

**DECISSION TREE TECHNIQUE DALAM
MENENTUKAN PENJURUSAN SISWA MENENGAH
KEJURUAN**

SKRIPSI



**Oleh:
Ika Purnama Sari
161510060**

**FAKULTAS TEKNIK DAN KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
UNIVERSITAS PUTERA BATAM
TAHUN 2020**

**DECISSION TREE TECHNIQUE DALAM
MENENTUKAN PENJURUSAN SISWA MENENGAH
KEJURUAN**

SKRIPSI

**Untuk memenuhi salah satu syarat
guna memperoleh gelar sarjana**



**Oleh:
Ika Purnama Sari
161510060**

**FAKULTAS TEKNIK DAN KOMPUTER
PROGRAM STUDI SISTEM INFORMASI
UNIVERSITAS PUTERA BATAM
TAHUN 2020**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Ika Purnama Sari
NPM : 161510060
Fakultas : Teknik dan Komputer
Program Studi : Sistem Informasi

Menyatakan bahwa “Skripsi” yang saya buat dengan judul:

DECISSION TREE TECHNIQUE DALAM MENENTUKAN PENJURUSAN SISWA MENENGAH KEJURUAN

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, didalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun.

Batam, 24 Juli 2020



Ika Purnama Sari
161510060

**DECISSION TREE TECHNIQUE DALAM MENENTUKAN
PENJURUSAN SISWA MENENGAH KEJURUAN**

SKRIPSI
Untuk memenuhi salah satu syarat
guna memperoleh gelar Sarjana

Oleh:
Ika Purnama Sari
161510060

Telah disetujui oleh Pembimbing pada tanggal
seperti tertera di bawah ini

Batam, 24 Juli 2020



Rika Harman, S.Kom., M.SI
Pembimbing

ABSTRAK

Banyak kasus dijumpai bahwa pemilihan jurusan yang tidak sesuai dengan kemampuan, kepribadian, minat dan bakat dapat mempengaruhi siswa dalam mengikuti pembelajaran. Penggunaan pendekatan *algoritma* klasifikasi data mining akan diterapkan untuk menentukan jurusan dalam bidang studi yang akan diambil oleh siswa, sehingga siswa tidak salah dalam memilih jurusan yang akan di tempuh selama belajar pada Sekolah Menengah Kejuruan (SMK). *Algoritma C4.5* digunakan untuk menentukan jurusan yang akan diambil oleh siswa sesuai dengan latar belakang, minat dan kemampuannya sendiri. Teknik *datamining* dengan metode klasifikasi algoritma C4.5 dalam menganalisa dalam menentukan penjurusan siswa SMK Putra Jaya School dengan data siswa yang tersedia dapat menemukan *rules* atau aturan keputusan untuk digunakan sebagai acuan atau standar dalam menilai tingkat penjurusan siswa yang diterima. Parameter pemilihan jurusan adalah test buta warna, kesehatan dan wawancara. Hasil pengujian dan evaluasi menunjukkan bahwa *Algoritma Decision Tree C4.5* akurat diterapkan untuk penentuan kesesuaian jurusan siswa pada Sekolah Menengah Kejuruan (SMK). Analisa pohon keputusan (*decission tree*) yang dihasilkan oleh perhitungan algoritma C4.5 menunjukkan bahwa variabel yang memiliki *gain* tertinggi atau yang merupakan faktor utama yang mempengaruhi tingkat penjurusan siswa SMK Putra Jaya School adalah Buta Warna, Kesehatan, dan Wawancara. Pengujian menggunakan software WEKA 3.9.2 juga menghasilkan hasil yang sama yaitu variabel yang menjadi faktor utama yang mempengaruhi tingkat penjurusan siswa adalah Tes Kesehatan.

Kata Kunci: Data Mining, Penentuan Jurusan Siswa, *Decision Tree C45*.

ABSTRACT

Many cases are found that the selection of majors that are not in accordance with the abilities, personality, interests and talents can influence students in participating in learning. The use of the data mining classification algorithm approach will be applied to determine the majors in the field of study to be taken by students, so students are not wrong in choosing the majors that will be taken while studying at the Vocational High School (SMK). C4.5 algorithm is used to determine the majors to be taken by students according to their own backgrounds, interests and abilities. The data mining technique using C4.5 algorithm classification method in analyzing in determining the direction of SMK Putra Jaya School students with available student data can find the rules or decision rules to be used as a reference or standard in assessing the direction of students accepted. Major selection parameters are color blindness, health test and interview. Test and evaluation results show that the Decision Tree C4.5 Algorithm is accurately applied to determine the suitability of student majors at Vocational High Schools (SMK). Analysis of the decision tree (decision tree) produced by the C4.5 algorithm calculation shows that the variable that has the highest gain or which is a major factor influencing the majors of SMK Putra Jaya School students is Color Blindness, Health, and Interview. Testing using WEKA 3.9.2 software also produces the same results, which are the variables that are the main factors that influence the level of student majors is the Health Test.

Keywords: Data Mining, Determination of Student Majors, Decision Tree C4.5.

KATA PENGANTAR

Puji syukur penulis ucapkan kepada Tuhan YME yang telah melimpahkan segala rahmat dan karuniaNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Sistem Informasi Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

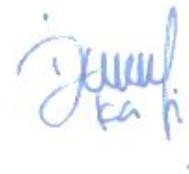
1. Rektor Universitas Putera Batam Ibu Dr. Nur Elfi Husda, S.Kom., M.SI.;
2. Ketua Program Studi Sistem Informasi Bapak Muhammat Rasid Ridho, S.Kom., M.SI.;
3. Bapak Muhammat Rasid Ridho, S.Kom., M.SI., selaku pembimbing akademik pada program studi Sistem Informasi di Universitas Putera Batam.
4. Bapak Rika Harman, S.Kom., M.SI. selaku pembimbing Skripsi pada Program Studi Sistem Informasi Universitas Putera Batam;
5. Dosen dan Staff Universitas Putera Batam;
6. Kedua orang tua penulis yang selalu mendoakan dan

menyemangati penulis hingga penulisan skripsi ini selesai;

7. Keluarga penulis yang selalu mendoakan dan memberikan motivasi kepada penulis agar penelitian ini selesai tepat waktu;
8. Teman-teman seperjuangan Sistem Informasi angkatan 2016 yang bersedia membagi ilmunya dan *sharing* pendapat dalam rangka pembuatan skripsi ini;
9. Teman-teman alumni maupun teman kalangan luar: Eko Tribuana Putra, Fanisa arifka, Siska rahayu, Silvia maulana, memberikan semangat dan motivasi dalam penyelesaian skripsi ini.
10. Semua pihak yang telah bersedia meluangkan waktu, tenaga dan pikirannya dalam memberikan data dan informasi selama penulis membuat skripsi yang tidak dapat penulis sebutkan satu-persatu.

Semoga Allah SWT membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Amin.

Batam, 24 Juli 2020



Ika Purnama Sari

DAFTAR ISI

| | |
|--|-------------|
| HALAMAN SAMPUL | i |
| HALAMAN JUDUL | ii |
| SURAT PERNYATAAN ORISINALITAS | iii |
| HALAMAN PENGESAHAN | iv |
| ABSTRAK | v |
| ABSTRACT | vi |
| KATA PENGANTAR | vii |
| DAFTAR ISI | ix |
| DAFTAR GAMBAR | xii |
| DAFTAR TABEL | xiii |
| DAFTAR RUMUS | xiv |
| BAB I PENDAHULUAN | 1 |
| 1.1. Latar Belakang | 1 |
| 1.2. Identifikasi Masalah | 6 |
| 1.3. Batasan Masalah..... | 6 |
| 1.4. Rumusan Masalah | 7 |
| 1.6. Manfaat Penelitian..... | 9 |
| 1.6.1. Manfaat Teoritis | 9 |
| 1.6.2. Manfaat Praktis | 10 |
| BAB II TINJAUAN PUSTAKA | 11 |
| 2.1. Teori | 11 |
| 2.1.1. Teori Umum | 11 |
| 2.1.2. Teori Khusus | 22 |

| | | |
|--|---|-----------|
| 2.2. | Kerangka Pemikiran | 23 |
| 2.3. | Hipotesis Penelitian | 24 |
| 2.4. | Penelitian Terdahulu..... | 25 |
| BAB III METODE PENELITIAN | | 28 |
| 3.1. | Desain Penelitian | 28 |
| 3.2. | Objek Penelitian | 30 |
| 3.2.1. | Lokasi Penelitian..... | 31 |
| 3.2.2. | Jadwal Penelitian..... | 32 |
| 3.3. | Populasi dan Sampel | 32 |
| 3.3.1. | Populasi..... | 33 |
| 3.3.2. | Sampel..... | 33 |
| 3.4. | Variabel Penelitian | 33 |
| 3.5. | Teknik Pengumpulan Data | 33 |
| 3.6. | Model Penelitian..... | 34 |
| BAB IV HASIL DAN PEMBAHASAN | | 36 |
| 4.1 | Hasil Penelitian | 36 |
| 4.2 | Pembahasan..... | 37 |
| 4.3 | Analisa Data | 41 |
| 4.3.1. | Pembersihan data | 41 |
| 4.3.2. | Seleksi data..... | 41 |
| 4.3.3. | Transformasi data..... | 41 |
| 4.3.4. | Pohon keputusan | 47 |
| 4.3.5. | Pengujian menggunakan Software WEKA | 57 |

| | |
|---|-----------|
| BAB V KESIMPULAN DAN SARAN | 63 |
| 5.1. Kesimpulan..... | 63 |
| 5.2. Saran..... | 64 |

DAFTAR PUSTAKA

Lampiran 1. Pendukung Penelitian

Lampiran 2. Daftar Riwayat Hidup

Lampiran 3. Surat Keterangan Penelitian

DAFTAR GAMBAR

| | |
|---|----|
| Gambar 1. 1 Data Penjurusan siswa SMK Putra Jaya School Th.2019/2020 | 4 |
| Gambar 2. 1 Pengelompokan Teknik Klasifikasi | 17 |
| Gambar 2. 2 Konsep dasar pohon keputusan | 18 |
| Gambar 2. 3 Kerangka pemikiran..... | 24 |
| Gambar 3. 1 Desain Penelitian | 28 |
| Gambar 3. 2 Lokasi Penelitian | 31 |
| Gambar 4. 1 Hasil Perhitungan <i>Node 1</i> | 52 |
| Gambar 4. 2 Pohon Keputusan <i>Node 2</i> | 56 |
| Gambar 4. 3 Format Data <i>Microsoft Excel</i> | 57 |
| Gambar 4. 4 Data dalam Notepad ++ | 58 |
| Gambar 4. 5 Halaman Awal <i>Software WEKA</i> | 58 |
| Gambar 4. 6 Membuka Data Pada Menu <i>Explorer</i> | 59 |
| Gambar 4. 7 Pemilihan Tabel Pada <i>Software WEKA</i> | 59 |
| Gambar 4. 8 Pemilihan <i>Model Classify</i> | 60 |
| Gambar 4. 9 <i>Classify Output</i> pada <i>WEKA</i> | 61 |
| Gambar 4. 10 <i>Visualize Tree</i> Pada <i>WEKA</i> | 61 |
| Gambar 4. 11 Hasil <i>Visualize Tree WEKA</i> | 62 |

DAFTAR TABEL

| | |
|---|----|
| Tabel 3. 1 Lokasi Penelitian | 32 |
| Tabel 4. 1 Format Data Siswa | 38 |
| Tabel 4. 2 Klasifikasi Atribut Buta Warna..... | 42 |
| Tabel 4. 3 Klasifikasi Kesehatan | 42 |
| Tabel 4. 4 Klasifikasi Wawancara..... | 43 |
| Tabel 4. 5 Data Pra-proses siswa..... | 44 |
| Tabel 4. 6 Hasil Perhitungan <i>Node 1</i> | 51 |
| Tabel 4. 8 Perhitungan Hasil <i>Node 2</i> | 55 |

DAFTAR RUMUS

| | |
|---|----|
| Rumus 2. 1 Perhitungan Gain | 21 |
| Rumus 2. 2 Perhitungan Entropy | 21 |

BAB I

PENDAHULUAN

1.1. Latar Belakang

Berdasarkan pendapat yang dikemukakan oleh (Sambani & Nuraeni, 2017) Mengacu pada Undang-Undang Republik Indonesia No 20 tahun 2003 perihal Sistem Pendidikan Nasional (Sisdiknas) menentukan bahwa tujuan dari Pendidikan Nasional yaitu untuk perkembangan kemampuan atau potensi peserta didik supaya jadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa dan berbudi pekerti luhur, berakhlak mulia, sehat, berilmu, dalam rangka mencerdaskan kehidupan bangsa. Sistem pendidikan nasional bisa menjamin pemerataan kesempatan pendidikan, meningkatkan mutu serta efisiensi dan juga relevansi manajemen pendidikan guna menghadapi tantangan sesuai dengan ketentuan perubahan kehidupan lokal, nasional, juga global sehingga harus dilakukan perbaikan pendidikan secara terencana, terarah, dan berkelanjutan.

Sedangkan pendapat lainnya (Rasyid, 2017) Pendidikan adalah sebuah sistem yang terdiri dari beberapa faktor yang saling berhubungan dan juga kompleks namun mempunyai tujuan besar yang sama yaitu menghasilkan sumber daya yang bermutu atau berkualitas. Pendidikan memberikan pengaruh terhadap kualitas SDM (Sumber Daya Manusia) yang sangat menentukan nasib bangsa. SMK (Sekolah Menengah Kejuruan) merupakan suatu struktur satuan pendidikan formal yang menyelenggarakan pendidikan kejuruan pada jenjang pendidikan menengah sebagai lanjutan dari SMP/MTs. Pendidikan Kejuruan merupakan

tahapan dari sistem pendidikan yang mempersiapkan seseorang supaya dapat bekerja pada suatu kelompok pekerjaan atau satu bidang pekerjaan dari pada bidang-bidang pekerjaan lain nya. berdasarkan pengertian bahwa di setiap jurusan bisa dipelajari dan dipahami lebih mendalam lagi dan juga bisa diterapkan didunia kerja.

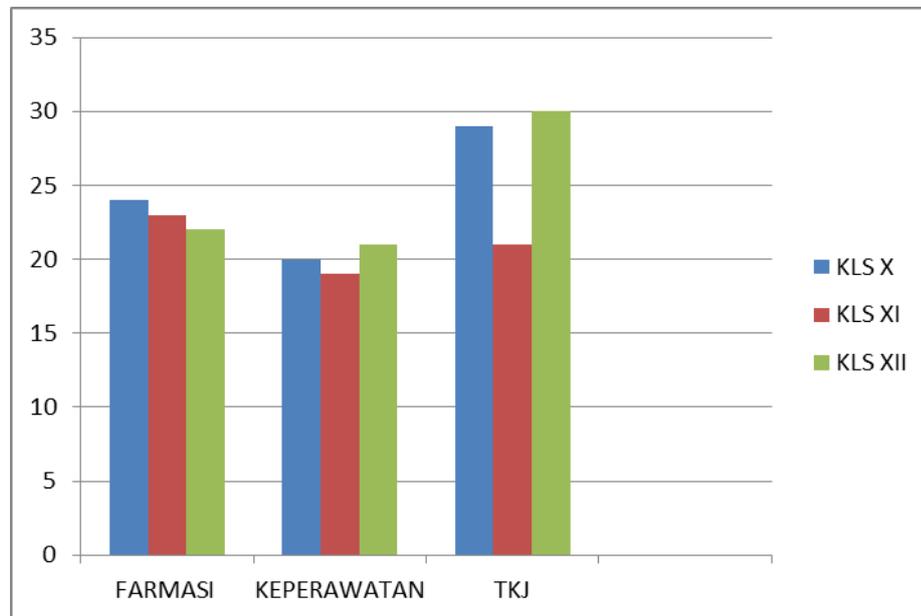
Mengutip pendapat dari (Sucipto et al., 2017) Siswa yang ingin melanjutkan SMA akan lebih mudah untuk menentukan suatu jurusan, sedangkan siswa yang ingin melanjutkan ke SMK akan lebih sulit karena banyak pilihan jurusan yang ada. Harus diketahui dalam tahap menentukan jurusan di SMK itu merupakan suatu hal yang sangat penting, karena nantinya siswa akan diberikan pembelajaran sesuai dengan jurusan yang sudah di pilih sesuai dengan minat dan bakat. Oleh karena itu, pihak sekolah harus mengetahui pola minat siswa yang sesuai jurusan yang sudah disediakan guna menghindari kesalahan dalam pemilihan jurusan. kebijakan tersebut juga nantinya mempermudah pihak sekolah untuk memilih jurusan mana yang tepat bagi calon siswa tersebut.

Seiring dengan perkembangan peradapan keilmuan proses pemilihan jurusan tersebut dapat dipermudah, hal tersebut memang tidak seperti membalikkan telapak tangan tetapi butuh pengetahuan dan proses lebih lanjut yang harus dilakukan. Menurut (Sambani & Nuraeni, 2017) pengkajian pola bisa dilaksanakan dengan menggunakan metode atau teknik klasifikasi data yang didapatkan dari proses *data mining*. *Data mining* biasanya dilakukan karena adanya sejumlah besar data yang bisa di gunakan untuk memperoleh hasil pengetahuan (*Knowledge*) dan pola yang digunakan untuk proses bisnis pada

perusahaan. Salah satu tahapan yang ada di data mining yaitu bagaimana meneliti data yang sudah ada guna membentuk sebuah model, kemudian model tersebut digunakan untuk mengenali pola data yang lain yang tidak terdapat pada basis data yang sudah disimpan. Setiap tahun, meningkatnya jumlah data siswa di sekolah SMK. Data tersebut tersimpan dalam *hard file* seperti hasil *print out* data, catatan buku dan dalam bentuk *soft file* di komputer. Dengan bertambahnya jumlah data yang cepat dan banyak yang dikumpulkan dan disimpan tiap tahunnya jauh melampaui kemampuan manusia dalam menganalisa data jika tidak memakai metode yang benar. Akan mengakibatkan data yang dikumpulkan jadi tidak bermanfaat. Untuk mengatasi hal tersebut diperlukan suatu metode dalam *data mining* yang bisa merubah sesuatu yang tidak bermanfaat menjadi sesuatu yang bisa dimanfaatkan juga bisa memberikan suatu informasi yang sangat penting.

Siswa yang akan mendaftar ke SMK Putra Jaya School Batam mengajukan pilihan jurusan dengan mengisi formulir yang sudah disediakan oleh pihak sekolah. Tes yang biasanya dilakukan pada saat penyeleksian calon siswa baru di SMK di antaranya tes kesehatan, tes wawancara, dan tes buta warna. Kemudian dari hasil tes di olah lagi guna memperoleh nilai yang memenuhi syarat untuk pilihan jurusan pertama. Jika nilai standar yang awal tidak tercukupi, maka dicocokkan untuk pilihan jurusan yang kedua. Jika sesuai dengan standar nilai yang kedua, maka siswa di masukkan ke pilihan jurusan yang kedua. Dan jika tidak sesuai dengan kedua pilihan jurusan tersebut, maka siswa dinyatakan tidak lulus atau tidak di terima di SMK tersebut. Demikian grafik data penjurusan siswa

di Sekolah Menengah Kejuruan (SMK) Putra Jaya School di Kota Batam, sebagai berikut.



Gambar 1. 1 Data Penjurusan siswa SMK Putra Jaya School Th.2019/2020

Sumber : (SMK Putra Jaya School, 2020)

Berdasarkan gambar pada grafik data penjurusan siswa Sekolah Menengah Kejuruan (SMK) Putra Jaya School tahun 2019, pada jurusan Farmasi di kelas X mengalami peningkatan dengan jumlah siswa 24 orang, dan pada kelas XI jurusan Farmasi mengalami penurunan dengan jumlah siswa 23 orang, kemudian pada kelas XII jurusan Farmasi mengalami penurunan lagi dengan jumlah siswa 22 orang. Jurusan Keperawatan pada kelas X yang memilih hanya 20 orang, sedangkan pada kelas XI yang memilih jurusan Keperawatan adalah 19 orang, dan pada kelas XII penjurusan Keperawatan mengalami peningkatan dengan jumlah siswa 21 orang. Dan yang terakhir yaitu jurusan Teknik Komputer dan Jaringan

pada kelas X mengalami peningkatan sebanyak 29 siswa, pada kelas XI jurusan tersebut mengalami penurunan sebanyak 21 siswa, dan pada kelas XII penjurusan tersebut mengalami kenaikan sebanyak 30 siswa.

Masalah yang di hadapi yaitu bagaimana menentu kan keputusan siswa yang terbaik yang sesuai dengan minat juga bakatnya. Dengan data yang begitu banyak dan waktu yang sangat singkat, sarana dan prasarana yang kurang memadai, banyaknya siswa yang mendaftar harus sesuai dengan jumlah pendidik yang disediakan di sekolah, panitia harus mengambil keputusan yang tepat. Panitia mengolah data dengan menggunakan *Microsoft Excel*, dan belum tersedianya program khusus untuk perhitungan sehingga dalam proses tersebut membutuhkan waktu yang sangat lama. Ketika memakai *Microsoft Excel* pihak sekolah merasa kesulitan dalam pengolahan data siswa baru. Berdasarkan masalah tersebut maka penelitian ini di lakukan dengan tujuan untuk membantu pihak sekolah dalam mempercepat proses penempatan jurusan yang sesuai dengan kompetensi siswa. Sistem Pendukung Keputusan (SPK) merupakan suatu system berbasis komputer yang dapat membantu seseorang dalam meningkatkan kinerjanya dalam dalam pengambilan keputusan. Sistem ini merupakan Sistem Pendukung Keputusan (SPK) yang dibangun menggunakan penggabungan metode *Decision Tree Technique* algoritma C.45 (Rusdiansyah, 2017).

1.2. Identifikasi Masalah

Berdasarkan permasalahan yang telah dijabarkan pada latar belakang penulis menyimpulkan adanya permasalahan, yaitu :

1. Sering mendapatkan komplain dari siswa yang ada tentang proses penjurusan yang ada, dalam hal ini sistem penjurusan yang tidak sesuai dengan kemampuan, serta minat siswa sehingga menimbulkan rasa ketidaknyamanan dari para siswa yang ada.
2. Tidak ada standar atau referensi dalam penentuan pemilihan jurusan secara jelas berdasarkan tingkat kemampuan dari masing-masing siswa yang baru masuk.
3. Tidak ada pemanfaatan dari data arsip sekolah yang lama sebagai bahan acuan untuk menentukan jurusan bagi siswa yang baru tentang patokan dalam penjurusan.

1.3 Batasan Masalah

Untuk lebih fokusnya permasalahan yang dibahas didalam penelitian ini maka perlu ditetapkan batasan terhadap masalah yang akan dibahas, yang antara lain adalah sebagai berikut:

1. Penelitian ini dilakukan di Sekolah Menengah Kejuruan (SMK) Putra Jaya School Kota Batam dengan menganalisa penentuan dari penjurusan siswa baru.

2. Menganalisa data penjurusan yang ada di Sekolah Menengah Kejuruan (SMK) Putra Jaya School dengan metode yang ada pada *data mining* menggunakan algoritma C4.5.
3. Data dari penelitian ini diambil berdasarkan kepada data pada 3 (tiga) bulan terakhir terhitung dari bulan April, Mei dan Juni tahun 2020, yang diambil langsung pada obyek penelitian.

1.4 Rumusan Masalah

Berdasarkan penjabaran yang telah dibahas diatas maka permasalahan dapat dirumuskan yang antara lain adalah sebagai berikut:

1. Bagaimanakah penerapan Data Mining khususnya teknik klasifikasi menggunakan Algoritma C4.5 dalam menganalisa penjurusan siswa pada SMK Putra Jaya School Batam?
2. Apakah teknik klasifikasi menggunakan Algoritma C4.5 dapat menemukan pola dalam penjurusan siswa yaitu berupa pohon keputusan yang dibentuk berdasarkan perhitungan *entropy* dan *Gain*?
3. Bagaimanakah mengimplementasikan teknik klasifikasi Algoritma C4.5 sebagai alat untuk menemukan pola yang digunakan dalam penjurusan siswa?

1.5 Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan, yaitu:

1. Pengujian metode Algoritma C4.5 dan klasifikasi dalam wujud penerapan dari *data mining* itu sendiri yang digunakan dalam pemilihan jurusan yang sesuai bagi siswa baru SMK Putra Jaya School Kota Batam.
2. Menerapkan *data mining* khususnya teknik klasifikasi menggunakan Algoriitma C4.5 dalam menganalisa penjurusan yang tepat pada SMK Putra Jaya School Kota Batam dan mengurangi resiko kesalahan penjurusan.
3. Untuk mendapatkan keputusan tentang referensi atau standar dalam pemilihan jurusan yang sesuai dengan keinginan dan kemampuan siswa dengan menggunakan teknik dalam *data mining* yaitu algoritma C4.5.
4. Bisa mengambil keputusan yang baik, dimana keputusan ini dapat memuaskan semua pihak yang membutuhkan tidak terkecuali para siswa dan pihak manajemen sekolah.

1.6. Manfaat Penelitian

Penelitian ini diharapkan bermanfaat bagi pembaca. Manfaat dari penelitian yang dilakukan adalah sebagai berikut:

1.6.1. Manfaat Teoritis

Manfaat teoritis yang di harapkan dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Memperkuat teori tentang algoritma C4.5 dalam menentukan penjurusan siswa di SMK Putra Jaya School Kota Batam.
2. Sebagai referensi pada penelitian-penelitian selanjutnya yang berhubungan dengan teknik *data mining* dengan metode algoritma C4.5.
3. Dapat mengembangkan keilmuan tentang *datamining* khususnya Algoritma C4.5 dalam hal mengenai data-data yang selama ini dianggap tidak berguna atau dibuang oleh pihak yang mempunyai data.
4. Membuktikan lagi kepada para peneliti selanjutnya bahwa data-data yang semuanya dapat dimanfaatkan dengan baik dan mempunyai nilai guna yang tinggi jika dimanfaatkan dengan tepat.

1.6.2. Manfaat Praktis

Selain manfaat teoritis, maka praktis yang diharapkan dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Dijadikan acuan atau pedoman standar untuk menentukan penjurusan siswa di SMK Putra Jaya School Kota Batam sehingga dapat memberikan nilai lebih bagi kesemuanya yang terlibat dalam system ini.
2. Memberikan masukan kepada sekolah SMK Putra Jaya School Kota Batam dalam menentukan penjurusan siswa perlu adanya aturan-aturan standar yang perlu dilakukan.
3. Meningkatkan kualitas sarana dan prasarana dalam menentukan penjurusan siswa pada SMK Putra Jaya School khususnya serta SMK lainnya dikota Batam pada umumnya.

BAB II

TINJAUAN PUSTAKA

2.1. Teori

Berikut merupakan penjabaran mengenai teori yang akan dikajikan penulis didalam penelitian ini, adapun teori yang dimaksud atas dua teori besar yaitu teori umum dan teori khusus, adapun uraiannya yaitu:

2.1.1. Teori Umum

Berikut merupakan penjelasan tentang teori umum yang di gunakan untuk proposal penelitian ini, teori umum ini dimaksudkan untuk memperkuat isi dari teori yang ada. Hal ini dimaksudkan supaya penelitian yang dilakukan hasilnya lebih baik. Adapun isi dari teori umum yaitu:

2.1.1.1 *Knowledge Discovery in Database (KDD)*

Data mining juga dikatakan KDD (*Knowledge Discovery in Database*) membentuk aktifitas dengan cakupan pengumpulan, pemakaian historis yang akan menghasilkan kesesuaian, hubungan / pola pada *set data* yang berukuran besar. Menurut (Elisa & Ervina, 2018) KDD (*knowledge discovery in database*) adalah salah satu tahapan non-trival untuk pengidentifikasian pola pada data yang di mana pola tersebut bersifat sah, baru, juga bisa bermanfaat.

Knowledge Discovery in Database (KDD) merupakan proses untuk menemukan informasi yang berguna dalam *database*. Seluruh proses KDD biasanya terdiri dari langkah-langkah, yaitu memahami bidang aplikasi, membuat

data target yang ditetapkan dari data mentah yang tersimpan dalam *database*, pembersihan data dan *preprocessing* data (Fiandra et al., 2017) .

Menurut (Elisa & Ervina, 2018) tahap-tahap dalam *Knowledge Discovery in Database* (KDD) terdiri dari:

1. *Cleaning Data*

Dalam tahapan ini di lakukan proses pembersihan data dari data tidak konsisten dan pengurangan data (*noise*).

2. *Data Integration*

Tahapan untuk penggabungan data dari berbagai macam sumber yang berbeda-beda.

3. *Data Selection*

Dalam tahapan ini, data di seleksi dari data base dilakukan dengan cara memilih atau memilah data yang sesuai dengan tujuan analisa.

4. *Data Transformation*

Pada tahapan ini data dirubah menjadi sebuah data yang sesuai dalam tahapan *data mining*.

5. *Data mining*

Dalam tahapan ini, di lakukan implementasi metode *data mining* yang sesuai dengan metode khusus guna mendapatkan suatu pola dari data.

6. *Patten Evaluation*

Dalam tahapan ini dilakukan identifikasi pola atau bentuk pada data.

7. *Knowledge Presentation*

Dalam proses ini, hasil *data mining* berbentuk informasi yang di hasilkan dan di presentasikan kemudian bisa diberikan kepada tempat objek nantinya.

2.1.1.2. *Data Mining*

Istilah *data mining* mulai dikenal sejak tahun 1990, ketika pekerjaan pemanfaatan data menjadi sesuatu yang penting dalam berbagai bidang, mulai dari bidang akademik, bisnis hingga medis. Munculnya *datamining* didasarkan pada jumlah data yang tersimpan dalam basis data semakin besar (Nur et al., 2015) . Menurut pendapat (Elisa & Ervina, 2018) *Data mining* adalah sebuah proses dimana satu atau lebih metode pembelajaran sistem komputer yang diberikan tugas untuk mengekstrak dan menganalisa pengetahuan secara langsung atau sekumpulan proses dalam menggali nilai tambah dari sekumpulan data yang tidak diketahui secara manual.

Menurut (Harwati & Sudiya, 2016) *data mining* adalah proses menemukan pola atau informasi menarik dalam data yang dipilih menggunakan teknik atau metode. Dengan teknologi yang semakin canggih, basis data sekarang bias untuk menyimpan data kapasitas besar. Definisi lain menurut (Sumpena et al., 2018) *Datamining* merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan didalam database besar. Salah satu kesulitan untuk mendefinisikan *data mining* adalah kenyataan bahwa *data mining* mewarisi banyak aspek dan

teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu. Berawal dari beberapa disiplin ilmu, *data mining* bertujuan untuk memperbaiki teknik tradisional sehingga bias menangani:

1. Jumlah data yang besar
2. Dimensi data yang tinggi
3. Data yang heterogen dan berbeda sifat

Menurut (Elisa & Ervina, 2018) *datamining* memiliki karakteristik sebagai berikut:

1. *Data mining* mempunyai hubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
2. *Data mining* biasa menggunakan data yang sangat besar, biasanya data yang sangat besar digunakan untuk membuat hasil lebih dipercaya.
3. *Data mining* berguna untuk membuat keputusan yang kritis, terutama dalam strategi.

2.1.1.3. Pengelompokan *Data Mining*

Menurut fungsionalitas nya, tugas *data mining* terbagi menjadi 6 bagian yaitu (Suyanto, 2017):

1. (*Classification*) Klasifikasi yaitu men generalisasi struktur yang diketahui untuk diaplikasikan pada data-data baru.
2. Klasterisasi (*Clustering*) yaitu mengelompokkan data, yang tidak diketahui label kelasnya, kedalam sejumlah kelompok tertentu sesuai dengan aturan kemiripannya.
3. Regresi (*Regression*) yaitu menemukan suatu fungsi yang memodelkan data dengan galat (kesalahan prediksi) seminimal mungkin.
4. Deteksi Anomali (*Anomaly Detection*) yaitu mengidentifikasi data yang tidak umum, bias berupa *outlier* (pencilan), perubahan atau deviasi yang mungkin sangat penting dan perlu investigasi lebih lanjut.
5. Pembelajaran aturan asosiasi (*Association rule learning*) atau pemodelan kebergantungan (*dependency modeling*) Mencari relasi antar variable.
6. Perangkuman (*Summarization*) yaitu menyediakan representasi data yang lebih sederhana, meliputi visualisasi dan pembuatan laporan.

2.1.1.4. Kasifikasi

Bagian yang sangat penting dalam *data mining* adalah teknik klasifikasi, yaitu bagaimana mempelajari sekumpulan data sehingga dihasilkan aturan yang bias mengklasifikasi atau mengenali data-data baru yang belum pernah dipelajari.

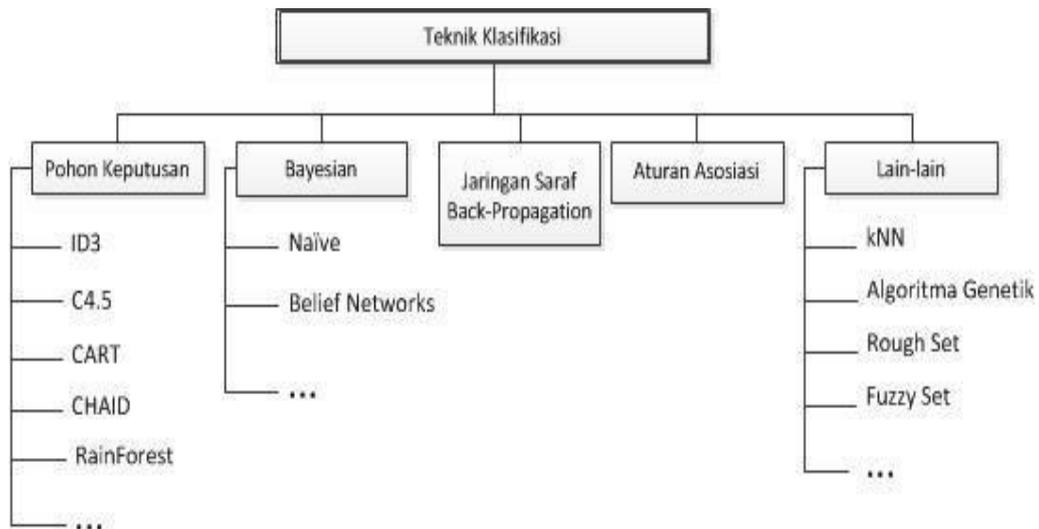
Klasifikasi dapat didefinisikan sebagai proses untuk menyatakan suatu objek data sebagai salah satu kategori (kelas) yang telah didefinisikan (Suyanto, 2017) .

Sedangkan menurut pendapat (Elisa & Ervina, 2018) klasifikasi merupakan sebuah langkah dalam menemukan kekayaan-kekayaan yang menyerupai dalam suatu himpunan objek dalam *database* dan mengelompokkannya kedalam kelas yang beda berdasarkan klasifikasi yang ditentukan. Proses klasifikasi mempunyai dua langkah yakni proses pembelajaran data dianalisa dengan algoritma klasifikasi hingga dibentuk aturan-aturan klasifikasi. Langkah yang ke dua yaitu tahap pengujian dalam memperkirakan keakuratan dari aturan klasifikasi.

Berikut ini adalah beberapa model kasifikasi yang paling sering digunakan (Elisa & Ervina, 2018) antara lain adalah:

1. Pembentukan Deskriptif dibutuhkan sebagai alat penjabaran dari kelas berbeda untuk membedakan suatu objek.
2. Pembentukan Prediktif dibutuhkan untuk memperkirakan judul kelas untuk menrecord sesuatu yang tidak dikenali.

Pada saat melakukan proses klasifikasi ada beberapa teknik yang harus diketahui diantaranya adalah metode Bayesian, jaringan saraf, pola asosiasi, *back propagation*, dan pohon keputusan. Agar lebih jelas nya adapun teknik klasifikasi bisa dilihat pada gambar 2.1 berikut ini.



Gambar 2. 1 Pengelompokkan Teknik Klasifikasi
Sumber: (Jamhur, 2016)

2.1.1.5. Pohon Keputusan (*Decision Tree*)

Menurut pendapat (Suyanto, 2017) pohon keputusan adalah salah satu metode klasifikasi yang populer dan banyak digunakan secara praktis. Metode ini berusaha menemukan model klasifikasi yang tahan terhadap derau. Salah satu metode pohon keputusan yang sangat populer adalah *Iterative Dychotomizer Version 3* (ID3). Pohon keputusan merupakan salah satu bentuk implementasi pola dengan menerapkan berbagai tahapan dan proses yang terdapat dalam kegiatan klasifikasi dan juga kegiatan prediksi.

Arsitektur dari pohon keputusan memang dibuat mirip dengan struktur yang terdapat pada sebuah pohon dan penjelasannya, yaitu sebagai berikut:

1. Simpul bagian akar

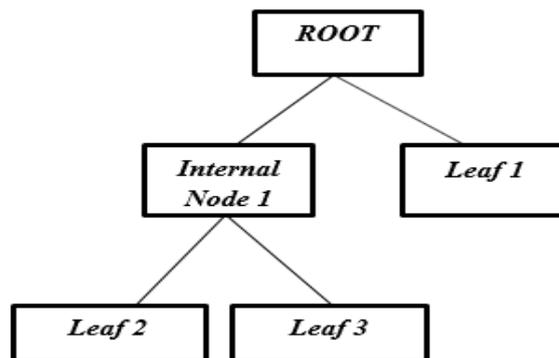
Simpul dari bagian akar terletak pada bagian paling atas dan pertama dari struktur yang terdapat pada pohon keputusan.

2. Simpul internal

Simpul dari bagian ini adalah percabangan lanjut dari simpul bagian akar. Dalam sebuah simpul ini apabila terdapat satu masukan maka dapat mengeluarkan keluaran dalam jumlah maksimal sebanyak dua.

3. Simpul daun

Simpul daun merupakan simpul yang berada pada ujung bagian dari pohon. Dalam simpul daun terdapat sebuah masukan namun tidak memiliki keluaran.



Gambar 2. 2 Konsep dasar pohon keputusan

Konsep pohon keputusan terdapat 3 jenis *node* sebagai berikut (Elisa & Ervina, 2018) :

1. *Rood Node* yaitu *node* paling utama, *node* ini tidak memiliki masukan dan biasanya juga tidak memiliki keluaran atau memiliki banyak keluaran.
2. *Internal Node* yaitu *node* percabangan, *node* ini cuma memiliki satu masukan serta mempunyai banyak keluaran.
3. *Terminal Node* atau *Leaf Node* adalah *node* akhir, *node* cuma memiliki masukan saja tanpa ada keluaran.

Peranan pohon keputusan sebagai alat bantu untuk mengambil keputusan yang telah dikembangkan oleh manusia sejak perkembangan teori pohon. Menurut (Harryanto & Hansun, 2017) pohon keputusan ini memiliki beberapa kekurangan dan kelebihan diantaranya yaitu:

1. Kelebihan pohon keputusan
 - a. Daerah pengambilan keputusan yang kompleks dapat diubah menjadi sederhana.
 - b. Dapat menghilangkan perhitungan yang tidak penting karena proses pengujian hanya berdasarkan kriteria yang diperlukan saja.
 - c. Proses pemilihan fitur dari *internal node* yang berbeda lebih fleksibel. Fitur yang telah dipilih ini akan menjadi pembeda antara kriteria yang satu dengankriteria yang lainnya.

- d. Metode ini dapat menghindari munculnya permasalahan dengan cara menggunakan kriteria dengan jumlah yang sedikit pada *node internal* tanpa mengurangi kualitas keputusan yang dihasilkan.
2. Kekurangan pohon keputusan
- a. Dapat terjadi *overlap* apabila hasil keputusan dan kriteria yang digunakan jumlahnya sangat banyak. Hal ini juga dapat berakibat bertambahnya waktu yang digunakan untuk pengambilan keputusan dan jumlah memori yang dibutuhkan semakin tinggi.
 - b. Akumulasi jumlah *error* dari setiap tingkat pohon keputusan besar.
 - c. Mendesain pohon keputusan yang optimal sulit.
 - d. Kualitas keputusan yang didapatkan sangat tergantung dengan bagaimana pohon tersebut didesain.

2.1.1.6. Algoritma C4.5

Algoritma C4.5 merupakan peningkatan dari ID3, peningkatannya dalam mengalahkan *missing value*, biasanya mengalahkan *continue data* dan *praining*. Algoritma C4.5 adalah sebuah teknik dalam membentuk suatu pohon keputusan berdasarkan data yang sudah diperoleh sebelumnya. Teknik utama yang mendasari algoritma ini yaitu pembentukan pohon keputusan berdasarkan kontes atribut yang mempunyai nilai gain paling tinggi berdasarkan nilai entropy atribut tersebut sebagai langkah dasar dari algoritma ini adalah pembuatan pohon keputusan berdasarkan pemilihan atribut yang mempunyai nilai *gain* yang tertinggi berdasarkan nilai *entropy* atribut tersebut sebagai tahap atribut

klasifikasi. Dalam mencari *gain* yang tertinggi digunakan rumus dibawah ini (Elisa & Ervina, 2018) .

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

Rumus 2. 1 Perhitungan Gain

Dimana:

S : Himpunan kasus

A : Atribut

N : Jumlah partisi atribut A

|Si| : Jumlah kasus pada partisi ke-i

|S| : Jumlah kasus dalam S

Sementara itu, untuk perhitungan nilai *entropy* dapat dilihat pada persamaan berikut.

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

Rumus 2. 2 Perhitungan Entropy

Dimana:

S : Himpunan kasus

A : Fitur

N : Jumlah partisi S

Pi : Proporsi dari Si terhadap S

Hasil dari perhitungan *gain* dan *entropy* ini bisa di uji dengan *WEKA* (*Software Waikato Environment for Knowledge Analysis*). Secara istilah, *entropy* adalah perbedaan atau persamaan. Dalam *data mining*, *entropy* diartikan sebagai sebuah ukuran untuk menghitung persamaan (heterogenitas) dalam suatu kelompok data. Semakin sama sautu kelompok data maka nilai *entropy* yang dihasilkan semkin besar. (Suyanto, 2017) .

2.1.2. Teori Khusus

Berikut merupakan teori yang dipakai dimana, penjelasan dapat diuraikan antara lain sebagai berikut.

2.1.2.1.Penjurusan

Penjurusan ini bertujuan agar nanti siswa dapat menyelesaikan sekolah sesuai pengajaran kejuruan (Rasyid, 2017). dengan kemampuannya sebelum melanjutkan ke jenjang yang lebih tinggi yaitu perguruan tinggi (Muhidin et al., 2019) . Pendidikan kejuruan merupakan layanan social yang efisien jika sesuai dengan kebutuhan seseorang yang memang memerlukan dan memang paling efektif jika dilakukan lewat

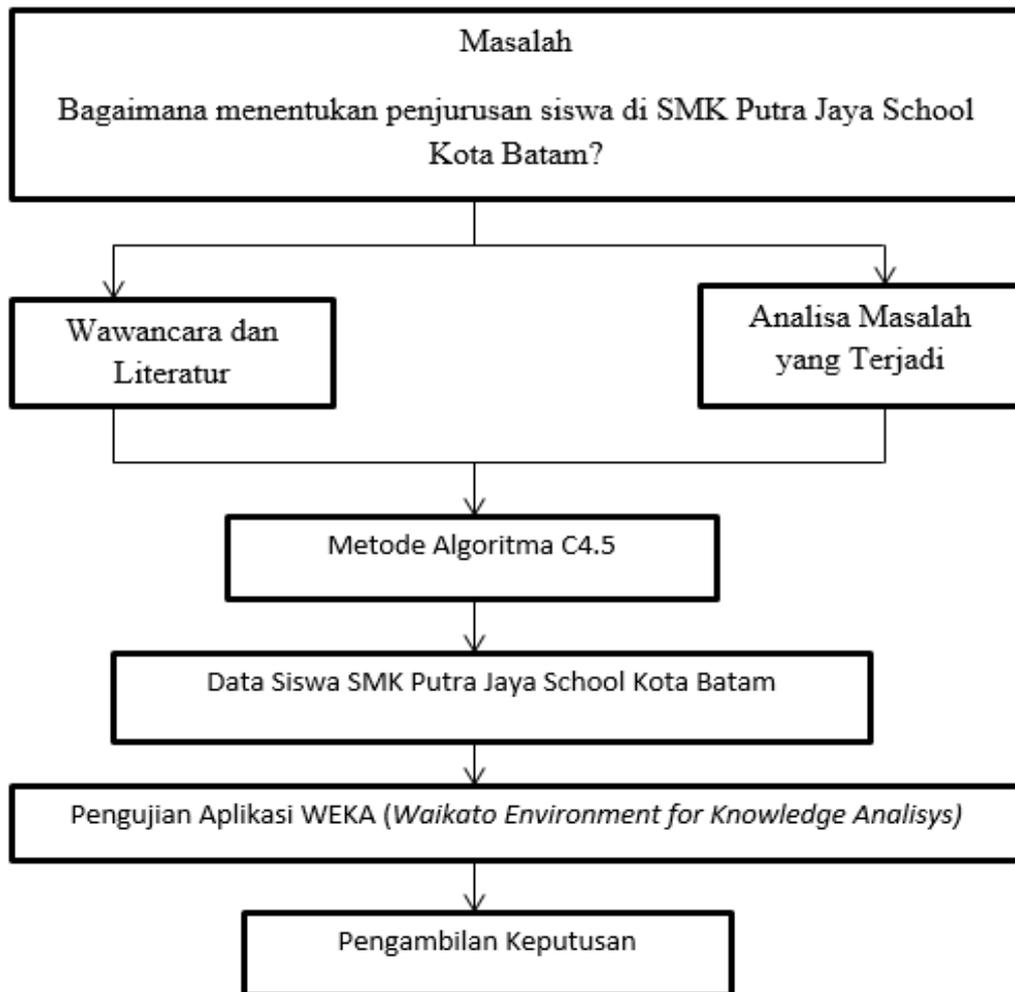
2.1.2.2. WEKA (*Waikato Environment for Knowledge Analysis*)

Menurut pendapat *WEKA (Waikato Environment for Knowledge Analysis)* merupakan bagian alat pembejalaran unstuck mesin, *WEKA* ini dibedarkan di Universitas Waikota yang berada diSelandia Baru yang dipergunkan utuk pendidikan dan penelitian. *WEKA* adalah suatu software yang mempunyai sifat *open source*, maka perangkat lunak ini dapat diakses oleh siapapun. Pengujian

menggunakan WEKA akan jauh lebih praktis juga mudah dalam membentuk suatu pohon keputusan hampir samadengan perhitungan yang dilakukan secara manual. (Elisa & Ervina, 2018) .

2.2. Kerangka Pemikiran

Kerangka pemikiran adalah suatu pola yang menjelaskan mengenai alur logika yang akan digunakan pada penelitian. Kerangka pemikiran pada penelitian ini sebagai berikut:



Gambar 2. 3 Kerangka pemikiran

2.3. Hipotesis Penelitian

Hipotesis merupakan prasangka sementara kepada hasil penelitian yang akan di uji kebenarannya. Hipotesis penelitian ini adalah:

1. Metode *Data Mining* dengan Algoritma C4.5 yang akan digunakan untuk menganalisa data berdasarkan hasil data penjurusan.

2. Hasil berdasarkan analisa data dengan menggunakan Algoritma C4.5 dibutuhkan sebagai dasar atau panduan dalam menentukan hasil dalam penelitian ini yaitu menentukan penjurusan calon siswa baru.

2.4. Penelitian Terdahulu

Peneliti yang lainya menemukan beberapa hasil penelitian terdahulu yang bersangkutan dengan menganalisa menggunakan algoritma C4.5 yang dirangkum sebagai berikut:

1. Penelitian yang berjudul Klasifikasi Penerimaan Mahasiswa Baru Menggunakan Algoritma C4.5 Dan Adaboost yang dilakukan oleh (Agustin et al., 2017) mengatakan bahwa pola tersebut dapat membantu untuk mengambil keputusan penerimaan mahasiswa baru yang dapat lulus tepat waktu dan mahasiswa yang lulus terlambat dapat terprediksi lebih awal.
2. Penelitian yang berjudul Penerimaan Calon Siswa Baru Dan Penentuan Penjurusan Dengan Algoritma C4.5 SMK Plus PGRI 1 Cibinong yang dilakukan oleh (Sumpena et al., 2018) mengatakan bahwa penjurusan calon siswa baru dapat menggunakan aplikasi *datamining* menggunakan metode *decision tree* serta algoritma C4.5. Serta tingkat keakuratan dari perhitungan yang sudah dilakukan sebesar 90%.
3. Dalam sebuah penelitian dengan judul Penerapan Algoritma K-Means Pada Siswa Baru Sekolah Menengah Kejuruan Untuk Clustering Jurusan yang dilakukan oleh (Nur et al., 2015) mengatakan bahwa terdapat pusat *cluster* dengan *Cluster* 1=1.4;2.2;2.2, *Cluster*

$2=2.28;1.64;4$ dan *cluster 3=5;3;6*. Pusat cluster tersebut diperoleh beberapa aturan sehingga menghasilkan pusat cluster yang terbaik.

4. Penelitian yang berjudul Analisis Keputusan Menentukan Jurusan Pada Sekolah Menengah Kejuruan Dengan Metode *Simple Additive Weighting* yang dilakukan oleh (Rusdiansyah, 2017) mengatakan bahwa teknik *simple Additive Weighting (SAW)* yang dipergunakan dalam menyeleksi berdasarkan nilai penentuan jurusan guna memperoleh siswa yang memiliki kualitas.
5. Penelitian yang berjudul Penerapan Algoritma C4.5 untuk Klasifikasi Pola Penjurusan di Sekolah Menengah Kejuruan (SMK) Kota Tasikmalaya yang dilakukan oleh (Sambani & Nuraeni, 2017) mengatakan bahwa tingkat akurasi algoritma C4.5 pada penelitian ini sudah mencapai tingkat baik, sehingga dapat meningkatkan ketelitian dalam proses klasifikasi penentuan jurusan berdasarkan minat siswa.
6. Penelitian yang berjudul Implementasi Logika Fuzzy Dalam Menentukan Jurusan Bagi Siswa Baru Sekolah Menengah Kejuruan (SMK) Negeri 1 Air Putih yang dilakukan oleh (Irawan & Herviana, 2018) mengatakan bahwa berdasarkan hasil penelitian didapatkan hasil bahwa metode Tsukamoto dapat memberikan rekomendasi jurusan yang sesuai mata pelajaran yang memiliki nilai yang lebih tinggi.

7. Penelitian yang berjudul Indikator Pemilihan Jurusan Pada SMK Nusantara Menggunakan Metode SAW yang dilakukan oleh (Hasanah et al., 2018) mengatakan bahwa hasil yang didapatkan pada penelitian ini mempermudah siswa baru untuk menentukan dan mengakses jurusan yang akan dipilih oleh sesuai minat dan bakatnya. Penelitian dengan menggunakan metode ini mendapatkan hasil 100% dari 3 jurusan.

BAB III

METODE PENELITIAN

3.1. Desain Penelitian

Desain penelitian merupakan langkah dalam memecahkan masalah pada penelitian ini. Pada penelitian khususnya dalam laporan ini adalah mengenai penjurusan siswa Sekolah Menengah Kejuruan (SMK) Putra Jaya School Kota Batam. Model atau gambar desain dari penelitian yaitu:



Gambar 3. 1 Desain Penelitian

Berdasarkan pada gambar 3.1 diatas tahap-tahap desain penelitian bisa dijabarkan dibawah ini :

1. Mendeskripsikan Masalah

Penulis memberikan penjelasan permasalahan yang akan diteliti serta menginterpretasikan batasan dari masalah yang akan dicermati, hingga mendukung dalam menghasilkan sebuah penyelesaian dari masalah itu. Sehingga hasil dari penelitian akan lebih membantu karena penelitian tersebut dapat lebih terarah.

2. Analisa Masalah

Tahapan ini merupakan tahapan agar bias menginterpretasikan masalah dengan pedoman dari batasan masalah. Dengan melakukan penganalisaan dari masalah yang sudah ditetapkan. Penulis berupaya dalam melakukan analisa berdasarkan beberapa permasalahan yang ingin diangkat dan diteliti oleh penulis.

3. Mempelajari Literatur

Penulis berupaya mempelajari beberapa sumber teori yang mendukung penelitian dari jurnal untuk dijadikan pedoman bagi penulis dalam melakukan penelitian.

4. Mengumpulkan Data

Proses dalam mengumpulkan data dilakukan dengan teknik dari wawancara. Penulis melakukan kegiatan dalam wawancara dengan

memberikan pertanyaan yang berkaitan dengan data penjurusan siswa Sekolah Menengah Kejuruan (SMK) Putra Jaya School Kota Batam.

5. Analisa Teknik Pengolahan Data Dengan Algoritma C4.5

Penyusunan menggunakan algoritma C4.5 dari model sistem sampai dengan dibentuknya pohon keputusan, kemudian menghasilkan rule penjurusan siswa SMK.

6. Implementasi Algoritma C4.5

Penulis menentukan kelengkapan data yang dijadikan sebagai akar dalam *decision tree* dan kemudian dilakukan perhitungan nilai *gain* tertinggi dari keseluruhan kelengkapan yang terdapat pada data yang selanjutnya digunakan juga dalam proses membentuk struktur *decision tree*.

7. Pengujian Hasil

Pada tahapan terakhir ini, peneliti melakukan pengujian setelah selesai melakukan perhitungan terhadap nilai *gain* yang tertinggi dan membentuk *decision tree* dengan menggunakan bantuan *software WEKA* versi 3.9.2.

3.2. Objek Penelitian

Pada sub bab objek penelitian ini, penulis menentukan jadwal penelitian dan lokasi penelitian. Penelitian dilaksanakan disalah satu sekolah menengah kejuruan yaitu SMK Putra Jaya School.

3.2.1. Lokasi Penelitian

Penelitian ini berlokasi di Sekolah Menengah Kejuruan (SMK) Putra Jaya School beralamat di Jl. Laksamana Bintang, Komplek Gading Mas Blok A No. 01 Sei Panas Kota Batam.



Gambar 3. 2 Lokasi Penelitian

3.2.2. Jadwal Penelitian

Demikian penjelasan penulis mengenai rencana waktu yang akan dilakukan dalam melaksanakan kegiatan analisa mengenai penjurusan siswa Sekolah Menengah Kejuruan (SMK), adalah sebagai berikut.

Tabel 3. 1 Lokasi Penelitian

| nNo | Kegiatan | Maret-20 | April-20 | Mei-20 | Juni-20 |
|-----|----------------------|----------|----------|--------|---------|
| 1 | Pengajuan Judul | | | | |
| 2 | Bimbingan | | | | |
| 3 | Studi Kepustakaan | | | | |
| 4 | Pengajuan Penelitian | | | | |
| 5 | Rancangan Penelitian | | | | |
| 6 | Pengumpulan Data | | | | |
| 7 | Pengolahan data | | | | |
| 8 | Penyusunan laporan | | | | |

3.3. Populasi dan Sampel

Pada sub bab ini penulis akan menjelaskan mengenai keterangan yang dimiliki oleh populasi dan sampel yang dijadikan bahan dalam pelaksanaan kegiatan penelitian sebagai berikut.

3.3.1. Populasi

Populasi adalah keseluruhan objek yang akan diteliti juga wilayah generalisasi yang terdiri dari subyek yang menjadi kuantitas dan karakteristik tertentu yang ditetapkan kemudian dibuat kesimpulannya. Populasi yang digunakan dalam penelitian ini yaitu data penjurusan siswa/i Sekolah Menengah Kejuruan (SMK) Putra Jaya School Kota Batam pada tahun ajaran 2018/2019.

3.3.2. Sampel

Sampel adalah bagian dari jumlah dan karakteristik yang dimiliki oleh populasi tersebut.

3.4. Variabel Penelitian

Pengolahan data dilakukan sesuai data penjurusan siswa Sekolah Menengah Kejuruan (SMK) Putra Jaya School Kota Batam. Untuk menganalisa data pada penelitian ini diperlukan variabel yaitu sebagai berikut:

1. Tes buta warna
2. Tes kesehatan
3. Tes Wawancara

Untuk hasil akhir keputusannya adalah **Diterima** dan **Tidak Diterima**.

3.5. Teknik Pengumpulan Data

Pengumpulan data dilakukan untuk mendapatkan informasi yang dibutuhkan untuk mencapai tujuan penelitian yaitu sebagai berikut:

1. Wawancara

Wawancara merupakan teknik pengumpulan data yang dilakukan

melalui tatap muka dan Tanya jawab langsung antara peneliti dan narasumber. Pada metode ini peneliti langsung menanyakan beberapa hal tentang kebutuhan informasi berkaitan dengan masalah yang dihadapi panitia penerimaan siswa Sekolah Menengah Kejuruan (SMK).

2. Studi Literatur

Mempelajari bahan dari jurnal penelitian yang sudah dilakukan peneliti lainnya namun tetap berkaitan dengan penelitian yang dilakukan penulis.

3. Observasi

Mengamati kegiatan penerimaan siswa baru Sekolah Menengah Kejuruan (SMK) Putra Jaya School Kota Batam tahun ajaran 2018/2019.

4. Dokumentasi

Penulis mengumpulkan bukti dari data penerimaan siswa di Sekolah Menengah Kejuruan (SMK) Putra Jaya School.

3.6. Model Penelitian

Penulis melakukan jenis model penelitian menggunakan *data mining* dan jenis algoritma C4.5. Tahapan awal dilakukan klasifikasi sesuai dengan kategori atau kriteria data, selanjutnya dilakukan perhitungan algoritma dari C4.5 guna memperoleh nilai *entropy* dan *gain* tertinggi pada saat membuat struktur dari *decision tree* dilakukan secara berulang hingga simpul pada *decision tree* tidak

memiliki turunan lagi sampai didapatkan pola berupa aturan dasar penilaian mengenai penjurusan siswa Sekolah Menengah Kejuruan (SMK) Putra Jaya School di Kota Batam.