

BAB III

METODOLOGI

Dalam penelitian ini metodologi memegang peranan penting guna mendapatkan data yang obyektik, valid dan selanjutnya digunakan untuk memecahkan permasalahan yang telah dirumuskan. Maka dengan demikian memecahkan metodologi sangat diperlukan dalam rangka mengumpulkan data untuk memecahkan suatu masalah sehingga menyusun laporan ini dapat dipertanggung jawabkan.

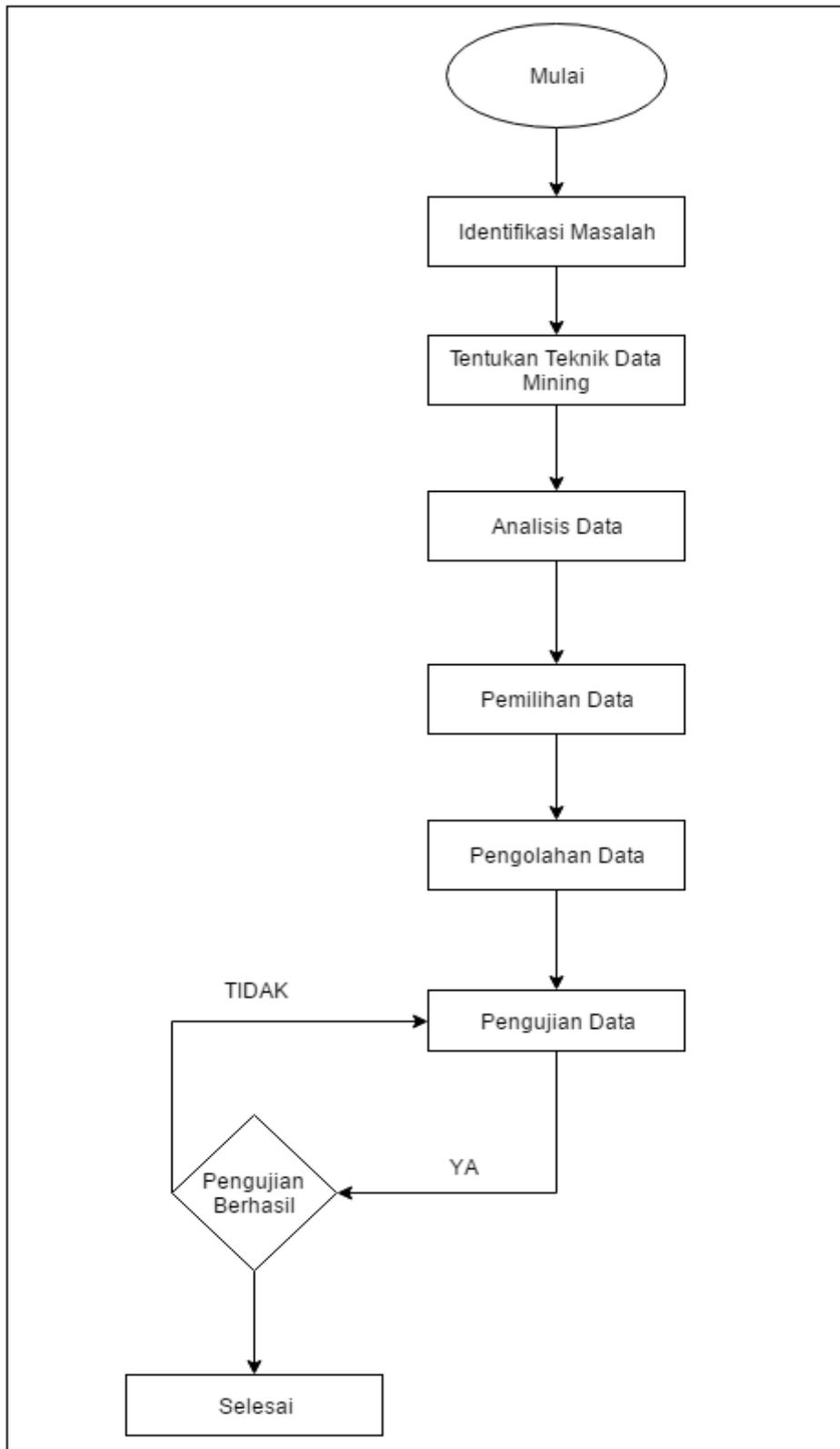
3.1. Prosedur Penelitian

Prosedur penelitian adalah langkah – langkah yang digunakan sebagai alat untuk mengumpulkan data dan menjawab pertanyaan – pertanyaan dalam penelitian. Didalam prosedur penelitian ini, penulis membahas tentang metode dan teknik pengumpulan data.

Penelitian eksperimen menggunakan data dalam penelitiannya dan menghasilkan kesimpulan yang mampu dibuktikan dengan pengamatan atau percobaan. Penelitian ini termasuk penelitian eksperimen dengan menggunakan data dosen Fakultas Teknik Universitas Muhammadiyah Yogyakarta. Prosedur penelitian akan dijelaskan pada gambar 3.1.

3.1.1. Identifikasi Masalah

Pada dasarnya semua penelitian akan selalu didahului dengan identifikasi masalah, hal ini berguna untuk peneliti agar fokus pada titik permasalahan. Dalam kasus ini peneliti ingin menganalisa dengan melakukan observasi klasifikasi lama kerja Fakultas Teknik di Universitas Muhammadiyah Yogyakarta dengan menggunakan algoritma C.45 sekaligus ingin mengetahui bagaimana pola dan cara kerja klasifikasi dalam *data mining* tersebut.



Gambar 3.1 Kerangka Pemikiran

3.1.2. Menentukan Metode Data mining

Setelah identifikasi tahap selanjutnya adalah menentukan metode apa yang cocok digunakan dalam teknik klasifikasi, setelah melakukan pengamatan dan observasi peneliti menentukan algoritma C.45.

3.1.3. Analisis Data

Pada tahap ini peneliti melakukan analisis data seperti darimana data didapatkan, menentukan atribut apa saja yang diperlukan dan juga variabel apa yang cocok. Peneliti mendapatkan data ini dari *database server* BSI (Biro Sistem Informasi) Universitas Muhammadiyah Yogyakarta. Untuk bisa mendapatkan data tersebut perlu adanya hak akses, artinya *data server* tidak bisa langsung diakses tanpa adanya izin dari *admin* karena data tersebut bersifat rahasia.

Perlu diketahui peneliti sudah tiga kali mengganti topik penelitian hal ini dikarenakan ketersediaan data. awalnya, peneliti ingin menganalisa “Penerimaan Pegawai Baru Menggunakan Metode *Decision Tree*” namun setelah menganalisa data ternyata data kurang lengkap dan terpaksa harus dihentikan. Topik yang kedua yaitu “Implementasi C.45 Terhadap Kinerja Dosen Menggunakan *Decision Tree*” namun setelah data dianalisa, penelitian juga terpaksa harus dihentikan karena data tidak mendukung. Lalu yang ketiga “Klasifikasi Lama Kerja Dosen Fakultas Teknik Universitas Muhammadiyah Yogyakarta Dengan Menerapkan Metode *Decision Tree* Dan Algoritma C.45. Pada analisa data yang ketiga ini data yang diperlukan bisa dipakai meskipun ada beberapa data tidak *valid*.

3.1.4. Pemilihan data

Setelah data dianalisa dan ketersediaanya juga ada di *database server* BSI, selanjutnya data mentah akan diolah menjadi data yang berbentuk informasi. Data dipilih berdasarkan keperluan penelitian dengan melakukan seleksi data sehingga menjadi *dataset*.

3.1.5. Pengolahan data

Setelah data sudah dipilih tahap penelitian selanjutnya yaitu pengolahan data, bisa dikatakan disini sangat menentukan keberhasilan dari penelitian tersebut karena mulai dari *dataset*, atribut, variabelnya harus sangat diperhatikan.

3.1.6. Pengujian Data

Pengujian juga merupakan tahapan yang sangat menentukan apakah pengujian yang dibuat telah layak digunakan atau belum. Jika pengujian belum menghasilkan luaran yang sesuai dengan kebutuhan, maka tahapan penelitian kembali ke pengolahan data, jika sudah memasuki tahapan ulangan atau kegagalan perlu adanya pemecahan masalah terhadap kegagalan tersebut dengan analisa pengolahan data.

3.2. Bahan Penelitian

Dalam melakukan penelitian ini, berikut bahan penelitian yang digunakan:

3.2.1. Literatur

Yaitu jurnal, buku dan artikel yang berhubungan dengan Algoritma C.45, *decision tree* dan *data mining* yang relevan dengan penelitian dan semuanya digunakan sebagai bahan referensi untuk memperdalam pemahaman teori – teori dan penerapan data.

3.2.2. Data Institusi Universitas

Data berasal dari *data server* BSI (Biro Sistem Informasi) Universitas Muhammadiyah Yogyakarta.

3.3. Alat Penelitian

Adapun alat penelitian yang digunakan dalam penelitian ini adalah:

3.3.1. SQL Server 2014

SQL Server merupakan *Relational Database Management System (RDBMS)* yang digunakan untuk menghubungkan data yang ada pada *server*. *SQL Server* digunakan untuk mengelola basis data, seperti menyediakan basis data secara fisik dan membuat table – tabel.

3.3.2. WEKA

WEKA (*Waikato Environment for Knowledge Analysis*) adalah suatu perangkat lunak pembelajaran mesin yang populer ditulis dengan *Java*, yang dikembangkan di Universitas Waikato di selandia baru. WEKA adalah perangkat lunak gratis yang tersedia di bawah GNU (*General Public License*). Pada WEKA menyediakan penggunaan teknik klasifikasi menggunakan pohon keputusan dengan algoritma J48. Teknik klasifikasi dan algoritma yang digunakan di WEKA disebut *classifier*.

WEKA terdiri dari koleksi algoritma *machine learning* yang dapat digunakan untuk melakukan generalisasi atau formulasi dari sekumpulan data *sampling*. Algoritma ini bisa diterapkan secara langsung kedalam *dataset* atau bisa juga dipanggil dari kode *java*. WEKA memiliki *tools* untuk data *re-processing*, *classification*, *regression*, *clustering*, *association rules*, dan *visualization*. WEKA mengorganisasi kelas-kelas kedalam paket-paket dan setiap kelas dipaket dapat mereferensi kelas lain dipaket lain. Kelas yang paling penting adalah *Classifier*, yang mendeklarasikan struktur umum dari skema klasifikasi dan prediksi. Kelas ini memiliki 2 metode, yaitu *build Classifier* dan *classify Instance*, yang harus diimplementasikan oleh kelas-kelas yang menginduk kekelas ini. Semua kelas yang mengimplementasikan algoritma klasifikasi menginduk kekelas *Classifier*, termasuk kelas J48. J48 yang menangani himpunan data dalam format ARFF, tidak mengandung kode untuk mengkonstruksi pohon keputusan. Kelas ini mereferensi kelas-kelas lain, kebanyakan dipaket WEKA. *Classifiers* J48, yang mengerjakan semua proses konstruksi pohon. Pada WEKA ada beberapa metode pemilihan variable dari suatu *dataset*, diantaranya *BestFirst*, *ExhaustiveSearch*, *FCBFSearch*, *GeneticSearch*, *GreedyStepwise*, *RaceSearch*, *RandomSearch*, *Ranker* dan *RankerSearch*.

3.3.3. RapidMiner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (*open source*). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap *data mining*, *text mining* dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator *data mining*, termasuk operator untuk *input*, *output*, data *preprocessing* dan *visualisasi*.

RapidMiner memiliki beberapa sifat sebagai berikut:

1. Ditulis dengan bahasa pemrograman *Java* sehingga dapat dijalankan di berbagai sistem operasi.
2. Proses penemuan pengetahuan dimodelkan sebagai *operator trees*
3. Representasi XML internal untuk memastikan format standar pertukaran data.

4. Bahasa *scripting* memungkinkan untuk eksperimen skala besar dan otomatisasi eksperimen.
5. Konsep *multi-layer* untuk menjamin tampilan data yang efisien dan menjamin penanganan data.
6. Memiliki GUI, *command line mode*, dan *Java API* yang dapat dipanggil dari program lain.

Tools yang digunakan dalam RapidMiner ini adalah:

1. *Panel operator view*

Panel operator view merupakan induk langkah kerja dari proses analisa

2. *Panel repository view*

merupakan komponen utama dalam *Design Perspective* selain *Operator View*. *View* ini dapat digunakan untuk mengelola dan menata proses Analisis menjadi proyek dan pada saat yang sama juga dapat digunakan sebagai sumber data dan yang berkaitan dengan meta data.

3. *Process view*

Process View menunjukkan langkah – langkah tertentu dalam proses analisis dan sebagai penghubung langkah – langkah tersebut.

4. *Parameter view*

Digunakan untuk mengatur fungsionalitas dari beberapa operator yang membutuhkannya agar dapat dijalankan sesuai dengan proses analisa.

3.3.4. Microsoft Excel 2013

Microsoft excel adalah sebuah program aplikasi lembar kerja *spreadsheet* yang dibuat dan didistribusikan oleh Microsoft Corporation yang dapat dijalankan pada Microsoft Windows dan Mac OS.

Microsoft excel merupakan suatu *software* yang digunakan untuk membuat tabel dalam *software* ini juga kita dapat menggunakan rumus – rumus guna menjumlahkan, mengurangi, mencari nilai rata - rata, mencari nilai *max* ataupun

min, serta menampilkan persyaratan suatu kondisi secara langsung tanpa harus menuliskan nya satu persatu.