

## BAB II KERANGKA TEORITIS

### A. Landasan Teori

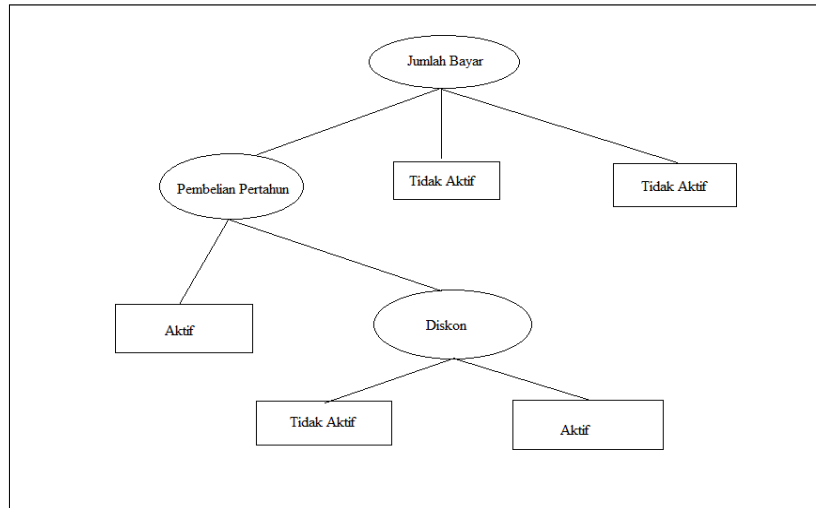
#### 1. Data Mining

Menurut (Deny Jollyta, William Ramdhan, Muhammad Zarlis”Konsep Data Mining Dan Penerapan”,2020:22) menyatakan bahwa data mining merupakan proses untuk menggali (*mining*) pengetahuan dan informasi baru dari data yang berjumlah banyak pada data warehouse,dengan menggunakan kecerdasan buatan (*Artificial Intelligence*),statistic dan matematika. Data mining merupakan teknologi yang diharapkan dapat menjembatani komunikasi antara data dan pemakainya. Beberapa teknik dan sifat data mining menurut (Hermawati, 2013:14) adalah sebagai berikut :

- (a) Klasterisasi Adalah mempartisi data-set menjadi beberapa sub-net atau kelompok sedemikian rupa sehingga elemen-elemen dari suatu kelompok tertentu memiliki set property yang di share bersama, dengan tingkat similaritas yang tinggi dalam suatu kelompok yang rendah. Disebut juga dengan “unsupervised learning”;
- (b) Regresi Adalah memprediksi nilai dari suatu variabel kontinyu yang diberikan berdasarkan nilai adari variabel yang lain, dengan mengasumsikan sebuah model ketergantungan linier atau nonlinier;
- (c) Klasifikasi Adalah menentukan sebuah record data baru ke salah satu dari beberapa kategori (kelas) yang telah didefinisikan sebelumnya dan disebut juga dengan “supervised learning”;
- (d) Kaidah Asosiasi (association rule) Adalah mendeteksi kumpulan atribut-atribut yang muncul bersamaan (co-occur) dalam frekuensi yang sering dan membentuk sejumlah kaidah dari kumpulan-kumpulan tersebut;

#### 2. Pohon Keputusan (Decision tree)

Pohon keputusan atau decision tree adalah pemetaan, mengenai alternatif-alternatif pemecahan masalah yang dapat diambil dari masalah tersebut. Pohon tersebut juga memperlihatkan faktor-faktor kemukninan/probabilitas yang akan mempengaruhi alternatif-alternatif keputusan tersebut, disertai dengan estimasi hasil akhir yang akan didapat bila kita mengambil alternatif keputusan tersebut.(Retno Tri Vulandari & Magfirah, 2018).



**Gambar 2.1 Contoh Pohon Keputusan**

Decision tree merupakan metode klasifikasi yang paling sering digunakan. Karena dalam pengerjaannya tidak memerlukan waktu yang lama dan hasilnya pun mudah untuk dipahami dan banyak penelitian dalam kasus ini sering menggunakan decision tree untuk mendapatkan hasil yang maksimal

Ayu Rizqi Oktaviana & Magfirah (2018) menyatakan pada decision tree terdapat 3 jenis node, yaitu :

- (a) Root Node, merupakan node paling atas, pada node ini tidak ada input dan bisa tidak mempunyai output lebih dari satu.
- (b) Internal Node, merupakan node percabangan, pada node ini hanya terdapat satu input dan mempunyai output minimal 2.
- (c) Leaf Node atau terminal node, merupakan node terakhir, pada node ini hanya terdapat satu input dan tidak mempunyai output

Kelebihan yang dimiliki decision tree antara lain :

- (a) Hasil analisa berupa diagram pohon yang mudah dimengerti.
- (b) Mudah untuk dibangun, serta membutuhkan data percobaan yang
- (c) lebih dari sedikit dibandingkan algoritma klasifikasi lainnya.
- (d) Mampu mengolah data nominal dan kontinyu.
- (e) Menggunakan teknik statistik sehingga dapat divalidasi.
- (f) Akurasi yang dihasilkan mampu menandingi teknik klasifikasi lainnya.

### 3. Algoritma C4.5

Ayu Rizqi Oktaviana & Magfirah (2018) menyatakan algoritma C4.5 merupakan suatu deretan algoritma untuk permasalahan klasifikasi didalam sebuah mesin dan himpunan data. Dengan nilai data yang bervariasi, dimana kejadian diuraikan oleh koleksi atribut dan mempunyai salah satu dari satu set kelas yang eksklusif.

Kemudian hasil dari pembelajaran selanjutnya akan digunakan untuk mengolah data-data yang baru yang disebut test dataset. Karena algoritma C4.5 digunakan untuk melakukan klasifikasi, jadi hasil dari pengolahan test dataset berupa pengelompokan data ke dalam kelas-kelasnya. Umumnya, langkah-langkah algoritma C4.5 yang digunakan untuk membentuk pohon keputusan adalah Untuk memilih atribut sebagai root, didasarkan pada nilai gain tertinggi dari atribut yang tersedia. Sementara itu, untuk mendapat nilai gain tertinggi kita harus menghitung nilai entropy dari semua nilai didalam atribut. Entropy berperan sebagai parameter untuk mengukur varian dari data sampel. Setelah nilai entropy dalam data sampel diketahui, atribut yang paling berpengaruh akan menjadi pengukur dalam pengklasifikasian data, ukuran ini disebut sebagai Information gain. Pada penelitian, untuk memilih atribut sebagai akar didasarkan pada nilai information gain tertinggi dari atribut yang ada. Berikut rumus perhitungan yang digunakan (Magfirah, 2018).

(a) Entropy, merupakan langkah awal dalam perhitungan algoritma C4.5.

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

Keterangan :

S : himpunan kasus

n : jumlah partisi S

Pi : proporsi dari Si terhadap S

(b) Information gain, merupakan kriteria yang digunakan untuk memilih suatu atribut yang popular, dapat dihitung dengan cara pengelompokan berdasarkan masing-masing atribut dalam satu data. Notasi information gain adalah Gain (S,A) yang berarti dalam data atribut A relative terhadap output

$$Gain(S, A) = Entropy(S) - \sum_i^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Keterangan :

- S : himpunan kasus
- A : atribut
- N : jumlah partisi atribut A
- |Si| : jumlah kasus pada partisi ke-i
- |S| : jumlah kasus dalam

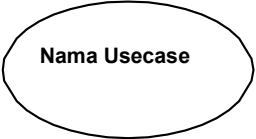
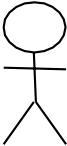
#### 4. Unified Modeling Language (UML)



Rosa A.S & M. Shalahuddin (2016:137) menyatakan UML dapat memenuhi kebutuhan pemodelan berbentuk visual untuk menspesifikasikan, menggambarkan, membangun, dokumentasi dari perangkat lunak, UML berfungsi sebagai pemodelan dan komunikasi sebuah sistem dengan menggunakan diagram dan teks pendukung sebagai bahasa visual.

##### a. Usecase Diagram

Sholiq (2006:7) menyatakan Diagram usecase menggambarkan interaksi antara usecase dan aktor dimana aktor dapat berupa orang, sistem, atau peralatan yang berinteraksi dengan sistem yang dibangun, use case menggambarkan fungsi sistem dan persyaratan persyaratan yang harus dipenuhi dari sudut pandang pemakaian.

**Tabel 2. 1 Simbol Usecase Diagram**

Simbol	Deskripsi
<b>Use case</b> 	Fungsionalitas yang disediakan sistem sebagai unit-unit yang saling bertukar pesan antar unit atau aktor; biasanya dinyatakan dengan menggunakan kata kerja diawal <i>frase</i> nama <i>Use Case</i>
<b>Aktor / actor</b> 	Orang, proses, atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat diluar sistem informaasi yang akan dibuat itu sendiri, jadi walaupun simbol dari aktor adalah gambar orang, tapi aktorbelum tentu merupakan orang: biasanya dinyatakan menggunakan kata benda diawal <i>frase</i> nama aktor

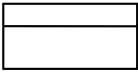


<b>Asosiasi / association</b>  	Komunikasi antara aktor dan <i>use case</i> yang berpartisipasi pada <i>use case</i> atau <i>use case</i> memiliki interaksi dengan aktor
<b>Ekstensi / extend</b>  <<exten d>>	Relasi <i>use case</i> tambahan ke sebuah <i>use case</i> dinamakan <i>use case</i> yang ditambahkan dapat berdiri sendiri walau tanpa <i>use case</i> tambahan itu; mirip dengan prinsip <i>inheritance</i> pada pemrograman berorientasi objek; biasanya <i>use case</i> tambahan memiliki nama depan yang sama dengan <i>use case</i> yang ditambahkan.
<b>Generalisasi / generalization</b>  	Hubungan generalisasi dan spesialisasi (umum-khusus) antara dua buah <i>use case</i> dimana fungsi yang satu adalah fungsi yang lebih umum dari lainnya.





Sumber : (Sholiq 2006, 7)

#### b. Class Diagram

Menurut (Sholiq 2006, 13) diagram class menggambarkan interaksi antara kelas dan sistem, kelas memiliki informasi dan tingkah laku (Behavior) yang berhubungan dengan informasi tersebut, Sebuah kelas pada diagram kelas dibuat untuk tipe objek pada diagram sequence.

Tabel 2. 2 Simbol Class Diagram

Simbol	Deskripsi
Kelas  	Kelas pada struktur sistem
Antarmuka / <i>Interface</i>  	Sama dengan konsep <i>interface</i> dalam pemrograman berorientasi objek
Asosiasi / <i>Association</i>  	Relasi antar kelas dengan makna umum, asosiasi biasanya juga disertai dengan <i>multiplicity</i>





Asosiasi berarah / <i>Directed Association</i> 	Relasi antar kelas dengan makna kelas satu digunakan oleh kelas yang lain, asosiasi biasanya juga disertai dengan <i>multiplicity</i>
Generalisasi 	Relasi antar kelas dengan makna generalisasi-spesialisasi (umum khusus)
Kebergantungan / <i>Dependency</i> 	Relasi antar kelas dengan makna kebergantungan antarkelas
Agregasi / <i>Aggregation</i> 	Relasi antarkelas dengan makna semua-bagian (whole-part)


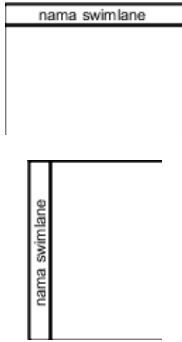
Sumber : (Sholiq 2006, 7)

### c. Activity Diagram

Menurut (Sholiq 2006, 8) diagram activity menggambarkan sebuah aliran fungsionalitas sistem, diagram activity juga dapat dipakai untuk menggambarkan aliran kerja bisnis (Business Workflow) dan dapat juga dipakai untuk menggambarkan aliran kejadian (Flow of Event) di dalam usecase.

Tabel 2. 3 Simbol Activity Diagram

Simbol	Deskripsi
<b>Status awal</b> 	Status awal aktivitas sistem, sebuah diagram aktivitas memiliki sebuah status awal
<b>Aktivitas</b> 	Aktivitas yang dilakukan sistem, aktivitas biasanya diawali dengan kata kerja
<b>Percabangan / <i>decision</i></b> 	Asosiasi percabangan dimana jika adapilihan aktivitas lebih dari satu
<b>Penggabungan / <i>join</i></b> 	Asosiasi penggabungan dimana lebih darisatu aktivitas digabungkan menjadi satu

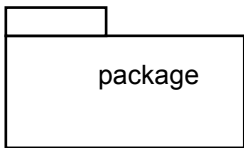
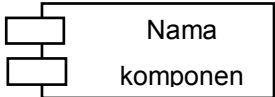
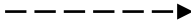
<b>Status akhir</b>		Status akhir yang dilakukan sistem, sebuah diagram aktivitas memiliki sebuah status akhir
<b>Swimlane</b>  <b>Atau</b>		Memisahkan organisasi bisnis yang bertanggung jawab terhadap aktivitas yang terjadi

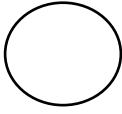

Sumber : (Sholiq 2006, 8)

#### d. Sequence Diagram

Diagram sequence menurut (Sholiq 2006, 9-10) Berfungsi untuk menggambarkan aliran fungsionalitas dalam usecase diagram ini juga menunjukkan langkah langkah proses yang terjadi dalam usecase.

**Tabel 2. 4 Simbol Sequence Diagram**

Simbol	Deskripsi
package 	Package adalah sebuah wadah dari satu atau lebih komponen
komponen 	Berisi koponen komponen dalam sistem
Kebergantungan/dependency 	Kebergantungan antar komponen, arah panah mengarah pada komponen

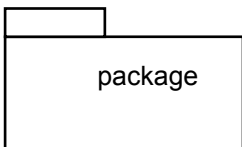
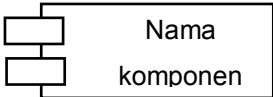

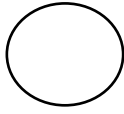

Antarmuka/Interface 	Antarmuka komponen
Link 	Relasi antar package atau komponen

Sumber : (Sholiq 2006, 9-10)

**e. Simbol Component Diagram**

Diagram komponen menurut (Sholiq 2006,15-16) berfungsi untuk menggambarkan model secara fisik komponen berupa perangkat lunak pada sistem dan keterhubungan antara perangkat lunak dan sistem, terdapat dua komponen pada diagram komponen yaitu executable dan kode pustaka (libraries code).

**Tabel 2. 5 Simbol Component Diagram**

Simbol	Deskripsi
package 	Package adalah sebuah wadah dari satu atau lebih komponen
komponen 	Berisi koponen komponen dalam sistem
Kebergantungan/dependency 	Kebergantungan antar komponen, arah panah mengarah pada komponen
Antarmuka/Interface 	Antarmuka komponen
Link 	Relasi antar package atau komponen

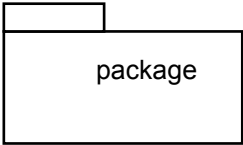
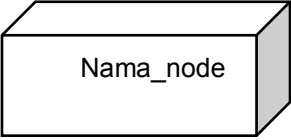
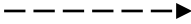

Sumber : (Sholiq 2006, 15-16)



**f. Simbol Deployment Diagram**

Diagram deployment menurut (Sholiq 2006, 17) menggambarkan rancangan fisik jaringan dimana akan terdapat banyak komponen dalam diagram ini, terdapat node pada diagram ini dimana node berisi banyak sub sistem yang dijalankan pada peralatan fisik yang terpisah.

**Tabel 2. 6 Simbol Deployment Diagram**

Simbol	Deskripsi
package 	Package adalah sebuah wadah dari satu atau lebih komponen
Node 	Berisi perangkat lunak atau perangkat keras
Kebergantungan/dependency 	Kebergantungan antar node, arah panah mengarah pada node
Link 	Relasi antar node

Sumber : (Sholiq 2006, 17)

**5. Bahasa Pemrograman**

**a. Database**

Menurut (Rosa A.S 2016, 43) basis data adalah sistem yang sudah terkomputerisasi yang berfungsi untuk memelihara data atau informasi yang sudah di olah sebelumnya dan membuat informasi atau data tersebut tersedia saat di butuhkan di masa yang akan datang.

**b. MySQL**

Menurut (Dr. Gede Indrawan 2021, 3) Mysql adalah sebuah server database yang bersifat open source atau gratis digunakan, perangkat lunak database ini berbentuk database relational yang menggunakan bahasa permintaan yang bernama SQL (Structured query language).

### c. PHP

Menurut Supono & Putratama (2018: 1) mengemukakan bahwa “PHP (PHP: hypertext preprocessor) adalah suatu bahasa pemrograman yang digunakan untuk menterjemahkan basis kode program menjadi kode mesin yang dapat dimengerti oleh komputer yang bersifat server-side yang ditambahkan ke HTML”.

### d. Hypertext Markup Language (HTML)

Menurut (Wardana,2016:3) Hypertext markup language (HTML) merupakan bahasa pemrograman dasar untuk mengelola website. Akan tetapi, HTML hanya terbatas pada pembuatan website statis (website yang tidak dapat berinteraksi aktif dengan user). Maka dari itu, HTML biasa dikombinasikan dengan bahasa pemrograman web lainnya.

## 6. Web Server

### a. Webserver

Menurut (Yudho yudhanto 2018, 10) Webserver memiliki dua pengertian jika merujuk pada hardware webserver berfungsi untuk menyimpan data seperti HTML dokumen, gambar sedangkan jika merujuk pada software fungsi webserver adalah pusat yang mengontrol proses permintaan yang diterima oleh browser.

### b. Apache

Menurut (Tim EMS 2015, 17) Apache adalah sebuah perangkat lunak server yang free dan open source, Apache dapat dijalankan sebagai layanan/ service yang memungkinkan penggunaanya dapat menghidupkan dan mematikannya.

## B. Tinjauan Studi

Penelitian rujukan merupakan acuan yang dibutuhkan seorang peneliti untuk melakukan penelitian. Penelitian rujukan ini diambil berdasarkan kesamaan metode yang digunakan yaitu Algoritma C4.5. Banyak penelitian yang menggunakan metode ini dalam berbagai kasus.

- (1) Penelitian ini dilakukan oleh **Muhammad Yunus, Hanandriya Ramadhan , Dimas Rizki Aji , Agus Yulianto** dengan judul “**Penerapan Metode Data Mining C4.5 Untuk Pemilihan Beasiswa Kartu Indonesia pintar**” Berdasarkan hasil uji coba dan pembahasan yang sudah dilakukan untuk pemilihan penerima Kartu Indonesia pintar di SMPN 38 Jakarta dapat diambil kesimpulan bahwa menggunakan metode klasifikasi dengan algoritma C4.5

bisa selesaikan. Data siswa dapat diprediksi dengan accuracy yang cukup tinggi, hasil tersebut dapat diperoleh dari data training dan data testing dengan accuracy mencapai 90,24% sehingga dengan hasil tersebut algoritma C4.5 sangat baik dalam Pemilihan penerima Kartu Indonesia Pintar. Dengan hasil accuracy sebesar 90,24% dapat dijadikan pendukung untuk digunakan oleh pihak SMPN 38 Jakarta dalam pemilihan penerima Kartu Indonesia Pintar dengan algoritma C4.5.

- (2) Penelitian ini dilakukan oleh **Agustiena Merdekawa dengan judul “Sistem Pendukung Keputusan Penerimaan Beasiswa Menggunakan Algoritma C4.5”** Penelitian ini digunakan untuk mendukung keputusan pemberian beasiswa dengan menggunakan algoritma C4.5 yang dapat membantu pemberian beasiswa bagi siswa atau siswi sekolah dasar (SD), Sekolah Lanjutan Tingkat Atas (SLTP), dan Sekolah Lanjutan Tingkat Atas (SLTA). Dalam penelitian ini didapat kriteria penentu utama beasiswa yaitu nilai raport, jika nilai raportnya tidak ada maka dengan melihat data peringkat, jika peringkat 1 maka dapat beasiswa, tetapi jika tidak ada peringkat tidak mendapat beasiswa. Atribut yang lainnya tidak termasuk dalam Pemilihan beasiswa. Penelitian ini melakukan pengujian dengan menggunakan confusion matrix dan kurva ROC/AUC (Area Under Cover), dengan hasil 69,73%.
- (3) Penelitian ini dilakukan oleh **Erfan Hasmin, Sitti Aisa dengan judul “Penerapan Algoritma C4.5 Untuk Pemilihan Penerima Beasiswa”** dari hasil penelitian maka bisa ditarik kesimpulan bahwa penerapan algoritma C.45 untuk rekomendasi penerima beasiswa sudah memiliki akurasi yang cukup baik dimana penyelesaiannya telah memiliki aturan yang telah dilakukan proses klasifikasi dengan datamining. Dengan demikian, aplikasi ini membantu bagian kemahasiswaan dalam seleksi calon penerima beasiswa di tahun selanjutnya
- (4) Penelitian ini dilakukan oleh **Nani Nurani, Sufajar Bustianto, Aswan Supriadi Sunge dengan judul “Penerapan Data Mining C4.5 Untuk Beasiswa Di SDIT AZZAHRA Cibitung “** berdasarkan pengujian data siswa Tahun Ajaran 2018-2019 diperoleh perhitungan klasifikasi data mining dengan model decision tree menggunakan algoritma C 4.5 yang dievaluasi dengan confusion matrix menghasilkan angka accuracy 91,64% $\pm$ 6,67%, class precision 96,33%, dan class recall 87,67% menggunakan rapidMiner. Dan berdasarkan pengujian data siswa Tahun Ajaran 2019-2020 diperoleh perhitungan klasifikasi data mining dengan model decision tree menggunakan

algoritma C 4.5 yang dievaluasi dengan confusion matrix menghasilkan angka accuracy 93,65% +/-6,91%, class precision 62,50%, dan class recall 50,00% menggunakan rapidMiner.

- (5) Penelitian ini dilakukan oleh **Muhammad Arif Rahman dengan judul “Algoritma C4.5 Untuk Menentukan Mahasiswa Penerima Beasiswa Kartu Indonesia Pintar Kuliah”** dari penelitian dan pengujian yang telah dilakukan dapat disimpulkan sebagai berikut : Bahwa data mining classification dengan menggunakan metode pohon keputusan dengan Algoritma C45 untuk menentukan mahasiswa penerima beasiswa pada Program Pascasarjana IAIN Raden Intan Lampung dapat dilakukan. Sejumlah kelebihan dalam penggunaan Algoritma C45 dalam membangun pohon keputusan penerima beasiswa adalah kemampuannya menangani data kontinu maupun data nominal, mengingat bahwa hampir seluruh atribut kriteria penerima beasiswa yang digunakan bertipe data kontinu. Dari sampel yang diambil yaitu sebanyak 40 (empat puluh mahasiswa) calon penerima beasiswa pada Program Pascasarjana IAIN Raden Intan Lampung dapat dilihat pada pohon keputusan yaitu sebanyak 18 (delapan belas) mahasiswa yang tidak layak menjadi penerima beasiswa karena mempunyai IPK <3,00, kemudian sebanyak 8 (delapan) mahasiswa yang tidak layak menjadi penerima beasiswa karena mempunyai masa kerja <5 tahun dan pekerja Non PNS, sehingga dihasilkan sebanyak 14 (empat belas) mahasiswa yang layak menjadi penerima beasiswa karena telah memenuhi kriteria penerima beasiswa dari segi IPK, pekerjaan dan masa kerja yang telah ditentukan sebelumnya.
- (6) Penelitian ini dilakukan oleh **Adi Suradi dengan judul “Penerapan Data Mining Untuk Menentukan Rekomendasi Beasiswa Ditingkat SMP Dengan Metode Algoritma C4.5”** sebelum dilakukan penelitian prosedur pengolahan data yang dilakukan meliputi kegiatan pengumpulan data, pengelompokan, pencocokan data dengan biodata siswa, perkiraan siswa penerima, dan menyusun laporan. Sehingga pemberian beasiswa cukup lama yaitu kurang lebih 3 sampai 4 minggu dengan adanya metode Algoritma C4.5 pemberian beasiswa bisa dilakukan selama waktu 3 – 4 hari. Berdasarkan hasil pengujian metode pohon keputusan terhadap data siswa dapat disimpulkan bahwa, penerapan pohon keputusan terhadap data siswa SMPN 2 Cikarang Selatan memiliki tingkat akurasi yang cukup baik dalam menyelesaikan kalsifikasi rekomendasi beasiswa miskin. Metode pohon keputusan merupakan metode yang cukup sesuai untuk menyelesaikan studi

kasus dalam pemilihan siswa yang mendapatkan rekomendasi Beasiswa Miskin

- (7) Penelitian ini dilakukan oleh **Arief Prajoko , Rahmat Widiya Sembiring , Saifullah dengan judul “Penerapan Algoritma C4.5 Dalam Klasifikasi Penerima Kartu Keluarga Sejahtera (KKS)”** Berdasarkan seluruh hasil tahapan penelitian yang telah dilakukan pada Penerapan Algoritma C4.5 dalam mengklasifikasi masyarakat yang berhak menerima bantuan Kartu Keluarga Sejahtera (KKS) secara layak dapat disimpulkan sebagai berikut:
- (a) Bahwa C2 (Penghasilan Kepala Keluarga) memiliki gain tertinggi yaitu 0.46880137.
  - (b) Permasalahan menentukan masyarakat yang layak dan tidak layak dapat diselesaikan menggunakan teknik data mining, yaitu dengan Algoritma C4.5. menghasilkan 8 (Delapan) rules dan Tingkat akurasi yang dihasilkan oleh metode tersebut adalah 100,00%.
  - (c) Dengan adanya penerapan data mining algoritma C4.5 diharapkan mampu memberikan solusi kepada perangkat desa Nagori Mekar Rejo dalam menentukan masyarakat yang berhak dalam menerima bantuanKartu Keluarga Sejahtera (KKS):
- (8) Penelitian ini dilakukan oleh **Jajam Haerul Jaman , Novia Indriyani Puji Astuti dengan judul “Penerapan Metode Data Mining C4.5 Untuk Pemilihan Beasiswa Ditingkat SD”** pengujian Algoritma C4.5 diukur berdasarkan nilai accuracy, precision, dan error rate. Pengujian kinerja menerapkan information gain dengan 7 kriteria yaitu: JTO, nilai, jarak, pekerjaan, JPO, penerima kartu indonesia pintar, penerima KPS menghasilkan nilai akurasi sebesar 93,58%, precision sebesar 95,5%, error rate 6,14% dikarenakan memiliki nilai ketepatan prediksi secara benar tertinggi dan persentase dari total data yang diprediksi secara salah terendah. Hasil ini merupakan model terbaik jika dibandingkan dengan skenario yang lain untuk memprediksi penerima beasiswa di SDN Karawang Kulon III karena Semua algoritma klasifikasi berusaha membentuk model yang memiliki akurasi tinggi (Error Rate yang rendah)
- (9) Penelitian ini dilakukan oleh **Erni Ermawati , Taufik Hidayatulloh dengan judul “Penerapan Algoritma C4.5 Pada Sistem Penunjang Keputusan Pemilihan Penerima Raskin”** berdasarkan pembahasan pada bab-bab sebelumnya, dapat diambil kesimpulan dari pembuatan aplikasi penerapan algoritma C4.5 untuk Pemilihan penerima RASKIN berbasis web Dalam upaya membantu Kepala Seksi Kesejahteraan Rakyat (KASIKESRA) dan Kepala

Desa (KADES) serta para perangkat desa lainnya yang bertugas dalam menentukan penerima RASKIN, program SPK ini dapat menjadi alternatif pemecahan masalah, diantaranya yaitu .

- (a) Sistem pendukung keputusan dibuat agar membantu para pengguna khususnya kepala seksi kesejahteraan rakyat (KASIKESRA) dan kepala desa (KADES) serta para perangkat desa dalam pengambilan keputusan, tanpa harus melakukan rempug desa yang berulang dan berkepanjangan.
  - (b) Sistem pendukung keputusan ini dirancang dalam bentuk program untuk kelayakan penerima RASKIN berdasarkan kriteria yang ada. Algoritma C4.5 ini dapat diterapkan untuk memprediksi masyarakat penerima RASKIN dengan entropy terbesar 0,9954 yakni petani dan gain terbesarnya adalah bahan bakar memasak 0,5000 serta dilengkapi dengan hasil uji kelayakan dengan begitu para pengguna akan lebih cepat dan tepat sasaran dalam menentukan penerima RASKIN
- (10) Penelitian ini dilakukan oleh **Andro Paul V. M., Indra Gunawan, Bahrudi Efendi Damanik, Iin Parlina dan Widodo Saputra dengan judul” Penerapan Data Mining Menggunakan Algoritma C4.5 Dalam Menentukan Kelayakan Penerima Bantuan Bedah Rumah Pada Desa Tiga Golok)”** Penelitian ini dapat disimpulkan bahwa analisa dan perhitungan menggunakan algoritma ini sangat membantu proses Pemilihan keputusan untuk menentukan masyarakat yang layak menerima bantuan bedah rumah dan masyarakat yang tidak layak menerima bantuan bedah rumah. Karena dengan perhitungan ini dapat mempersingkat waktu dalam pengambilan keputusan. Pohon Keputusan (decision tree) memberikan gambaran secara jelas atribut mana yang merupakan prioritas sebagai penentu keputusan karena dengan perhitungan Decision tree ini, akan menentukan atribut yang paling utama serta kelas akar dan kelas daun. Dengan penerapan data mining menggunakan algoritma C4.5 maka dihasilkan suatu keputusan keluarga yang mendapatkan bantuan bedah rumah menjadi tepat sasaran.

**Tabel 2.7 Penelitian Rujukan**

No	Nama Peneliti	Judul	Jurnal	Kontribusi
1	Muhammad Yunus, Hanandriya Ramadha,	Penerapan Metode Data Mining C4.5 Untuk Pemilihan Beasiswa Kartu Indonesia Pintar (KARTU INDONESIA PINTAR)	Paradigma - Jurnal Komputer dan Informatika Vol 23 no 2 2021	Kontribusi penelitian ini adalah mendapatkan cara menentukan pohon keputusan

	Dimas Rizki Aji , Agus Yulianto (2021)			tentang Pemilihan beasiswa
2	Agustiena Merdeka wa(2018)	Sistem Pendukung Keputusan Penerimaan Beasiswa Menggunakan Algoritma C4.5	Jurnal Pendidikan Teknologi dan Kejuruan vol 15 hal 113-123 2018	Kontribusi penelitian ini adalah pengambilan uji hasil menggunakan confussion matrix dan pengambilan metode Algoritma C4.5  Kelemahan penelitian ini adalah belum adanya aplikasi untuk penghitungan
3	Erfan Hasmin, Sitti Aisa(2019)	Penerapan Algoritma C4.5 Untuk Pemilihan Penerima Beasiswa Mahasiswa	CogITo Smart Journal vol 5 hal 308 2019	Kontribusi penelitian ini adalah pengambilan metode algoritma C4.5 dan cara perhitungan melalui aplikasi  Kelemahan penelitian ini adalah kesimpulan yang ditulis oleh peneliti masih kurang
4	Nani Nurani, Sufajar Bustianto, Aswan Supriadi Sunge (2016)	Penerapan Data Mining C4.5 Untuk Beasiswa Di SDIT AZZAHRA Cibitung	Jurnal data mining	Kontribusi penelitian ini adalah pengambilan uji hasil menggunakan confussion matrix dan pengambilan metode menggunakan algoritma C4.5
5	Muhammad Arif Rahman(2018 )	Algoritma C4.5 Untuk Menentukan Mahasiswa Penerima Beasiswa KARTU INDONESIA PINTAR Kuliah	Jurnal Teknologi Informasi Magister Darmajaya vol 1 118-128	Kontribusi penelitian ini adalah pengambilan metode algoritma C4.5 dan cara perhitungan C4.5  Kelemahan penelitian ini adalah tidak terdapat nilai akurasi

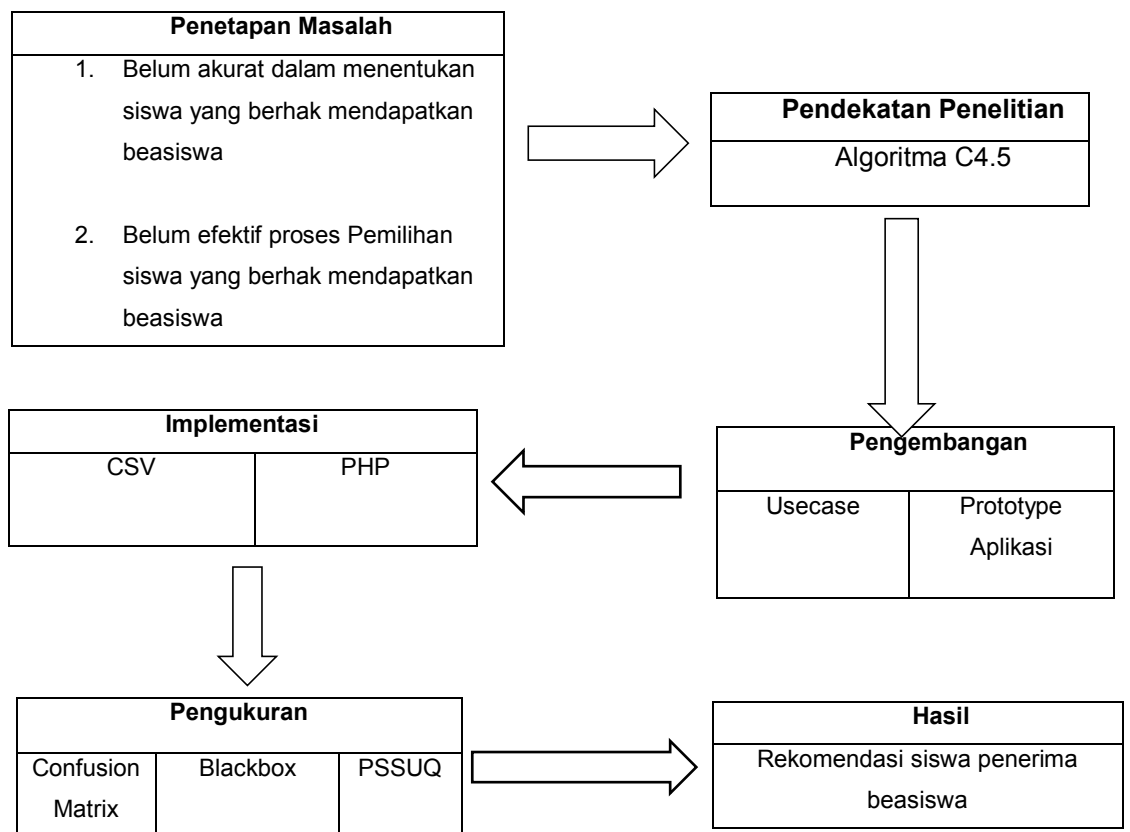
6	Adi Suradi(2018)	Penerapan DataMining Untuk Menentukan Rekomendasi Beasiswa Ditingkat SMP Dengan Metode Algoritma C4.5	Jurnal Teknik Informatika pelita bangsa	Kontribusi penelitian ini adalah pengambilan metode algoritma C4.5 dan cara perhitungan C4.5  Kelemahan penelitian ini adalah belum ada perhitungan algoritma c4.5 munggunakan aplikasi
7	Arief Prajoko , Rahmat Widiya Sembiring , Saifullah(2021 )	Penerapan Algoritma C4.5 Dalam Klasifikasi Penerima Kartu Keluarga Sejahtera (KKS)	Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika) vol 6 171	Kontribusi penelitian ini adalah pengambilan metode algoritma C4.5 dan cara penghitungan confusion matrix  Kelemahan penelitian ini adalah belum adanya aplikasi untuk perhitungan
8	Jajam Haerul Jaman, Novia Indriyani Puji Astuti(2020)	Penerapan Metode Data Mining C4.5 Untuk Pemilihan Beasiswa Ditingkat	Techno Xplore vol 3 25-29	Kontribusi penelitian ini adalah pengambilan uji hasil menggunakan confusion matrix dan pengambilan metode menggunakan algoritma C4.5  Kelemahan penelitian ini adalah tidak disampaikannya langkah langkah perhitungan C4.5
9	Erni Ermawati, Taufik Hidayatulloh	Penerapan Algoritma C4.5 Pada Sistem Penunjang Keputusan Pemilihan Penerima Raskin	Seminar Nasional Ilmu Pengetahuan dan Teknologi Komputer Nusa Mandiri 123 - 134	Kontribusi penelitian ini adalah pengambilan uji hasil menggunakan confusion matrix dan pengambilan metode menggunakan algoritma C4.5



10	Paul V. M., Indra Gunawan, Bahru Efendi Damanik, lin Parlina dan Widodo Saputra (2021)	Penerapan Data Mining Menggunakan Algoritma C4.5 Dalam Menentukan Kelayakan Penerima Bantuan Bedah Rumah Pada Desa Tiga Golok	Teknois : Jurnal Ilmiah Teknologi Informasi dan Sains Vol 8 17 - 28	Kontribusi penelitian ini adalah pengambilan uji hasil menggunakan confussion matrix dan pengambilan metode menggunakan algoritma C4.5
----	---	--	--	---

Berdasarkan uraian maka terdapat kontribusi dari penelitian yaitu adanya kesamaan variabel dan kesamaan metode yang digunakan yaitu penerapan metode algoritma C4.5 atau metode *Decision tree*. Kontribusi lainnya adalah bidang yang digunakan adalah bidang yang sama yaitu pemilihan siswa penerima beasiswa.

### C. Kerangka Pemikiran



Gambar 2.2 Kerangka Pemikiran

Kerangka pemikiran pada gambar 2.2 dapat dijelaskan sebagai berikut :

- (1) Penetapan masalah mencakup fenomena yaitu Belum Akurat dan Efektif dalam menentukan siswa yang berhak mendapatkan beasiswa
- (2) Pendekatan penelitian yaitu metode yang ditetapkan atau digunakan yaitu C4.5
- (3) Pengembangan yaitu tahap Prototype Aplikasi dan membuat diagram Usecase.
- (4) Implementasi yaitu tahap menerapkan metode C4.5 kedalam sistem aplikasi dan pembuatan coding lalu melakukan uji hasil dari sistem tersebut.
- (5) Pengukuran yaitu melakukan pengujian terhadap metode C4.5 yang ditetapkan ke sistem menggunakan confusion matrix, blackbox, dan PSSUQ.
- (6) Hasil yaitu sistem informasi menampilkan hasil rekomendasi siswa yang berhak mendapatkan beasiswa tersebut.

#### **D. Hipotesis**

Berdasarkan permasalahan yang dihadapi yaitu rekomendasi siswa penerima beasiswa kartu indonesia pintar ditingkat sekolah belum efektif dan belum akurat, maka perlu suatu cara untuk mengatasi permasalahan tersebut. Dalam teori data mining ada beberapa metode yang dapat melakukan pengklasifikasian untuk memprediksi suatu peluang dimasa depan berdasarkan dari pengalaman dimasa sebelumnya, diantaranya adalah metode algoritma C4.5. teori klasifikasi yang digunakan C4.5 yang akan menghasilkan hasil prediksi yang telah ditentukan nilai probabilitas akhir dari setiap kelas atau label dan melihat hasil akhir yang terbesar dari setiap kelas. Penelitian yang telah dilakukan oleh (Muhammad Yunus, 2021) berjudul “Penerapan Metode Data Mining Untuk Pemilihan Penerima Kartu Indonesia Pintar (KIP)” atribut yang digunakan oleh Muhammad Yunus yaitu Absensi, Nilai, Tanggungan Orang Tua, Bangunan Rumah, Pendapatan Orang Tua. Berdasarkan hal tersebut, maka dapat ditetapkan hipotesis pada penelitian ini yaitu metode Algoritma C4.5 diduga dapat memberikan rekomendasi siswa penerima beasiswa kartu indonesia pintar ditingkat sekolah.