksperimen dan evaluasi penelitian prediksi hasil pemilihan legislatif DPRD DKI Jakarta menggunakan algoritma C4.5 terbukti akurat, terlihat dari hasil yang didapat yaitu sebesar nilai akurasi sebesar 97.84% dan nilai AUC sebesar 0.970 dengan tingkat diagnosa Excellent Classification. |
| 2 | Muhammad Fauzul Arifin, Devi Fitriyah, (Universitas Mercu Buana Jakarta, 2017). | Penerapan Algoritma Klasifikasi C4.5 dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : | Perusahaan ini sering kali dihadapkan pada masalah penipuan, order fiktif, pembayaran macet sampe di tinggal oleh | Penerapan algoritma klasifikasi C4.5 akan dapat diimplementasikan pada rekomendasi penerimaan mitra penjualan di PT. Atria Artha Persada dengan variabel No. Tlp, alamat, status tempat, tempo, cara pembayaran, omset, dan |

| | | | | |
|---|---|---|---|--|
| | | PT Atria Artha Persada | mitra tanpa kabar. | jumlah pembelian, dilihat dari tingkat accuracy yang mencapai 96.26 % dan recall 71.43%., yang menyatakan bahwa perhitungan yang dilakukan akan mampu memprediksi dan merekomendasikan penerimaan mitra penjualan dengan baik. |
| 3 | Fina Nasari, (STMIK Potensi Utama, 2014) | Penerapan Algoritma C4.5 Dalam Pemilihan Bidang Peminatan Program Studi Sistem Informasi Di STMIK Potensi Utama Medan | Ketidaksesuaian bidang peminatan dengan minat dan keterampilan mahasiswa. | Berdasarkan perhitungan menggunakan algoritma C4.5 diperoleh factor dominan seseorang memilih bidang peminatan adalah berdasarkan nilai JK dengan tingkat kecocokan data hingga 80.14%. 2. Variable penelitian ini masih melihat data nilai dan jenis kelamin, untuk pengembangannya perlu dilihat juga minat dan bakat dari mahasiswa yang akan memilih peminatan, sehingga pemilihan peminatan akan lebih tepat. |
| 4 | Rizky Haqmanullah Pambudi, Budi Darma Setiawan, Indriati, (Fakultas Ilmu Komputer, Universitas Brawijaya. 2018) | Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal | Tingkat pendidikan Portugal berada ditingkat terendah disebabkan banyak siswa yang putus sekolah. Factor eksternal berpengaruh pada kegagalan | Pendidikan dalam kehidupan suatu negara memegang peranan yang sangat penting untuk menjamin kelangsungan hidup negara dan bangsa. Statistik menunjukkan bahwa tingkat pendidikan Portugal berada di peringkat terbawah yang disebabkan banyak siswa yang putus sekolah. Faktor eksternal berpengaruh pada |

| | | | | |
|---|-----------------|---|---|---|
| | | | <p>siswa dalam menyelesaikan studi khususnya dibidang studi matematika.</p> | <p>kegagalan siswa dalam menyelesaikan bidang studi khususnya bidang studi matematika. Algoritma C4.5 merupakan salah satu metode data mining untuk memprediksi kemampuan siswa dalam menyelesaikan bidang studi dilihat dari faktor eksternal siswa. Algoritma C4.5 digunakan untuk mengetahui tingkat akurasi prediksi kemampuan siswa sekolah menengah. Parameter pemilihan fitur adalah faktor-faktor yang mempengaruhi kemampuan siswa sekolah menengah dalam bidang studi matematika. Hasil pengujian dan analisis menunjukkan bahwa Algoritma Decision Tree C4.5 akurat diterapkan untuk prediksi nilai akhir siswa sekolah menengah dengan tingkat akurasi 60%.</p> |
| 5 | Juna Eska, 2016 | Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4.5 | | <p>bahwa pembelian wallpaper dengan menggunakan metode Data Mining khususnya Algoritma C4.5 akan bermanfaat sekali dalam proses pengambilan keputusan dalam pembelian wallpaper. Yang menjadi faktor tertinggi yang mempengaruhi penjualan</p> |

| | | | | |
|---|---|---|---|--|
| | | | | adalah faktor jumlah motif wallpaper. Faktor Harga, Ukuran, Kualitas Bahan, dan Warna tidak mempengaruhi pembelian karena wallpaper dengan harga mahal, ukuran yang kecil, kualitas bahan yang baik, dan warna yang sedikit ternyata masih diminati oleh pelanggan. |
| 6 | Lingga Hidayat, Dodi Siregar, Ilham Faisal, 2016 | Sistem Montoring dan Prediksi Prestasi Akademik Mahasiswa Menggunakan Algoritma Decision Tree C4.5 | sistem tidak dapat memprediksi prestasi akademik mahasiswa untuk periode berikutnya. | informasi yang telah dihasilkan dapat melakukan proses monitoring prestasi untuk setiap mahasiswa, Proses pengecekan nilai maupun pencarian data bisa dilakukan dengan cepat dan tepat. Keunggulan Sistem mampu menghasilkan informasi hasil prediksi untuk setiap mahasiswa. Dengan Menerapkan metode klasifikasi C4.5 untuk melakukan prediksi prestasi akademik mahasiswa sudah tepat. |

Perbedaan penelitian ini dengan penelitian yang dilakukan **Lingga Hidayat, Dodi Siregar, Ilham Faisal** dengan judul “**Sistem Montoring dan Prediksi Prestasi Akademik Mahasiswa Menggunakan Algoritma Decision Tree C4.5**” adalah pada jenis variabel dan output yang dihasilkan. Perbedaan lainnya adalah membangun sebuah sistem informasi yang dapat menampilkan informasi terkait dengan banyaknya permasalahan yang dilakukan siswa secara menyeluruh.

B. Landasan Teori

Dalam rangka memperoleh suatu pedoman guna lebih memperdalam masalah, maka perlu dikemukakan suatu landasan teori yang bersifat ilmiah. Dalam

landasan teori ini dikemukakan teori yang ada hubungannya dengan materi-materi yang digunakan untuk memecahkan masalah pada penelitian ini

1. Pengertian Bimbingan Konseling

Bimbingan dan konseling memegang tugas dan tanggung jawab yang penting untuk mengembangkan lingkungan, membangun interaksi dinamis antara individu dengan lingkungan, membelajarkan individu untuk mengembangkan, merubah dan memperbaiki perilaku. Bimbingan dan konseling bukanlah kegiatan pembelajaran dalam konteks adegan mengajar yang layaknya dilakukan guru sebagai pembelajaran bidang studi, melainkan layanan ahli dalam konteks memandirikan peserta didik. (H. Kamaluddin, 2011).

2. Model prediksi

Menurut eko prasetyo (2016:5) Pekerjaan ini berkaitan dengan pembuatan sebuah model yang dapat melakukan pemetaan dari setiap himpunan variabel ke setiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan baru yang didapat. Ada 2 jenis model prediksi, yaitu klasifikasi dan regresi. Klasifikasi digunakan untuk variabel target diskret, sedangkan regresi digunakan untuk variabel target kontinu. Contoh pekerjaan yang menggunakan jenis klasifikasi adalah melakukan deteksi jenis penyakit pasien berdasarkan sejumlah nilai-nilai parameter penyakit yang diderita masuk. pekerjaan ini termasuk jenis klasifikasi karena target yang diharapkan adalah diskret, hanya beberapa jenis kemungkinan nilai.

3. Pengembangan Sistem SDLC

Menurut Sanyoto Gondodiyoto (2007:514), siklus daur hidup sistem (*system life cycle*) adalah proses evalusioner yang terjadi dalam penerapan sistem atau sub sistem informasi berbasis komputer, mulai dari perancangan kebutuhan sistem sampai dioperasikan untuk kegiatan organisasi. Proses tersebut terdiri dari kegiatan perencanaan, analisis, rancangan (*design / construction*), penerapan (*system implementation*), dan penggunaan sistem atau sering disebut dengan istilah *production (operasionalisasi sistem)* sebagai suatu *system life* digunakan sesuai kebutuhan pengguna (*user*).

Pada tahap pengguna tersebut seluruh operasi sistem dilakukan oleh pengguna, sedangkan kegiatan perencanaan, analisis, rancangan dan penerapan dilakukan oleh teknis sistem informasi (*system development team*). Kegiatan yang dilakukan oleh team developer dinamakan siklus hidup pengembangan sistem (*system development life cycle*). Perlu dijelaskan bahwa

pada beberapa literature, tahap penerapan sistem sering disebut dengan istilah implementasi, tetapi sering kali istilah implementasi juga digunakan untuk seluruh kegiatan pengembangan sistem informasi.

Implementasi ialah merupakan proses pemasangan sistem yang baru direncanakan, termasuk semua perlengkapan dan system software yang dibeli. Implementasi juga sering didefinisikan sebagai kegiatan memperoleh dan mengintegrasikan (integration) seluruh sumber daya informasi, baik yang berwujud fisik maupun non-fisik, agar sistem informasi dapat dioperasikan untuk menghasilkan informasi sesuai kebutuhan para pengguna. Proses implementasi pada dasarnya terdiri dari berbagai tahap, yakni: perencanaan, perancangan, mendapatkan sumber daya hardware, software, database, fasilitas, fisik lain, pemograman, penyiapan lokasi, pemasangan dan pemeriksaan instalasi peralatan, pemasangan dan pemeriksaan software yang dibeli, konversi file, pemeriksaan akhir/serah terima, dokumentasi, serta mendidik calon pemakai sistem baru (*user training*).

Menurut Sanyoto Gondodiyoto (2007:516), sistem informasi dibangun menurut kaidah dan metode-metode tertentu yang disebut metodologi (*system development methodology*). Menurut berbagai text-book dan sumber lain, terdapat beberapa metodologi yang dapat diikuti. Antara lain yang disebut dengan waterfall system, software engineering (SE), structured system analysis and design (SSAD), soft system methodology (SSM), the prototyping approach dan masih banyak lagi metodologi yang dikemukakan tokoh-tokoh yang banyak berkecimpung dalam metodologi dengan pendekatan-pendekatan pengembangan sistem tersebut.

Perlu dicatat bahwa apapun metodologinya, tetapi sesungguhnya tahap-tahap pembangunan sistem informasi hakekatnya terdiri dari rangkaian kegiatan berikut :

Tabel 2. 2 Tahap-tahap Pembangunan Sistem

| Tahap | Keterangan |
|----------------------|---|
| Feasibility Study | Menentukan layak/tidaknya, cost-benefit satu proposed system. |
| Information Analysis | Menggali user requirements. |
| System design | Perancangan sistem user interface, sistem file, dan information processing functions yang akan dilakukan, dan sebagainya. |
| Program Development | Mendesain, coding, compiling, testing dan documenting program. |

| | |
|----------------------------------|---|
| Procedures and Forms Development | Mendesain system procedure dan form-form yang akan digunakan. |
| Acceptance Testing | Final test formal approval acceptance dari user. |
| Conversion | Implementasi , mengganti sistem lama dengan sistem baru. |
| Operation and Maintance | On-going production, operasional sistem, perawatan dan perbaikan, evaluasi atau usul sistem yang lebih baru lagi dikemudian hari. |

(Sumber : Sanyoto Gondodiyoto, 2007, p.516)

Meskipun terdapat berbagai konsep metodologi, namun siklus kegiatan pada hakekatnya sama dan dapat digambarkan secara garis besar sebagai berikut :



Gambar 2. 1 Siklus Pengembangan Aplikasi

(Sumber : Sanyoto Gondodiyoto, 2007, p.516)

Secara lebih rinci kegiatan-kegiatan dalam berbagai tahap pengembangan sistem aplikasi tersebut dapat diuraikan sebagai berikut :

a. The Planning Phase

Kegiatan-kegiatan yang dilakukan pada tahap ini antara lain adalah :

- 1) Mengenal masalah yang dihadapi (*Recognize the Problem*).
- 2) Merumuskan problem yang sesungguhnya (*Define the Problem*).
- 3) Menetapkan tujuan/sasaran (*Set. System Objectives*).
- 4) Identifikasi kendala/keterbatasan (*Identify System Constraint*).
- 5) Melakukan studi kelayakan (*Conduct a Feasibility Study*).
- 6) Menyiapkan proposal (*Prepare a System Study Proposal*).
- 7) Disetujui/tidaknya usulan (*Approve or Disapprove the Study Project*).
- 8) Membangun mekanisme control (*Establish a Control Mechanism*).

b. The Analysis Phase

Dalam tahap ini dilakukan studi tentang sistem yang berjalan saat ini (existing/ current system) dalam rangka menilai ada/tidaknya kelemahan dan apakah perlu disempurnakan atau disusun sistem pengganti, terdiri dari:

- 1) Membentuk tim atau proyek.
- 2) Merumuskan tujuan/kebutuhan informasi.
- 3) Merumuskan *system performance criteria*.
- 4) Menyiapkan *design proposal*.

c. Design Phase

System design merupakan rancangan tentang data dan proses yang diperlukan:

- 1) Menyusun *detailed system design*.
- 2) Identifikasi alternatif system configuration dan mengevaluasinya, serta memilih (*select the best configuration*).
- 3) Menyiapkan usulan *implementasi*.

d. The Implementation Phase

Pada tahap ini dilakukan acquisition dan integrase sumber fisik dan non fisik agar sistem dapat dioperasikan.

- 1) Perencanaan implementasi dan mengumumkannya.
- 2) Perolehan sumber daya hardware dan software.
- 3) Menyiapkan database.
- 4) Menyiapkan fasilitas fisik.
- 5) Pelatihan user.

e. The Use Phase

Antara lain kegiatan-kegiatan sebagai berikut :

- 1) Penggunaan sistem (use the system).
- 2) Evaluasi atau pemeriksaan (audit the system).

f. Maintain The System

- 1) Melakukan perbaikan
- 2) Menyiapkan usulan reengineering bila diperlukan.

4. Algoritma C4.5

Algoritma C4.5 merupakan kelompok algoritma pohon Keputusan (decision tree). Algoritma ini mempunyai input berupa training samples dan samples. Training samples data contoh yang akan digunakan untuk membangun sebuah tree yang telah diuji kebenarannya. Sedangkan samples

merupakan field-field data yang nantinya akan digunakan sebagai parameter dalam melakukan klasifikasi data.

Algoritma C4.5 merupakan pengembangan dari algoritma ID3, dimana pengembangan dilakukan dalam hal bisa mengatasi missing data, bisa mengatasi data kontinyu, pruning. Algoritma C4.5 memiliki kelebihan yaitu mudah dimengerti, fleksibel, dan menarik karena dapat divisualisasikan dalam bentuk gambar (pohon keputusan) Ada beberapa tahapan dalam membuat sebuah pohon keputusan dalam algoritma C4.5 (Rizky Haqmanullah Pambudi, dkk:2018), yaitu:

1. Mempersiapkan data training. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya atau data disebut data masa lalu dan sudah dikelompokkan dalam kelas – kelas tertentu.
2. Menghitung akar dari pohon. Aar akan diambil dari atribut yang akan dipilih, degan cara menghitung nilai gain dari masing-masing atribut nilai gain yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy. Untuk nilai entropy digunakan rumus :

$$Entropy(S) = - \sum_{i=1}^n p_i * \log_2 p_i \quad (1)$$

Keterangan :

S = Himpunan kasus

n = Jumlah partisi S

p_i = Proporsi S_i terhadap S

3. Menghitung nilai Gain menggunakan persamaan 2.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} Entropy(S_i) \quad (2)$$

Keterangan :

S = Himpunan kasus

A = Atribut

n = Jumlah Partisi atribut

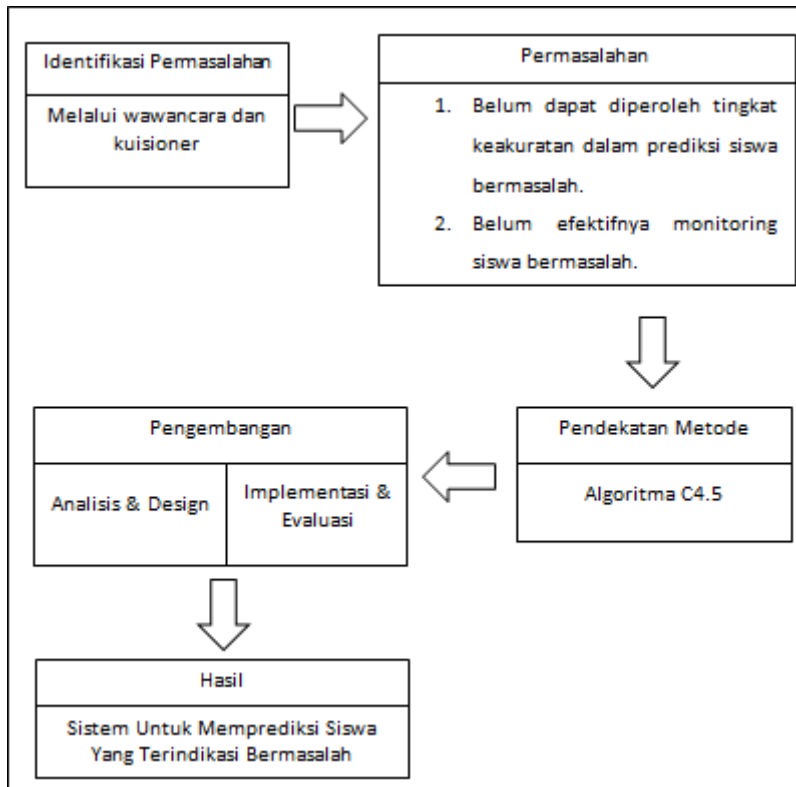
$|S_i|$ = Proporsi S_i terhadap S

$|S|$ = Jumlah kasus dalam S

4. Ulangi langkah ke 2 dan langkah ke 3 hingga semua record terpartisi
5. Proses partisi pohon keputusan akan berhenti saat :
 - a. Semua tupel dalam mode N mendapat kelas yang sama
 - b. Tidak ada lagi atribut di dalam tupel yang dipartisi lagi.
 - c. Tidak ada tupel di dalam cabang kosong.

C. Kerangka Pemikiran

Berdasarkan dukungan landasan teoritis yang diperoleh dari eksplorasi teori yang dijadikan rujukan penelitian, maka dapat disusun kerangka pemikiran sebagai berikut :



Gambar 2. 2 Kerangka Pemikiran

Kerangka pemikiran pada gambar 2.2 dapat dijelaskan sebagai berikut :

1. Identifikasi masalah di dapat dengan melalui kuisisioner.
2. Komponen permasalahan mencakup fenomena yaitu belum efektifnya monitoring terhadap siswa bermasalah yang berdampak pada prestasi akademik dan belum adanya alat / sistem yang membantu dalam memonitoring siswa bermasalah.
3. Pendekatan yang digunakan yaitu Algoritma C4.5.
4. Dalam pengembangan dibagi menjadi dua, yaitu tahap analisis dan design, dan tahap implementasi dan evaluasi.
5. Hasilnya yang di dapatkan adalah sebuah sistem untuk memprediksi siswa bimbingan konseling.

D. Hipotesis

Hipotesis dalam penelitian ini adalah penerapan metode Algoritma C4.5 diduga dapat memprediksi siswa yang terindikasi bermasalah di SMKN 4 Bogor.