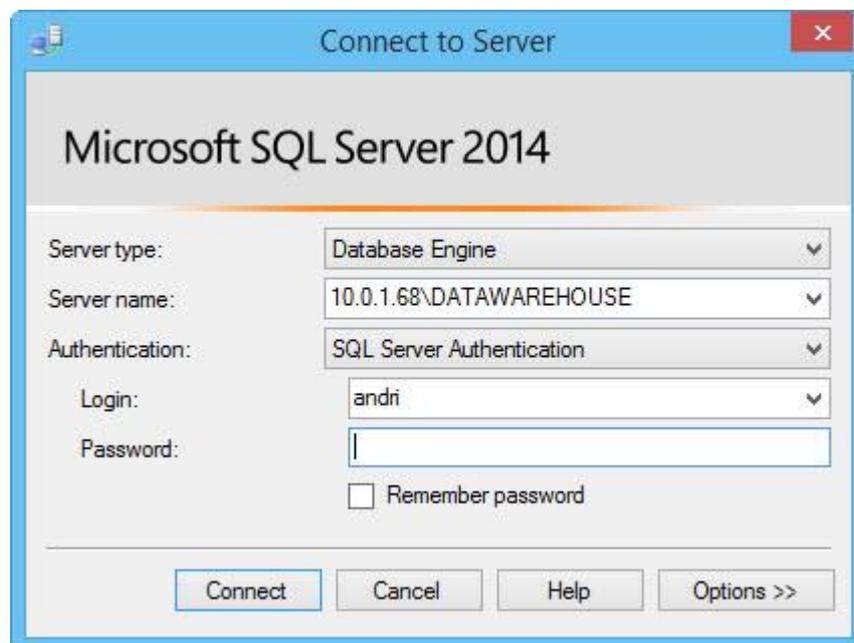


BAB IV

HASIL DAN PEMBAHASAN

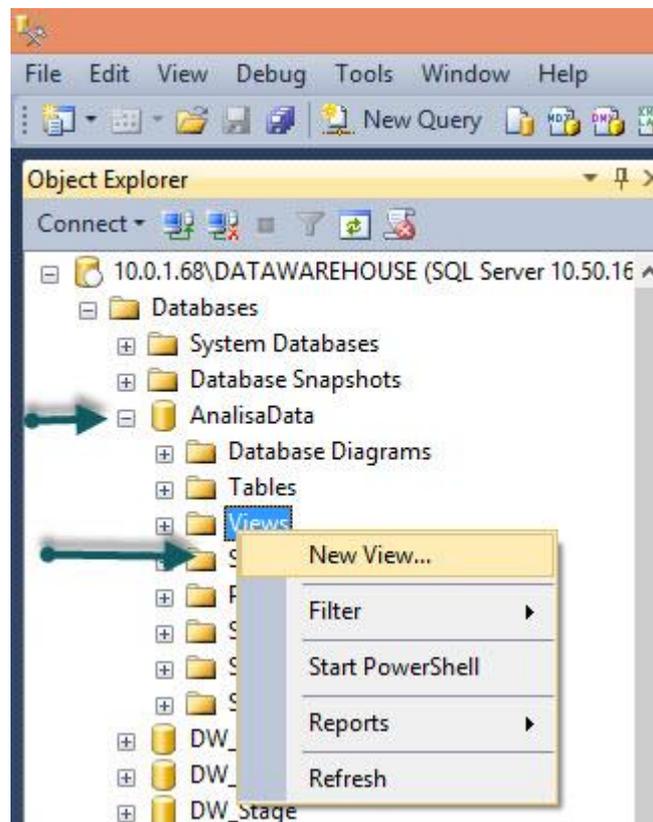
1.1. Pengumpulan data

Data yang digunakan dalam penelitian ini adalah *data warehouse Graduation* Universitas Muhammadiyah Yogyakarta pada Fakultas Teknik UMY tahun kelulusan 2013, 2014 dan 2015. *Software* yang digunakan untuk mengakses *data warehouse* ini yaitu *SQL Server Management 2014* dan untuk dapat mengakses *data warehouse* peneliti harus mempunyai hak akses ke *database Server* Universitas Muhammadiyah Yogyakarta. *Server name* dari *database server* BSI (Biro Sarana Informasi) adalah 10.0.1.68\DATAWAREHOUSE seperti pada gambar 4.1 dibawah ini



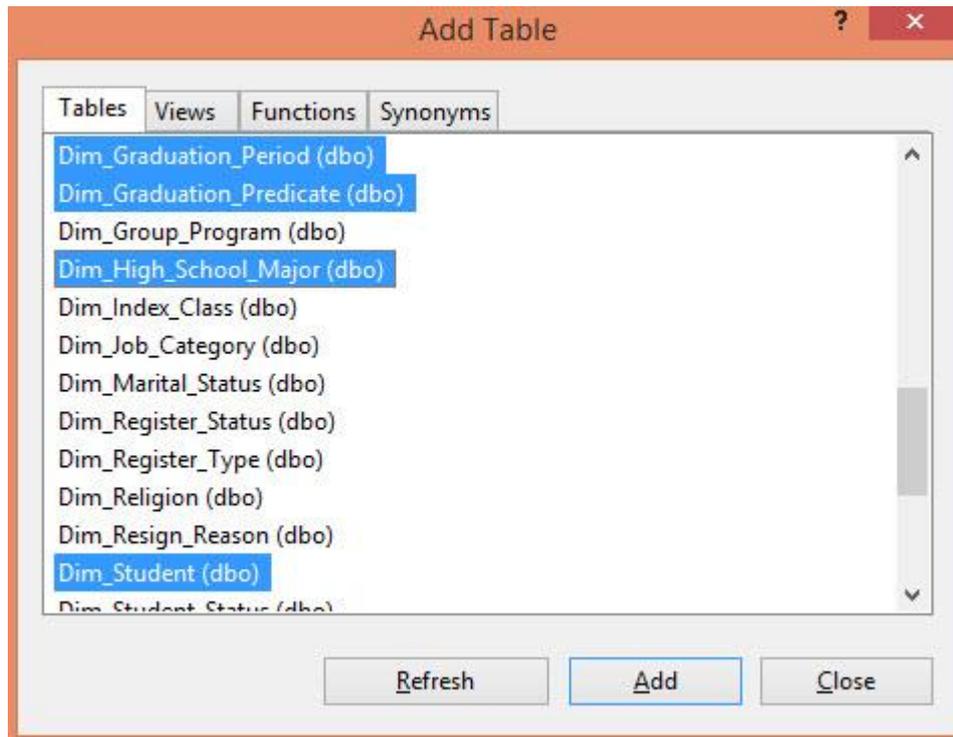
Gambar 4.1 Menghubungkan ke *server*.

Setelah masuk ke *database server* penulis membuat *database* baru di dalam folder analisa data agar lebih muda digunakan jika suatu waktu ingin menggunakan lagi data tersebut. Untuk membuat *database* baru penulis harus melakukan *expand* pada folder Analisis Data dan selanjutnya klik *new view* pada folder *view* seperti gambar 4.2 dibawah ini.



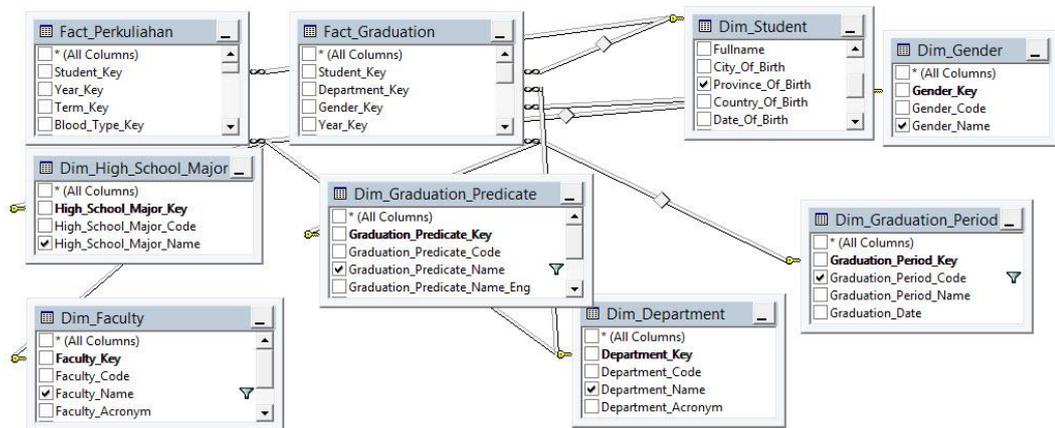
Gambar 4.2 *create database.*

Setelah klik *new view* maka langkah selanjutnya yaitu *add table* atau memilih tabel mana yang ingin digunakan seperti gambar 4.3 dibawah ini.



Gambar 4.3 Add table.

Data yang diambil dari *data warehouse* ini menggunakan dua tabel *factual* yaitu *fact_graduation* dan *fact_perkuliahan* sedangkan untuk tabel *dimensional* memiliki 7 tabel yaitu: *dim_student*, *dim_gender*, *dim_graduation_periode*, *dim_department*, *dim_graduation_predicate*, *dim_faculty*, *dim_high_major*. Pengambilan data berdasarkan atribut yang akan digunakan seperti gambar 4.4 dan penyeleksian data dari data asli (*data warehouse*) dapat dilihat di gambar 4.5 sebagai berikut.



Gambar 4.4 Tampilan view dari *data warehouse*.

Setelah memilih *atribut* yang ingin digunakan langkah selanjutnya yaitu memfilter data seperti gambar 4.5 dibawah ini.

Column	Alias	Table	Outp...	Sort Type	Sort Order	Filter	Or...	0 ^
Department_Name		Dim_Depa...	<input checked="" type="checkbox"/>					
Gender_Name		Dim_Gend...	<input checked="" type="checkbox"/>					
Graduation_Period_Code		Dim_Grad...	<input checked="" type="checkbox"/>	→		IN ('20131', '20132', '20133', '20141', '20142', '20143'...		
High_School_Major_Name		Dim_High...	<input checked="" type="checkbox"/>					
Province_Of_Birth		Dim_Stud...	<input checked="" type="checkbox"/>					
Faculty_Name		Dim_Facul...	<input checked="" type="checkbox"/>	→		= 'TEKNIK'		
Graduation_Predicate_Name		Dim_Grad...	<input checked="" type="checkbox"/>	→		IN ('Cumlaude', 'Sangat Memuaskan', 'Memuaskan')		

Gambar 4.5 Proses *filter* data dari *data warehouse*.

1.2. Seleksi Data (*data selection*)

Data selection adalah proses menganalisis data-data yang relevan dari *database* karena sering ditemukan bahwa tidak semua data dibutuhkan dalam proses *data mining*. Data tersebut dipilih dan diseleksi dari *database* untuk di analisis. Sumber data yang digunakan dalam penelitian ini berasal dari data mahasiswa yang telah lulus tahun 2013 sampai dengan 2015 pada Fakultas Teknik Universitas Muhammadiyah Yogyakarta. Dari semua data yang digunakan hanya *Province Of Birth*, *High School Major Name*, *Gender Name*, *Graduation*

Predicate Name. Karena informasi yang terkandung didalamnya sudah mewakili informasi yang dibutuhkan untuk dijadikan *indicator* penelitian.

1.3. Pembersihan Data (*cleaning data*)

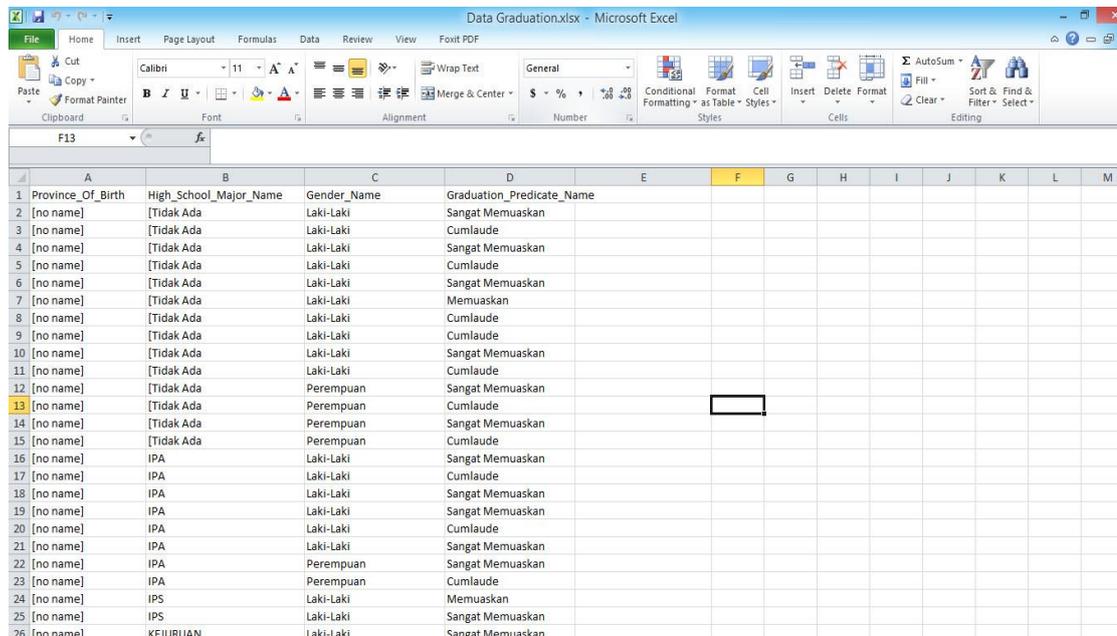
Setelah tahap pengumpulan data dan *filter* data maka tahap selanjutnya yaitu *cleaning data* agar tidak ada duplikasi data, memeriksa data yang inkonsisten dan memperbaiki kesalahan pada data seperti kesalahan cetak, sehingga data tersebut dapat diolah dan dilakukan proses *data mining*. Setelah semua data yang di butuhkan telah melalui tahap *cleaning data* maka data akan disimpan dalam *dataset* baru yang menggunakan *Microsoft Office Excel* dengan format csv. Data yang diambil dari fakultas Teknik ini ada data yang *unknown* yaitu data pada atribut *Province_Of_Birth* sebanyak 28 data dan atribut *High_School_Major_Name* sebanyak 28 data. Proses pembersihan data dapat di lihat di gambar 4.6.

```
SELECT DISTINCT
    dbo.Dim_Department.Department_Name, dbo.Dim_Gender.Gender_Name, dbo.Dim_Graduation_Period.Graduation_Period_Code, dbo.Dim_High_School_Major.High_School_Maj
    dbo.Dim_Student.Province_Of_Birth, dbo.Dim_Faculty.Faculty_Name, dbo.Dim_Graduation_Predicate.Graduation_Predicate_Name
FROM
    dbo.Dim_Department INNER JOIN
    dbo.Fact_Graduation ON dbo.Dim_Department.Department_Key = dbo.Fact_Graduation.Department_Key INNER JOIN
    dbo.Dim_Gender ON dbo.Fact_Graduation.Gender_Key = dbo.Dim_Gender.Gender_Key INNER JOIN
    dbo.Dim_Graduation_Period ON dbo.Fact_Graduation.Graduation_Period_Key = dbo.Dim_Graduation_Period.Graduation_Period_Key INNER JOIN
    dbo.Dim_Graduation_Predicate ON dbo.Fact_Graduation.Graduation_Predicate_Key = dbo.Dim_Graduation_Predicate.Graduation_Predicate_Key INNER JOIN
    dbo.Dim_Student ON dbo.Fact_Graduation.Student_Key = dbo.Dim_Student.Student_Key INNER JOIN
    dbo.Fact_Perkuliahan ON dbo.Dim_Department.Department_Key = dbo.Fact_Perkuliahan.Department_Key AND dbo.Dim_Gender.Gender_Key = dbo.Fact_Perkuliahan.Gender_Key
    dbo.Dim_Student.Student_Key = dbo.Fact_Perkuliahan.Student_Key INNER JOIN
    dbo.Dim_Faculty ON dbo.Fact_Perkuliahan.Faculty_Key = dbo.Dim_Faculty.Faculty_Key INNER JOIN
    dbo.Dim_High_School_Major ON dbo.Fact_Perkuliahan.High_School_Major_Key = dbo.Dim_High_School_Major.High_School_Major_Key
```

Gambar 4.6 *Cleaning data.*

1.4. Transformasi Data (*data transformation*).

Data Transformation adalah tahap mengubah data menjadi bentuk yang sesuai untuk diproses dalam *data mining*. Beberapa metode *data mining* membutuhkan format data yang khusus sebelum bisa di aplikasikan. Dalam penelitian ini data yang akan diproses dari *database SQL Server 2014 Management Studio* akan diubah menjadi *file CSV (comma delimited)* yang dapat digunakan untuk pengolahan data pada *Software RapidMiner* dan nama atribut data juga di ubah dari *Province_Of Birth, Senior_High_School, Gender_Name* dan *Graduation_Predicate_Name*. di ubah menjadi *Provinsi, Jenis Kelamin, Jurusan SMA, Predikat Kelulusan* dan nilai yang ada pada atribut *Provinsi* juga di ubah sesuai dengan kebutuhan. Gambar 4.7 adalah data yang belum di ubah oleh penulis atau yang belum di lakukan transformasi data.



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Province_Of_Birth	High_School_Major_Name	Gender_Name	Graduation_Predicate_Name									
2	[no name]	[Tidak Ada]	Laki-Laki	Sangat Memuaskan									
3	[no name]	[Tidak Ada]	Laki-Laki	Cumlaude									
4	[no name]	[Tidak Ada]	Laki-Laki	Sangat Memuaskan									
5	[no name]	[Tidak Ada]	Laki-Laki	Cumlaude									
6	[no name]	[Tidak Ada]	Laki-Laki	Sangat Memuaskan									
7	[no name]	[Tidak Ada]	Laki-Laki	Memuaskan									
8	[no name]	[Tidak Ada]	Laki-Laki	Cumlaude									
9	[no name]	[Tidak Ada]	Laki-Laki	Cumlaude									
10	[no name]	[Tidak Ada]	Laki-Laki	Sangat Memuaskan									
11	[no name]	[Tidak Ada]	Laki-Laki	Cumlaude									
12	[no name]	[Tidak Ada]	Perempuan	Sangat Memuaskan									
13	[no name]	[Tidak Ada]	Perempuan	Cumlaude									
14	[no name]	[Tidak Ada]	Perempuan	Sangat Memuaskan									
15	[no name]	[Tidak Ada]	Perempuan	Cumlaude									
16	[no name]	IPA	Laki-Laki	Sangat Memuaskan									
17	[no name]	IPA	Laki-Laki	Cumlaude									
18	[no name]	IPA	Laki-Laki	Sangat Memuaskan									
19	[no name]	IPA	Laki-Laki	Sangat Memuaskan									
20	[no name]	IPA	Laki-Laki	Cumlaude									
21	[no name]	IPA	Laki-Laki	Sangat Memuaskan									
22	[no name]	IPA	Perempuan	Sangat Memuaskan									
23	[no name]	IPA	Perempuan	Cumlaude									
24	[no name]	IPS	Laki-Laki	Memuaskan									
25	[no name]	IPS	Laki-Laki	Sangat Memuaskan									
26	[no name]	KEJURUAN	Laki-Laki	Sangat Memuaskan									

Gambar 4.7 Data format excel.

Selanjutnya penulis melakukan transformasi data agar penelitian ini bisa berjalan dengan baik dan gambar 4.8 di bawah ini adalah data yang telah terjadi transformasi data.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Provinsi	Jurusan SMA	Jenis Kelamin	Predikat Kelulusan										
2	A	Tidak Ada	Laki-Laki	Sangat Memuaskan										
3	A	Tidak Ada	Laki-Laki	Cumlaude										
4	A	Tidak Ada	Laki-Laki	Sangat Memuaskan										
5	A	Tidak Ada	Laki-Laki	Cumlaude										
6	A	Tidak Ada	Laki-Laki	Sangat Memuaskan										
7	A	Tidak Ada	Laki-Laki	Memuaskan										
8	A	Tidak Ada	Laki-Laki	Cumlaude										
9	A	Tidak Ada	Laki-Laki	Cumlaude										
10	A	Tidak Ada	Laki-Laki	Sangat Memuaskan										
11	A	Tidak Ada	Laki-Laki	Cumlaude										
12	A	Tidak Ada	Perempuan	Sangat Memuaskan										
13	A	Tidak Ada	Perempuan	Cumlaude										
14	A	Tidak Ada	Perempuan	Sangat Memuaskan										
15	A	Tidak Ada	Perempuan	Cumlaude										
16	A	IPA	Laki-Laki	Sangat Memuaskan										
17	A	IPA	Laki-Laki	Cumlaude										
18	A	IPA	Laki-Laki	Sangat Memuaskan										
19	A	IPA	Laki-Laki	Sangat Memuaskan										
20	A	IPA	Laki-Laki	Cumlaude										
21	A	IPA	Laki-Laki	Sangat Memuaskan										
22	A	IPA	Perempuan	Sangat Memuaskan										
23	A	IPA	Perempuan	Cumlaude										
24	A	IPS	Laki-Laki	Memuaskan										
25	A	IPS	Laki-Laki	Sangat Memuaskan										
26	A	Kejuruan	Laki-Laki	Sangat Memuaskan										

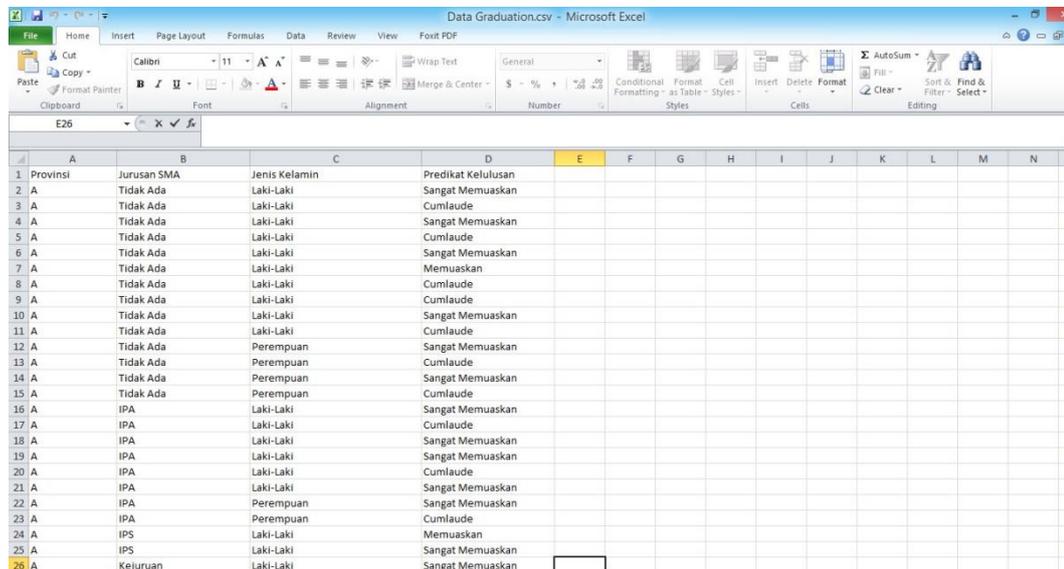
Gambar 4.8 Data yang telah diubah.

1.5. Implementasi

Pada tahap ini dilakukan pemodelan data, metode yang dipakai pada penelitian ini adalah *decision tree* (pohon keputusan) dengan menggunakan algoritma *ID3*. Data yang telah di kumpul, diseleksi dan di transformasi akan di kelola menggunakan metode *decision tree*. Metode ini adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian aturan keputusan.

Atribut data *graduation* tersebut dianalisis dengan menggunakan *software RapidMiner*. Data *graduation* fakultas teknik memiliki 272 *record data*. Gambar

4.9 adalah data graduation dari fakultas teknik dengan format .CSV yang akan diakses melalui *software RapidMiner*.



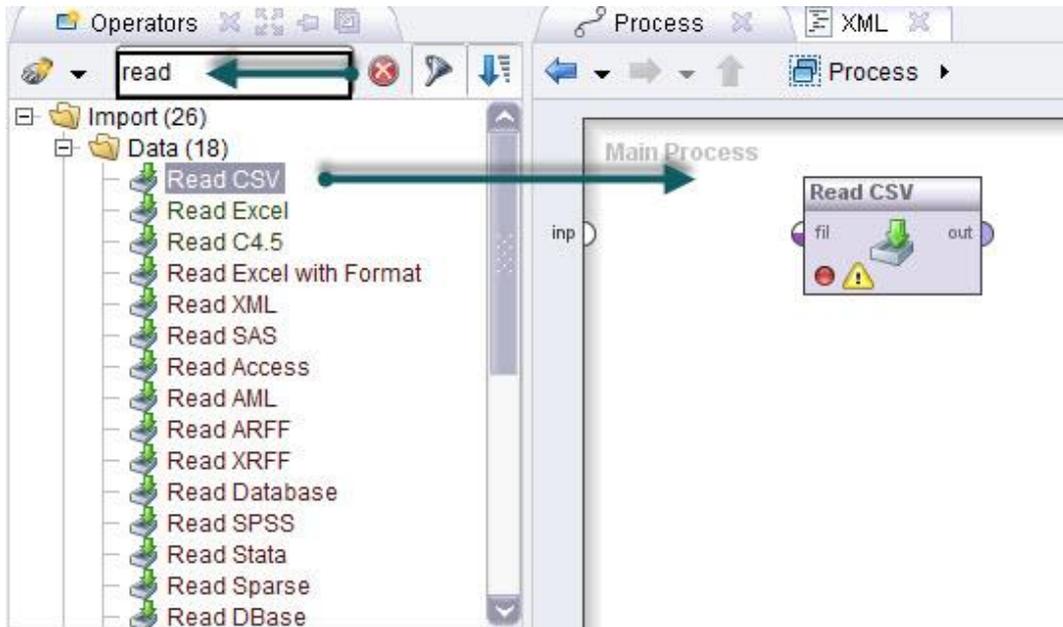
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Provinsi	Jurusan SMA	Jenis Kelamin	Predikat Kelulusan										
2	A	Tidak Ada	Laki-Laki	Sangat Memuaskan										
3	A	Tidak Ada	Laki-Laki	Cumlaude										
4	A	Tidak Ada	Laki-Laki	Sangat Memuaskan										
5	A	Tidak Ada	Laki-Laki	Cumlaude										
6	A	Tidak Ada	Laki-Laki	Sangat Memuaskan										
7	A	Tidak Ada	Laki-Laki	Memuaskan										
8	A	Tidak Ada	Laki-Laki	Cumlaude										
9	A	Tidak Ada	Laki-Laki	Cumlaude										
10	A	Tidak Ada	Laki-Laki	Sangat Memuaskan										
11	A	Tidak Ada	Laki-Laki	Cumlaude										
12	A	Tidak Ada	Perempuan	Sangat Memuaskan										
13	A	Tidak Ada	Perempuan	Cumlaude										
14	A	Tidak Ada	Perempuan	Sangat Memuaskan										
15	A	Tidak Ada	Perempuan	Cumlaude										
16	A	IPA	Laki-Laki	Sangat Memuaskan										
17	A	IPA	Laki-Laki	Cumlaude										
18	A	IPA	Laki-Laki	Sangat Memuaskan										
19	A	IPA	Laki-Laki	Sangat Memuaskan										
20	A	IPA	Laki-Laki	Cumlaude										
21	A	IPA	Laki-Laki	Sangat Memuaskan										
22	A	IPA	Perempuan	Sangat Memuaskan										
23	A	IPA	Perempuan	Cumlaude										
24	A	IPS	Laki-Laki	Memuaskan										
25	A	IPS	Laki-Laki	Sangat Memuaskan										
26	A	Kejuruan	Laki-Laki	Sangat Memuaskan										

Gambar 4.9 Data graduation fakultas teknik Format .csv.

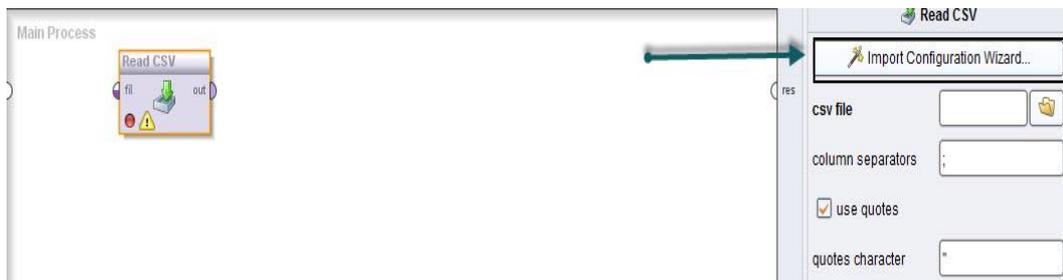
1.5.1. Pengujian *software RapidMiner*.

Atribut yang digunakan sebagai *label* adalah predikat kelulusan, penulis akan menganalisis faktor apa saja yang mempengaruhi tingkat kelulusan mahasiswa menggunakan data yang telah dipilih yaitu data *graduation* fakultas teknik. Data terlebih dahulu kita transformasi ke dalam format .csv agar bisa diakses menggunakan *software RapidMiner*.

Setelah data telah dibuat dalam bentuk format .csv selanjutnya lakukan importing data, untuk melakukan importing data dibutuhkan operator *read csv* lakukan *drag and drop* kedalam *view process* sehingga operator *read csv* muncul dalam *view process* seperti gambar 4.10. Setelah itu langkah selanjutnya langsung klik *Import Configuration Wizard* seperti gambar 4.11.

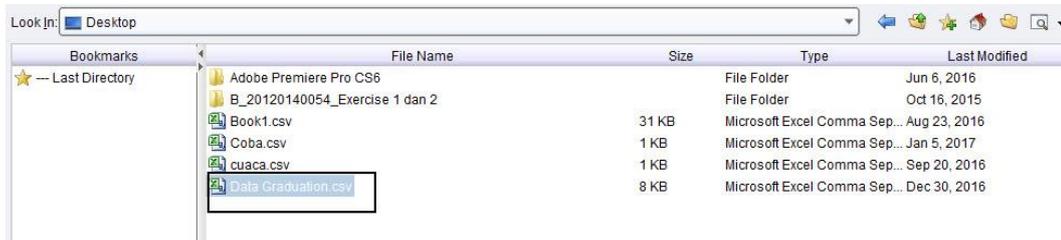


Gambar 4.10 *drag and drop read csv.*



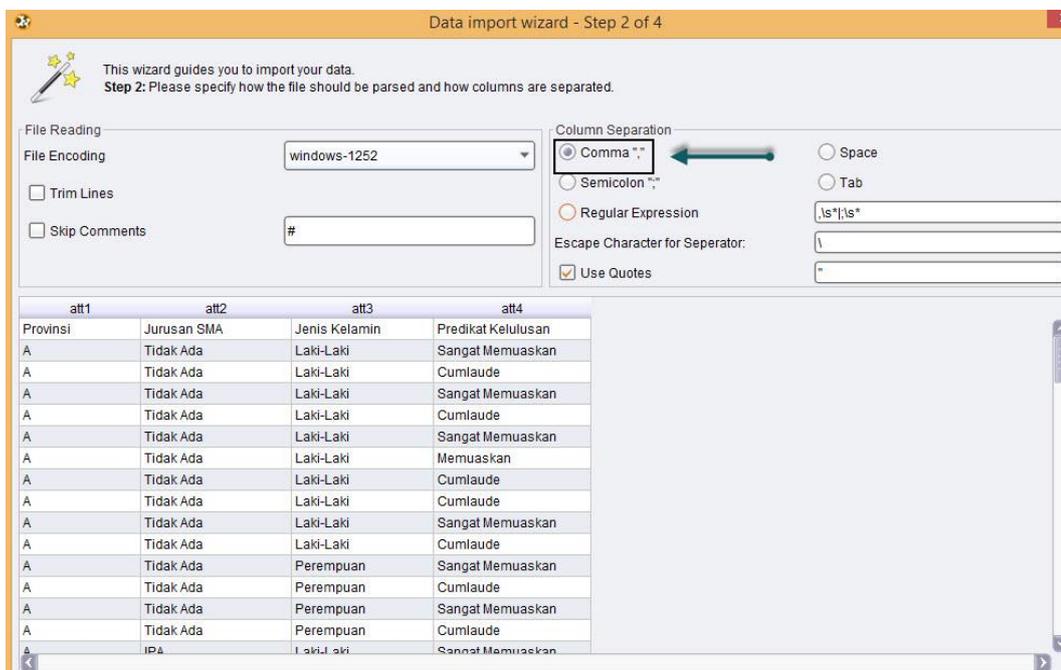
Gambar 4.11 *Import configuration wizard.*

Setelah klik *import configuration wizard* maka akan muncul form *data import wizard step 1* untuk memilih lokasi file yang digunakan untuk proses penelitian ini seperti gambar 4.12.



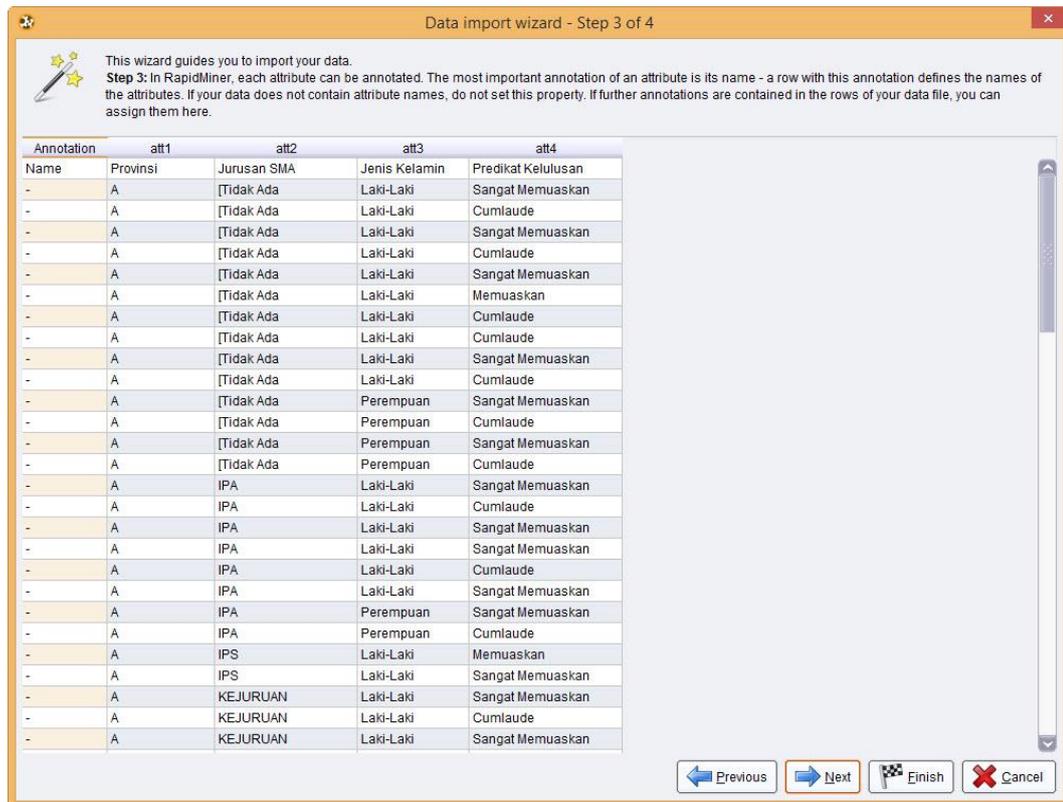
Gambar 4.12 Alur proses *import data*.

Setelah data telah kita pilih langkah selanjutnya yaitu klik **Next** dan akan muncul form *data import wizard step 2* seperti gambar 4.13.



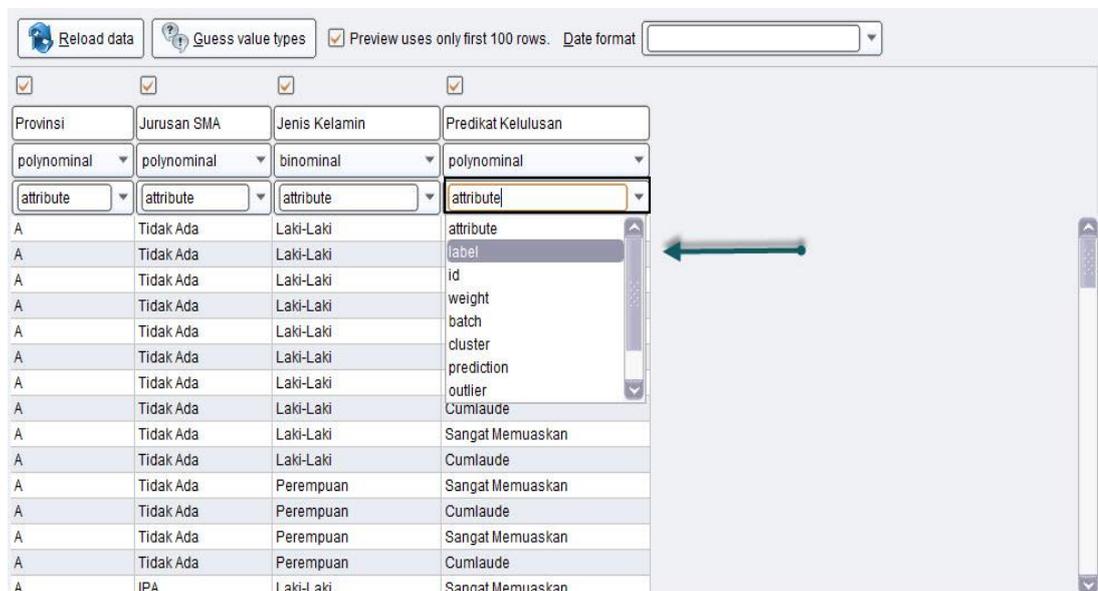
Gambar 4.13 Alur proses *import data*.

Setelah muncul *form* step ke 2 seperti gambar diatas lalu pilih *Column Separation Comma* karena *Comma* di gunakan untuk memisahkan atribut satu dengan atribut yang lainnya yang sebelumnya atribut tersebut masih menyatu. Untuk melanjutkan ke step selanjutnya klik **Next** maka akan muncul form *data import wizard step 3* seperti gambar 4.14.



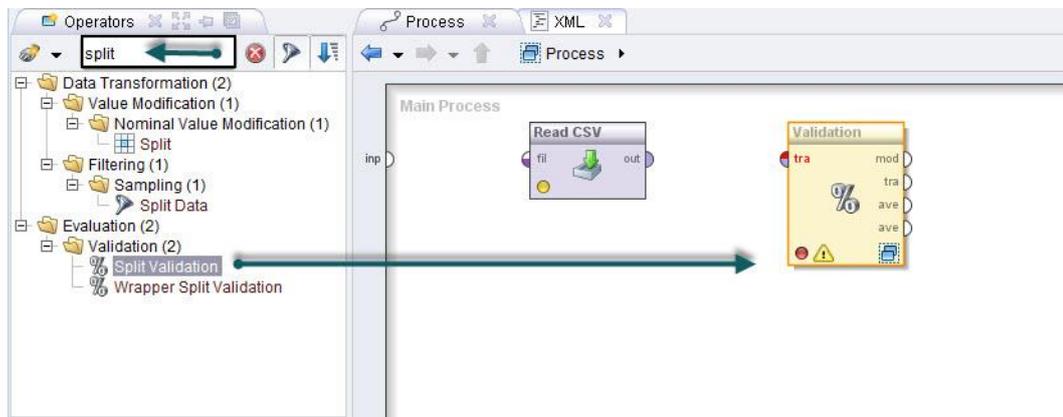
Gambar 4.14 Alur proses *import data*.

Pada step ke 3 ini tidak ada dilakukan apapun maka dari itu langsung ke *data import wizard step 4* dengan cara klik ➡ *Next* maka akan muncul form *data import wizard step 4* seperti gambar 4.15.



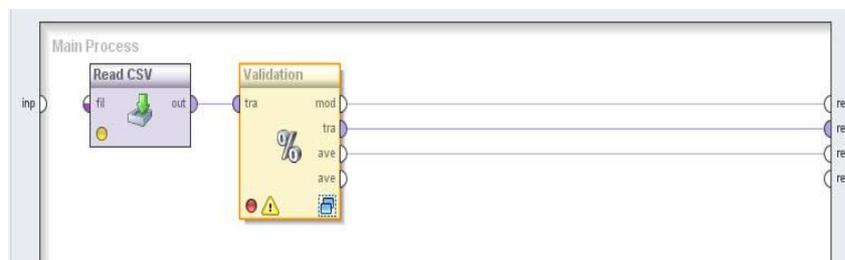
Gambar 4.15 Alur proses *import data*.

Setelah muncul *form data import wizard step 4* seperti gambar diatas lalu pilih salah satu atribut target karena pada klasifikasi tentu ada atribut target atau *label* dan atribut yang dipilih sebagai *label* yaitu atribut predikat. Setelah ditentukan label yang dipilih maka klik *Finish*. Setelah data selesai di *import* maka selanjutnya *drag and drop Split Validation* seperti gambar 4.16. Didalam *Split Validation* inilah akan memvalidasikan data yang dimodelkan kedalam algoritma dan fungsi *validation* ini adalah memaksimalkan nilai akurasi pengolahan data.



Gambar 4.16 Operator *read csv* dan *split validation*.

Selanjutnya, hubungkan operator *read csv* dengan *split validation* dengan menarik garis tabel *read csv* ke operator *split validation* dan menarik garis lagi dari operator *split validation* ke result di sisi kanan seperti gambar 4.17.

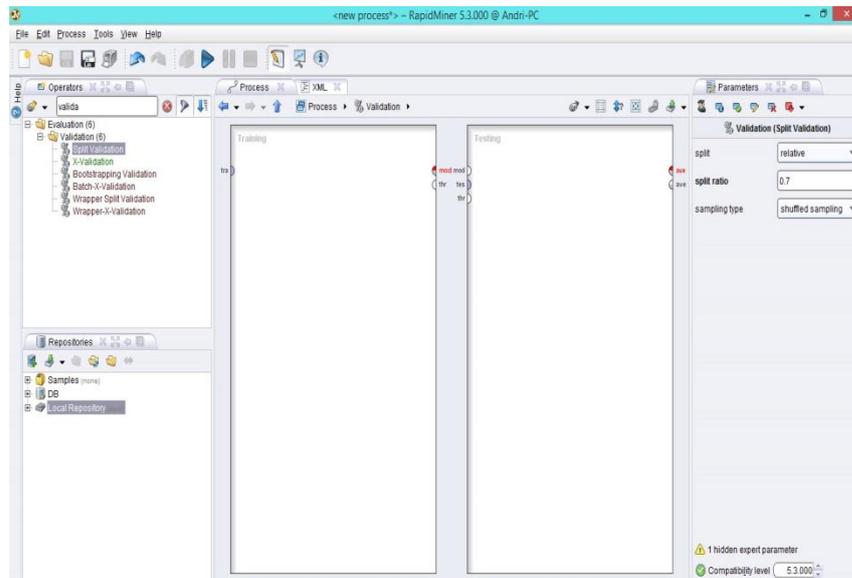


Gambar 4.17 Menghubungkan tabel *read csv* dengan operator *split validation*.

Operator *split validation* memiliki *port input* yaitu, *training example set* (tra) sebagai *port input* memperkirakan *ExampleSet* untuk melatih sebuah model (*training data set*). *ExampleSet* yang sama akan digunakan selama subproses pengujian untuk menguji model. Selain itu, operator ini juga memiliki *port output* sebagai berikut:

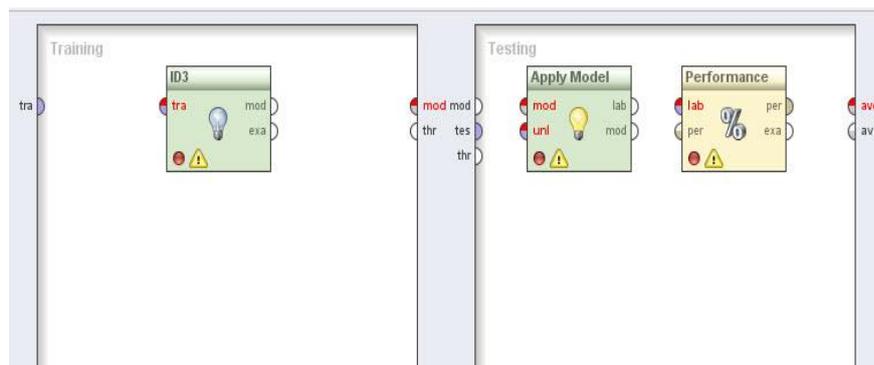
- *Model (mod)*, pelatihan *subprocess* harus mengembalikan sebuah model yang dilatih pada input *exampleset* dan model yang dibangun *exampleset* disampaikan melalui port ini.
- *Training ExampleSet (tra)*, the *exampleset* yang diberikan sebagai masukan pada port input pelatihan dilewatkan tanpa mengubah ke output melalui port ini. Port ini biasa digunakan untuk menggunakan kembali *exampleset* sama dioperator lebih lanjut atau untuk melihat *exampleset* dalam *workspace result*.
- *Averagable (ave)*, subproses pengujian harus mengembalikan *vector* kinerja. Hal ini biasanya dihasilkan dengan menerapkan model dan mengukur kinerjanya. Dua port tersebut diberikan tetapi hanya digunakan jika diperlukan. Kinerja statistic dihitung dengan skema estimasi hanya perkiraan (bukan perhitungan yang tepat) dari kinerja yang akan dicapai dengan model yang akan dibangun pada set data yang disampaikan secara lengkap.

Setelah *table operator csv* dan *split validation* telah di hubungkan maka langkah selanjutnya *double klik* pada operator *split validation* maka akan muncul form seperti gambar 4.18.



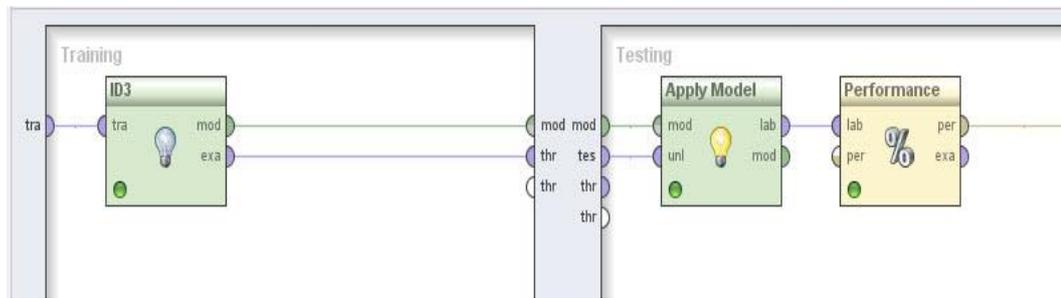
Gambar 4.18 Tampilan *split validation*.

Setelah muncul *form* seperti gambar diatas maka selanjutnya kita *drag and drop* algoritma *ID3* dari *operator* ke dalam box *training*, karena dalam penelitian kita menggunakan algoritma *ID3*, selain *ID3* kita juga *drag and drop* *apply model* dan *performance(classification)* Kedalam box *testing* Seperti gambar 4.19.



Gambar 4.19 Tampilan *split validation*.

Selanjutnya susun dan hubungkan port-port dari operator *ID3*, operator *Apply Model* dan operator *Performance* seperti terlihat pada gambar 4.20



Gambar 4.20 Susunan Operator *ID3*, *Apply Model*, *Peformance*.

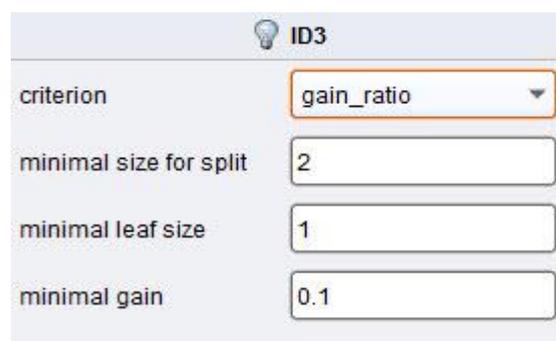
Pada operator *ID3* terdapat input *training set* (*tra*), port ini merupakan output dari operator *read csv*. Output dari operator lain juga dapat digunakan oleh port ini. Port ini menghasilkan *ExampleSet* yang dapat diproses menjadi *decision tree*. Selain itu pada operator ini juga terdapat output *model* (*mod*) dan *example set* (*exa*). *Mod* akan mengkonversi atribut yang dimasukan menjadi model keputusan dalam bentuk *decision tree*. *Exa* merupakan port yang menghasilkan output tanpa megubah inputan yang masuk melalui port ini. Port ini biasa digunakan untuk menggunakan kembali sama *ExampleSet* dioperator lebih lanjut atau untuk melihat *ExampelSet* dalam hasil *workspace*.

Pada operator *Apply Model* terdapat port input yaitu, *model* (*mod*) port ini memastikan bahwa peran atribut dari *ExampleSet* pada model yang dilatih konsisten dengan *ExampleSet* pada port input data *unlabeled*. *Unlabeled data* (*unl*) port ini memastikan bahwa peran atribut *ExampleSet* ini konsisten dengan *ExampleSet* pada model yang dikirim ke port input model dilatih. Port ini juga memiliki port output, diantaranya, *labeled data* (*lab*), model yang diberikan dalam input diterapkan pada *ExampleSet* terbaru disampaikan dari port ini. Beberapa informasi akan ditambahkan ke input *ExampleSet* sebelum dikirim melalui port

output. Dan *model (mod)*, model yang diberikan sebagai masukan dilewatkan tanpa megubah ke output melalui port ini.

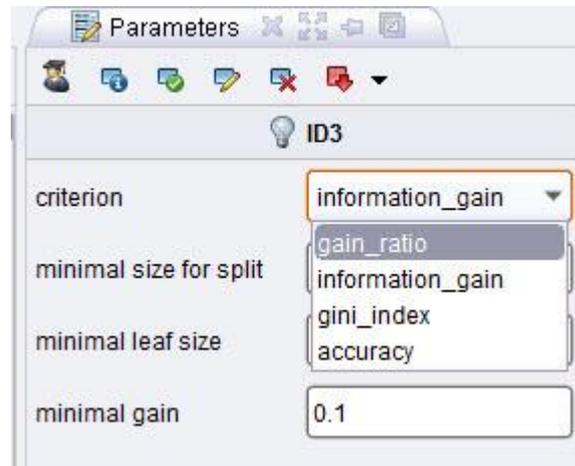
Operator *Performance* memiliki port input yaitu, *labelled data (lab)*, port ini mengharapkan *ExampleSet* berlabel. Dan *apply model* merupakan contoh yang baik dari operator yang menyediakan data yang berlabel. Pastikan bahwa *ExampleSet* memiliki atribut label dan atribut prediksi. *Performance (per)* ini adalah parameter opsional yang membutuhkan performance vector. Selain itu, operator ini juga memiliki port output yaitu, *performance (per)*, port ini memberikan *performance vector*. *Performance vector* adalah daftar nilai kinerja kriteria. *Example Set (exa)*, *example set* yang diberikan sebagai masukan dilewatkan tanpa mengubah ke output melalui port ini.

Langkah selanjutnya adalah mengatur *parameter* yang dibutuhkan. Setelah selesai menghubungkan port-port dari setiap operator atur *parameter ID3* seperti pada gambar 4.21 dan 4.22.



ID3	
criterion	gain_ratio
minimal size for split	2
minimal leaf size	1
minimal gain	0.1

Gambar 4.21 parameter ID3.



Gambar 4.22 *Criterion.*

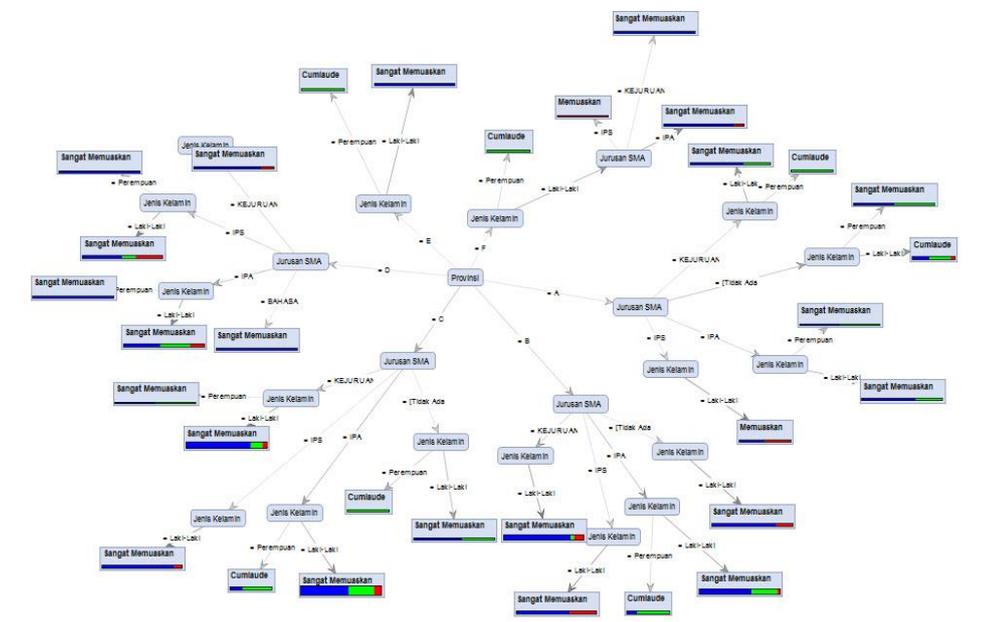
Langkah selanjutnya yaitu, memilih parameter *criterion ID3* dan parameter *criterion* yang digunakan adalah *information_gain* dengan metode ini, semua entropy dihitung. Kemudian atribut dengan entropi minimum yang dipilih untuk dilakukan perpecahan pohon (split). Metode ini memiliki bias dalam memilih atribut dengan sejumlah besar nilai. *Minimal size of split* adalah ukuran untuk membuat simpul-simpul pada *decision tree*. Simpul dibagi berdasarkan ukuran yang lebih besar dari atau sama dengan parameter *minimal size of split*. *Minimal leaf size* yaitu, pohon yang dihasilkan sedemikian rupa memiliki himpunan bagian simpul daun setidaknya sebanyak jumlah *minimal leaf size*. *Minimal gain* merupakan nilai gain minimal yang ditentukan untuk menghasilkan simpul pohon keputusan.

Setelah parameter diatur, klik ikon *run* pada *toolbar*, seperti pada gambar 4.23 untuk menampilkan hasilnya. Tunggu beberapa saat, komputer membutuhkan waktu untuk menyelesaikan perhitungan.



Gambar 4.23 Icon run

Setelah beberapa detik maka *RapidMiner* akan menampilkan hasil keputusan pada view result. Jika kita pilih *graph view* maka akan ditampilkan hasilnya berbentuk pohon keputusan (tree) seperti gambar 4.24



Gambar 4.24 Hasil berupa *graph* pohon keputusan

Tabel 4.1 Keterangan warna predikat kelulusan

Warna	Keterangan
Biru	Sangat Memuaskan
Merah	Memuaskan
Hijau	Cumaude

Hasil proses klasifikasi predicate kelulusan dengan metode *Decision tree* atau pohon keputusan ditunjukkan seperti gambar 4.24 diatas dapat dilihat bahwa atribut yang memiliki pengaruh paling tinggi untuk menentukan klasifikasi predicate kelulusan mahasiswa adalah *Provinsi* yang mana atribut ini menjadi node akar pertama. Selain menampilkan hasil *decision tree* berupa *graph* atau tampilan pohon keputusan, *RapidMiner* juga menyediakan *tool* untuk menampilkan hasil berupa *text view*, *table* dan *scatter plot view*.

```

Tree

Provinsi = A
| Jurusan SMA = IPA
| | Jenis Kelamin = Laki-Laki: Sangat Memuaskan {Sangat Memuaskan=4, Cumlaude=2, Memuaskan=0}
| | Jenis Kelamin = Perempuan: Sangat Memuaskan {Sangat Memuaskan=1, Cumlaude=1, Memuaskan=0}
| Jurusan SMA = IPS
| | Jenis Kelamin = Laki-Laki: Memuaskan {Sangat Memuaskan=1, Cumlaude=0, Memuaskan=1}
| Jurusan SMA = Kejuruan
| | Jenis Kelamin = Laki-Laki: Sangat Memuaskan {Sangat Memuaskan=2, Cumlaude=1, Memuaskan=0}
| | Jenis Kelamin = Perempuan: Cumlaude {Sangat Memuaskan=0, Cumlaude=1, Memuaskan=0}
| Jurusan SMA = Tidak Ada
| | Jenis Kelamin = Laki-Laki: Cumlaude {Sangat Memuaskan=4, Cumlaude=5, Memuaskan=1}
| | Jenis Kelamin = Perempuan: Sangat Memuaskan {Sangat Memuaskan=2, Cumlaude=2, Memuaskan=0}
Provinsi = B
| Jurusan SMA = IPA
| | Jenis Kelamin = Laki-Laki: Sangat Memuaskan {Sangat Memuaskan=18, Cumlaude=9, Memuaskan=1}
| | Jenis Kelamin = Perempuan: Cumlaude {Sangat Memuaskan=2, Cumlaude=6, Memuaskan=0}
| Jurusan SMA = IPS

```

Gambar 4.25 Hasil dari *text view* (*operator ID3*)

Seperti gambar 4.25 dapat dilihat penjelasan *RapidMiner* menggunakan *text view* untuk menjelaskan berapa jumlah mahasiswa yang cumlaude, sangat memuaskan dan memuaskan berdasarkan Jurusan SMA dan Jenis Kelamin di setiap provinsi.

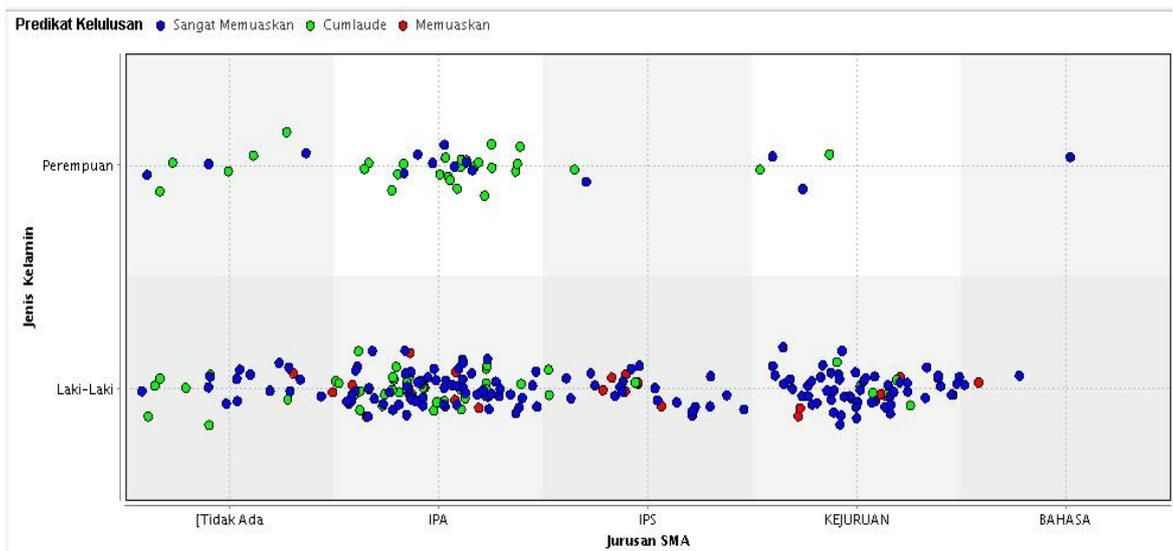
Table View Plot View

accuracy: 69.51%

	true Sangat Memuaskan	true Cumlaude	true Memuaskan	class precision
pred. Sangat Memuaskan	46	13	5	71.88%
pred. Cumlaude	7	11	0	61.11%
pred. Memuaskan	0	0	0	0.00%
class recall	86.79%	45.83%	0.00%	

Gambar 4.26 Hasil accuracy dari table (performanceVector)

Dapat dilihat pada gambar 4.26 tingkat accuracy dari *performance vector* yaitu 69.51%.



Gambar 4.27 Grafik *scatter plot view*.

Berdasarkan gambar 4.27 di atas bahwa seorang mahasiswa yang *jenis kelamin* laki-laki dan *Jurusan SMA* IPA ketika masih sekolah, memiliki probabilitas predikat kelulusan *Cumlaude* paling tinggi.

1.5.2. Algoritma ID3

Dalam penelitian algoritma yang digunakan adalah ID3 dan untuk memudahkan menjelaskan tentang algoritma ID3 dalam penelitian ini maka penulis membuat tabel data yang digunakan dalam penelitian seperti tabel 4.2

Tabel 4.2 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

NO	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusan
1	A	[Tidak Ada	Laki-Laki	Sangat Memuaskan
2	A	[Tidak Ada	Laki-Laki	Cumlaude
3	A	[Tidak Ada	Laki-Laki	Sangat Memuaskan
4	A	[Tidak Ada	Laki-Laki	Cumlaude
5	A	[Tidak Ada	Laki-Laki	Sangat Memuaskan
6	A	[Tidak Ada	Laki-Laki	Memuaskan
7	A	[Tidak Ada	Laki-Laki	Cumlaude
8	A	[Tidak Ada	Laki-Laki	Cumlaude
9	A	[Tidak Ada	Laki-Laki	Sangat Memuaskan
10	A	[Tidak Ada	Laki-Laki	Cumlaude
11	A	[Tidak Ada	Perempuan	Sangat Memuaskan
12	A	[Tidak Ada	Perempuan	Cumlaude
13	A	[Tidak Ada	Perempuan	Sangat Memuaskan
14	A	[Tidak Ada	Perempuan	Cumlaude
15	A	IPA	Laki-Laki	Sangat Memuaskan
16	A	IPA	Laki-Laki	Cumlaude
17	A	IPA	Laki-Laki	Sangat Memuaskan
18	A	IPA	Laki-Laki	Sangat Memuaskan
19	A	IPA	Laki-Laki	Cumlaude
20	A	IPA	Laki-Laki	Sangat Memuaskan

Tabel 4.3 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

NO	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusan
21	A	IPA	Perempuan	Sangat Memuaskan
22	A	IPA	Perempuan	Cumlaude
23	A	IPS	Laki-Laki	Memuaskan
24	A	IPS	Laki-Laki	Sangat Memuaskan
25	A	KEJURUAN	Laki-Laki	Sangat Memuaskan
26	A	KEJURUAN	Laki-Laki	Cumlaude
27	A	KEJURUAN	Laki-Laki	Sangat Memuaskan
28	A	KEJURUAN	Perempuan	Cumlaude
29	B	[Tidak Ada	Laki-Laki	Sangat Memuaskan
30	B	IPA	Laki-Laki	Cumlaude
31	B	IPA	Laki-Laki	Cumlaude
32	B	IPA	Laki-Laki	Cumlaude
33	B	IPA	Laki-Laki	Sangat Memuaskan
34	B	IPA	Perempuan	Cumlaude
35	B	IPA	Perempuan	Cumlaude
36	B	IPS	Laki-Laki	Memuaskan
37	B	IPS	Laki-Laki	Sangat Memuaskan
38	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
39	B	IPA	Laki-Laki	Memuaskan
40	B	IPA	Perempuan	Cumlaude
41	B	IPA	Perempuan	Sangat Memuaskan
42	B	IPS	Laki-Laki	Sangat Memuaskan
43	B	IPS	Laki-Laki	Sangat Memuaskan
44	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
45	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
46	B	KEJURUAN	Laki-Laki	Sangat Memuaskan

Tabel 4.4 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

NO	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusan
47	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
48	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
49	C	[Tidak Ada	Laki-Laki	Sangat Memuaskan
50	C	IPA	Laki-Laki	Cumlaude
51	C	IPA	Laki-Laki	Sangat Memuaskan
52	C	IPA	Laki-Laki	Cumlaude
53	C	IPA	Laki-Laki	Cumlaude
54	C	IPA	Laki-Laki	Sangat Memuaskan
55	C	IPA	Laki-Laki	Cumlaude
56	C	IPA	Laki-Laki	Memuaskan
57	C	IPA	Laki-Laki	Sangat Memuaskan
58	C	IPA	Laki-Laki	Sangat Memuaskan
59	C	IPA	Laki-Laki	Sangat Memuaskan
60	C	IPA	Laki-Laki	Sangat Memuaskan
61	C	IPA	Laki-Laki	Sangat Memuaskan
62	C	IPA	Perempuan	Sangat Memuaskan
63	C	IPA	Perempuan	Cumlaude
64	C	IPA	Perempuan	Cumlaude
65	C	IPS	Laki-Laki	Sangat Memuaskan
66	C	IPS	Laki-Laki	Sangat Memuaskan
67	C	IPS	Laki-Laki	Sangat Memuaskan
68	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
69	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
70	C	KEJURUAN	Laki-Laki	Cumlaude
71	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
72	C	KEJURUAN	Laki-Laki	Memuaskan

Tabel 4.5 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

NO	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusan
73	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
74	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
75	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
76	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
77	C	KEJURUAN	Laki-Laki	Cumlaude
78	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
79	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
80	C	KEJURUAN	Laki-Laki	Cumlaude
81	C	KEJURUAN	Perempuan	Sangat Memuaskan
82	C	IPA	Laki-Laki	Sangat Memuaskan
83	C	IPA	Perempuan	Cumlaude
84	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
85	B	[Tidak Ada	Laki-Laki	Sangat Memuaskan
86	B	IPA	Laki-Laki	Cumlaude
87	B	IPA	Laki-Laki	Cumlaude
88	B	IPA	Laki-Laki	Sangat Memuaskan
89	B	IPA	Perempuan	Cumlaude
90	B	IPS	Laki-Laki	Memuaskan
91	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
92	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
93	C	[Tidak Ada	Perempuan	Cumlaude
94	C	IPA	Laki-Laki	Sangat Memuaskan
95	C	IPA	Laki-Laki	Sangat Memuaskan
96	C	IPA	Laki-Laki	Sangat Memuaskan
97	C	IPA	Laki-Laki	Sangat Memuaskan
98	C	IPA	Laki-Laki	Memuaskan

Tabel 4.6 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

NO	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusan
99	C	IPA	Laki-Laki	Sangat Memuaskan
100	C	IPA	Laki-Laki	Cumlaude
101	C	IPA	Laki-Laki	Cumlaude
102	C	IPA	Laki-Laki	Memuaskan
103	C	IPA	Laki-Laki	Sangat Memuaskan
104	C	IPA	Laki-Laki	Cumlaude
105	C	IPA	Laki-Laki	Cumlaude
106	C	IPA	Perempuan	Cumlaude
107	C	IPA	Perempuan	Sangat Memuaskan
108	C	IPS	Laki-Laki	Sangat Memuaskan
109	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
110	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
111	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
112	C	[Tidak Ada	Laki-Laki	Sangat Memuaskan
113	C	[Tidak Ada	Laki-Laki	Cumlaude
114	C	[Tidak Ada	Laki-Laki	Sangat Memuaskan
115	C	[Tidak Ada	Perempuan	Cumlaude
116	C	[Tidak Ada	Perempuan	Cumlaude
117	C	IPA	Laki-Laki	Sangat Memuaskan
118	C	IPA	Laki-Laki	Sangat Memuaskan
119	C	IPA	Laki-Laki	Sangat Memuaskan
120	C	IPA	Laki-Laki	Memuaskan
121	C	IPA	Laki-Laki	Cumlaude
122	C	IPA	Laki-Laki	Sangat Memuaskan
123	C	IPA	Laki-Laki	Sangat Memuaskan
124	C	IPA	Laki-Laki	Cumlaude

Tabel 4.7 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

NO	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusan
125	C	IPA	Laki-Laki	Sangat Memuaskan
126	C	IPA	Laki-Laki	Sangat Memuaskan
127	C	IPA	Laki-Laki	Sangat Memuaskan
128	C	IPA	Laki-Laki	Cumlaude
129	C	IPA	Laki-Laki	Sangat Memuaskan
130	C	IPA	Laki-Laki	Cumlaude
131	C	IPA	Laki-Laki	Sangat Memuaskan
132	C	IPA	Laki-Laki	Sangat Memuaskan
133	C	IPA	Laki-Laki	Cumlaude
134	C	IPA	Laki-Laki	Sangat Memuaskan
135	C	IPA	Laki-Laki	Sangat Memuaskan
136	C	IPA	Laki-Laki	Sangat Memuaskan
137	C	IPA	Laki-Laki	Sangat Memuaskan
138	C	IPA	Laki-Laki	Sangat Memuaskan
139	C	IPA	Laki-Laki	Cumlaude
140	C	IPA	Perempuan	Cumlaude
141	C	IPA	Perempuan	Cumlaude
142	C	IPA	Perempuan	Cumlaude
143	C	IPA	Perempuan	Cumlaude
144	C	IPA	Perempuan	Sangat Memuaskan
145	C	IPA	Perempuan	Cumlaude
146	C	IPA	Perempuan	Sangat Memuaskan
147	C	IPA	Perempuan	Cumlaude
148	C	IPS	Laki-Laki	Sangat Memuaskan
149	C	IPS	Laki-Laki	Sangat Memuaskan
150	C	IPS	Laki-Laki	Memuaskan

Tabel 4.8 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

NO	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusan
151	C	IPS	Laki-Laki	Sangat Memuaskan
152	C	IPS	Laki-Laki	Sangat Memuaskan
153	C	IPS	Laki-Laki	Sangat Memuaskan
154	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
155	C	KEJURUAN	Laki-Laki	Cumlaude
156	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
157	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
158	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
159	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
160	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
161	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
162	C	KEJURUAN	Laki-Laki	Memuaskan
163	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
164	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
165	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
166	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
167	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
168	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
169	C	KEJURUAN	Laki-Laki	Sangat Memuaskan
170	C	[Tidak Ada	Laki-Laki	Cumlaude
171	C	IPA	Laki-Laki	Cumlaude
172	C	IPA	Laki-Laki	Cumlaude
173	C	IPA	Laki-Laki	Sangat Memuaskan
174	C	IPA	Laki-Laki	Memuaskan
175	C	IPA	Laki-Laki	Sangat Memuaskan
176	C	IPA	Laki-Laki	Cumlaude

Tabel 4.9 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

NO	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusa
177	C	IPA	Perempuan	Sangat Memuaskan
178	C	IPA	Perempuan	Cumlaude
179	C	KEJURUAN	Laki-Laki	Cumlaude
180	C	KEJURUAN	Perempuan	Cumlaude
181	D	IPA	Laki-Laki	Cumlaude
182	D	IPA	Perempuan	Sangat Memuaskan
183	D	IPS	Laki-Laki	Memuaskan
184	D	IPS	Perempuan	Sangat Memuaskan
185	D	KEJURUAN	Laki-Laki	Sangat Memuaskan
186	D	IPA	Laki-Laki	Cumlaude
187	D	IPA	Laki-Laki	Sangat Memuaskan
188	D	KEJURUAN	Laki-Laki	Sangat Memuaskan
189	D	KEJURUAN	Laki-Laki	Memuaskan
190	D	BAHASA	Perempuan	Sangat Memuaskan
191	D	IPA	Laki-Laki	Cumlaude
192	D	IPA	Laki-Laki	Sangat Memuaskan
193	D	IPA	Laki-Laki	Cumlaude
194	D	IPA	Laki-Laki	Sangat Memuaskan
195	D	IPA	Laki-Laki	Sangat Memuaskan
196	D	IPS	Laki-Laki	Memuaskan
197	D	KEJURUAN	Laki-Laki	Sangat Memuaskan
198	D	KEJURUAN	Laki-Laki	Sangat Memuaskan
199	D	IPA	Laki-Laki	Memuaskan
200	D	IPS	Laki-Laki	Sangat Memuaskan
201	D	IPS	Laki-Laki	Sangat Memuaskan
202	D	IPS	Laki-Laki	Cumlaude

Tabel 4.10 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

NO	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusan
203	D	IPS	Laki-Laki	Sangat Memuaskan
204	D	KEJURUAN	Laki-Laki	Sangat Memuaskan
205	D	IPA	Laki-Laki	Sangat Memuaskan
206	D	IPA	Laki-Laki	Memuaskan
207	B	IPA	Laki-Laki	Sangat Memuaskan
208	B	IPA	Perempuan	Sangat Memuaskan
209	B	IPA	Perempuan	Cumlaude
210	B	IPS	Laki-Laki	Sangat Memuaskan
211	B	[Tidak Ada	Laki-Laki	Sangat Memuaskan
212	B	[Tidak Ada	Laki-Laki	Sangat Memuaskan
213	B	IPA	Laki-Laki	Sangat Memuaskan
214	B	IPA	Laki-Laki	Sangat Memuaskan
215	B	IPA	Laki-Laki	Cumlaude
216	B	IPA	Laki-Laki	Cumlaude
217	B	IPA	Laki-Laki	Sangat Memuaskan
218	B	IPA	Perempuan	Cumlaude
219	F	IPA	Laki-Laki	Sangat Memuaskan
220	F	IPA	Laki-Laki	Memuaskan
221	F	IPA	Laki-Laki	Sangat Memuaskan
222	F	IPA	Laki-Laki	Sangat Memuaskan
223	F	IPS	Laki-Laki	Memuaskan
224	B	IPA	Laki-Laki	Sangat Memuaskan
225	B	IPA	Laki-Laki	Sangat Memuaskan
226	B	IPA	Laki-Laki	Sangat Memuaskan
227	F	IPA	Laki-Laki	Sangat Memuaskan
228	F	IPA	Perempuan	Cumlaude

Tabel 4.11 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

No	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusan
229	F	KEJURUAN	Laki-Laki	Sangat Memuaskan
230	F	IPA	Laki-Laki	Sangat Memuaskan
231	F	IPA	Perempuan	Cumlaude
232	F	IPA	Laki-Laki	Sangat Memuaskan
233	F	IPA	Laki-Laki	Sangat Memuaskan
234	F	IPA	Perempuan	Cumlaude
235	B	[Tidak Ada	Laki-Laki	Memuaskan
236	B	IPA	Laki-Laki	Cumlaude
237	B	IPA	Laki-Laki	Sangat Memuaskan
238	B	IPA	Laki-Laki	Cumlaude
239	B	IPA	Laki-Laki	Sangat Memuaskan
240	B	IPA	Laki-Laki	Sangat Memuaskan
241	B	IPS	Laki-Laki	Memuaskan
242	B	IPS	Laki-Laki	Sangat Memuaskan
243	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
244	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
245	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
246	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
247	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
248	B	KEJURUAN	Laki-Laki	Memuaskan
249	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
250	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
251	E	[Tidak Ada	Laki-Laki	Sangat Memuaskan
252	E	IPA	Laki-Laki	Sangat Memuaskan
253	E	KEJURUAN	Laki-Laki	Sangat Memuaskan
254	E	IPS	Laki-Laki	Sangat Memuaskan

Tabel 4.12 Data kelulusan mahasiswa Fakultas Teknik tahun 2013-2015 (UMY)

NO	Provinsi	Jurusan SMA	Jenis kelamin	Predikat kelulusan
255	E	IPA	Laki-Laki	Sangat Memuaskan
256	E	IPA	Perempuan	Cumlaude
257	B	IPA	Laki-Laki	Sangat Memuaskan
258	B	IPA	Laki-Laki	Sangat Memuaskan
259	B	IPA	Laki-Laki	Sangat Memuaskan
260	B	IPA	Laki-Laki	Sangat Memuaskan
261	B	IPS	Laki-Laki	Sangat Memuaskan
262	B	KEJURUAN	Laki-Laki	Memuaskan
263	B	KEJURUAN	Laki-Laki	Memuaskan
264	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
265	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
266	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
267	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
268	B	IPA	Laki-Laki	Sangat Memuaskan
269	B	IPA	Laki-Laki	Sangat Memuaskan
270	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
271	B	KEJURUAN	Laki-Laki	Sangat Memuaskan
272	B	KEJURUAN	Laki-Laki	Cumlaude

Tabel 4.13 Keterangan tentang atribut provinsi

Provinsi	Keterangan	Jumlah
A	Tidak Ada	28
B	Sumatera dan Sekitarnya	74
C	Jawa Dan Sekitarnya	125
D	Kalimantan	26
E	Sulawesi	6
F	Papua, Nusa Tenggara, Maluku dan sekitarnya	13

Tabel 4.2 diatas terdiri dari 272 *record* 272 data yang digunakan dalam penelitian ini. Dalam kasus yang ada dalam tabel 4.2 akan dibuat pohon keputusan untuk menentukan faktor apa saja yang membuat mahasiswa Fakultas Teknik cumlaude menggunakan atribut *Provinsi, Jenis Kelamin dan Jurusan SMA* dan atribut *Predikat kelulusan* sebagai atribut target. Cara untuk membangun pohon keputusan menggunakan Algoritma *ID3* seperti berikut:

1. Pilih atribut sebagai akar.
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, harus didasarkan pada nilai *gain* tertinggi dari Setiap atribut yang ada. Untuk mendapatkan nilai *gain* dari setiap atribut penulis harus lebih dulu menghitung nilai *entropy TOTAL* dan nilai *entropy* tiap atribut. Hasil perhitungan *entropy* dan *Gain* yang dilakukan oleh penulis di tunjukan pada table 4.4.

16			Laki-laki	229	157	25	47	1.19	
17			Perempuan	43	14	0	29	0.91	

Untuk menghitung nilai *entropy* **TOTAL** yang ada pada baris ke 1 dengan menggunakan cara seperti berikut.

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

- $Entropy(Total) = (-\frac{171}{272} * \log_2(\frac{171}{272})) + (-\frac{25}{272} * \log_2(\frac{25}{272})) + (-\frac{76}{272} * \log_2(\frac{76}{272}))$

$$Entropy(Total) = 1.251$$

Setelah menghitung *entropy* **TOTAL** dan mendapatkan hasil, selanjutnya penulis menghitung semua *entropy* yang pada atribut Provinsi yang ada pada baris ke 3.

- $Entropy(A) = (-\frac{14}{28} * \log_2(\frac{14}{28})) + (-\frac{2}{28} * \log_2(\frac{2}{28})) + (-\frac{12}{28} * \log_2(\frac{12}{28}))$

$$Entropy(A) = 1.295$$

- $Entropy(B) = (-\frac{50}{74} * \log_2(\frac{50}{74})) + (-\frac{8}{74} * \log_2(\frac{8}{74})) + (-\frac{16}{74} * \log_2(\frac{16}{74}))$

$$Entropy(B) = 1.206$$

Perhitungan *entropy* pada atribut **Provinsi** diatas hanya 2 dari 6 dan yang lain terlampir pada lampiran begitu juga proses perhitungan semua *entropy* pada atribut lain yaitu pada atribut **Jurusan SMA** dan **Kelamin**. Setelah *entropy* semua atribut telah dihitung langkah selanjutnya menghitung nilai *gain* setiap atribut karena nilai *gain* tertinggi akan menjadi node akar. Dan atribut yang pertama kali di hitung nilai *gain*nya yaitu atribut **Provinsi**

Untuk menghitung nilai *gain* atribut Provinsi yaitu dengan menggunakan cara seperti berikut:

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

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

$$Gain(Total, Propinsi) = 1.25 - \left(\left(\frac{28}{272} * 1.295 \right) + \left(\frac{74}{272} * 1.206 \right) + \left(\frac{125}{272} * 0.678 \right) + \left(\frac{26}{272} * 1.345 \right) + \left(\frac{6}{272} * 0.65 \right) + \left(\frac{13}{272} * 1.334 \right) \right)$$

$$Gain(Total, Propinsi) = 1.251 - ((0.133) + (0.328) + (0.311) + (0.128) + (0.014) + (0.063))$$

$$Gain(Total, Propinsi) = 1.251 - 0.977$$

$$= 0.274$$

Perhitungan di atas merupakan hasil proses perhitungan nilai *gain* atribut **provinsi** dan nilai *gain* yang didapat yaitu **0.274** yang dimana merupakan nilai *gain* tertinggi diantara atribut yang lainnya dengan begitu **Provinsi** dapat menjadi node akar. Ada 6 nilai atribut dari **Provinsi**, yaitu bisa di lihat dari tabel 4.4 mulai dari baris ke 3 sampai ke baris 8.

Berdasarkan penelitian yang dilakukan, dapat disimpulkan bahwa algoritma *ID3* bisa digunakan untuk mengelompokan tingkat kelulusan mahasiswa berdasarkan atribut Provinsi, Jurusan SMA, dan Jenis kelamin dari data yang dilatih didapatkan 7 kelompok yaitu:

1. Mahasiswa cumlaude = IPA dan laki-laki sebanyak 32 mahasiswa dari 272

Mahasiswa (11%)

2. Mahasiswa cumlaude= IPA dan Perempuan sebanyak 22 mahasiswa dari
272 mahasiswa (8%)
3. Mahasiswa cumlaude = Tidak ada dan Laki-laki sebanyak 7 mahasiswa
dari 272 mahasiswa (2,6%)
4. Mahasiswa cumlaude = Tidak ada dan perempuan sebanyak 5 mahasiswa
dari 272 mahasiswa (1,9%)
5. Mahasiswa cumlaude = Kejuruan dan Laki-laki sebanyak 7 mahasiswa
Dari 272 mahasiswa (2,6%)
6. Mahasiswa cumlaude = Kejuruan dan Perempuan sebanyak 2 mahasiswa
Dari 272 mahasiswa (0,7%)
7. Mahasiswa cumlaude = IPS dan Laki-laki hanya 1 mahasiswa dari 272
Mahasiswa (0,3%).