

BAB IV

HASIL PENGUJIAN DAN PEMBAHASAN

Hasil pengujian adalah hasil *final* yang berarti penelitian telah selesai dilakukan, semua yang berkaitan dengan pengujian akan dibahas pada bab ini mulai dari pengolahan *dataset* hingga pengujian data pada *software*. Pada penelitian ini peneliti memakai data dosen Fakultas Teknik Universitas Muhammadiyah Yogyakarta dengan atribut NIDN (Nomor Induk Dosen Nasional), nama, usia, alamat, status, lama kerja dan *work unit name*. Jumlah data yang dipakai berjumlah 31 yang terbagi atas empat fakultas yaitu Fakultas Teknik Sipil, Teknik Elektro, Teknik Mesin dan Teknologi Informasi.

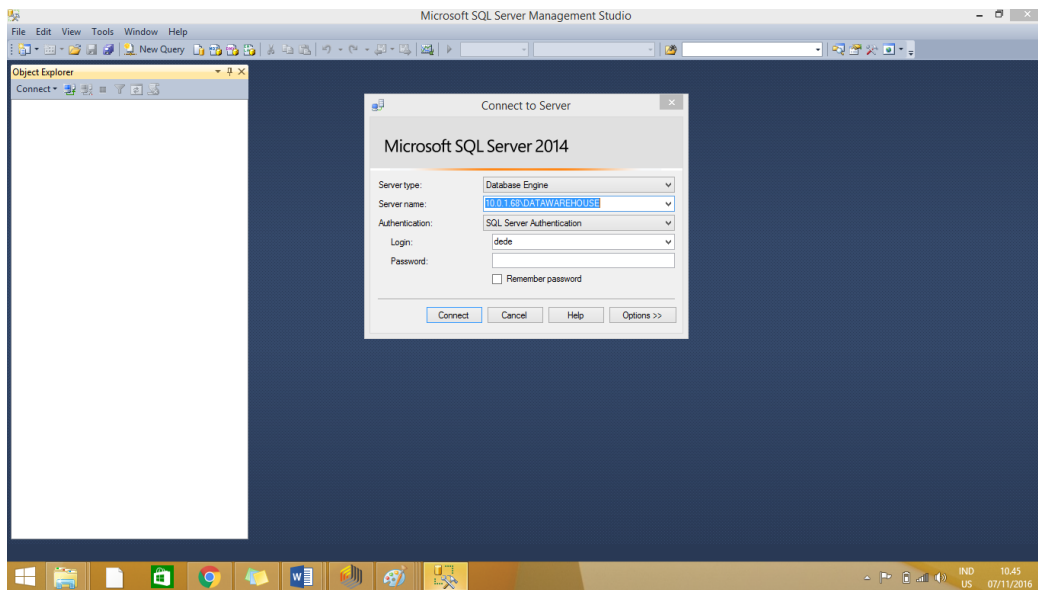
4.1. Pengambilan Data

Pengambilan data ini menggunakan *software sql server 2014* berikut antar muka *sql server 2014* seperti gambar 4.1.



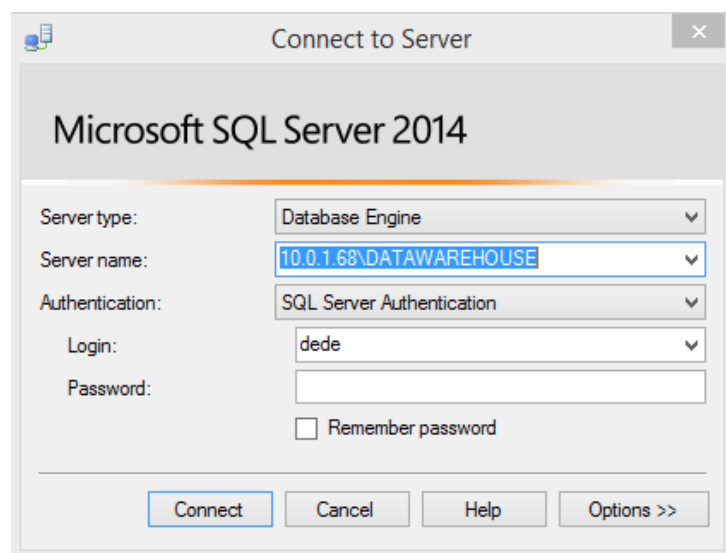
Gambar 4.1 Antar Muka *Sql Server 2014*

Sql server 2014 ini merupakan untuk menghubungkan data yang ada pada *server* ini bisa mengelola data secara fisik dan membuat tabel – tabel. Setelah menunggu beberapa detik akan muncul antar muka awal dari *sql server 2014* seperti pada gambar 4.2.



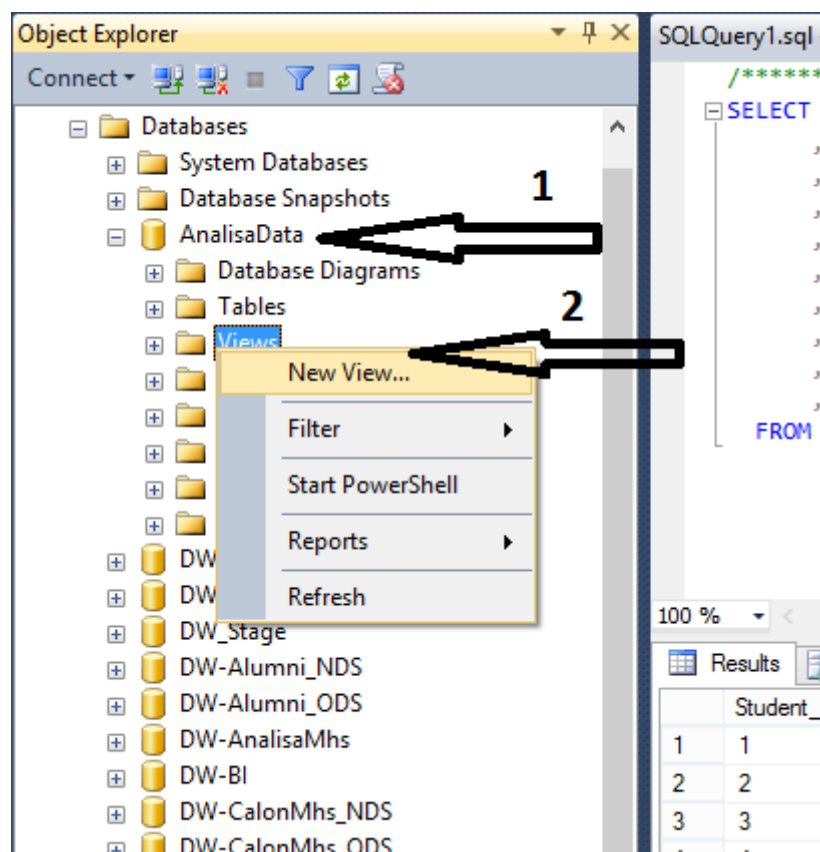
Gambar 4.2 Antar Muka Awal *Sql Server 2014*

Untuk masuk ke *database server* peneliti harus *login* terlebih dahulu, untuk *login* diperlukan hak akses yang sudah diberikan kepada admin. *Server name* dari *database server* BSI (Biro Sistem Informasi) adalah 10.0.1.68\DATAWAREHOUSE seperti pada gambar 4.3.



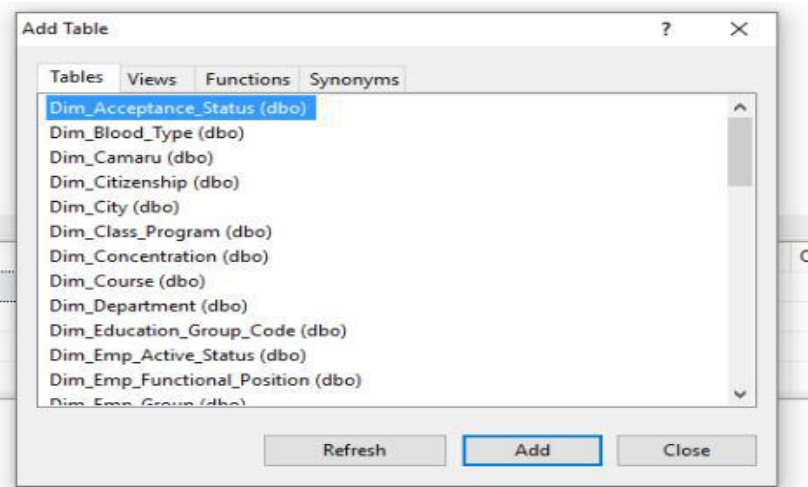
Gambar 4.3 Menghubungkan ke *Server*

login peneliti adalah “dede” dan *password* (dirahasiakan), peneliti menggunakan *login* dari penelitian lain karena milik peneliti sendiri terkendala dengan *password* yang tidak bisa diakses jadi peneliti menggunakan *login* yang sama dengan peneliti lain yang topiknya sama dengan peneliti. Setelah masuk ke *database server* memilih dimana data – data tersebut disimpan, yaitu di *database* AnalisaData yang bisa dilihat pada gambar 4.4.



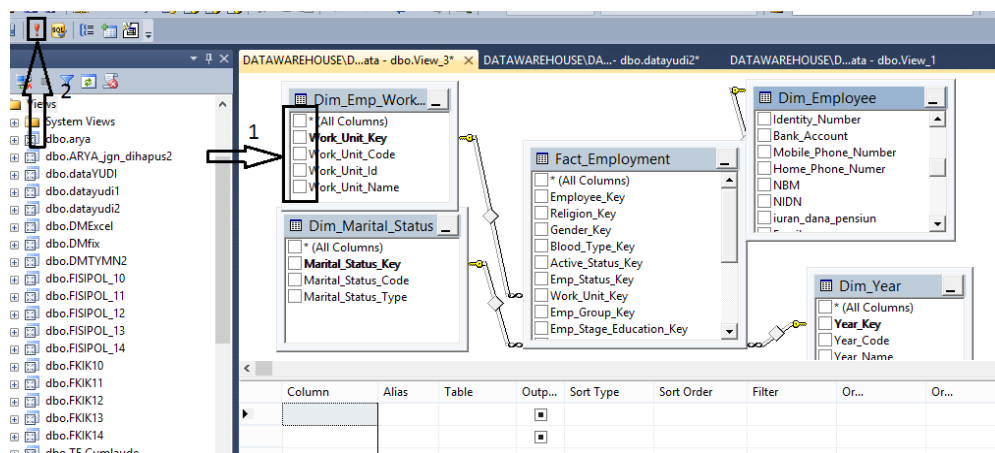
Gambar 4.4 Database server BSI

Pada *database* AnalisaData terdapat folder *view*, peneliti ingin membuat sebuah data baru dari data yang sudah ada agar lebih mudah dan bisa digunakan jika sewaktu – waktu ingin mengambil data tersebut. Langkah – langkahnya klik tanda tambah pada *database* AnalisaData lalu klik kanan pada folder *view* lalu *new view*.



Gambar 4.5 Membuat Tabel Baru

Pada antar muka *add table* seperti gambar diatas langkah selanjutnya memilih tabel – tabel mana saja yang dibutuhkan untuk membuat data baru. Peneliti menggunakan tabel *Dim_Employee* yang berisi data – data dari seluruh karyawan Universitas Muhammadiyah Yogyakarta seperti nama, alamat, NIDN, nomor *handphone* dan masih banyak lagi. Tabel *Dim_Emp_Work_Unit* berisi data tempat dimana karyawan bekerja tabel *Dim_Marital_Status* sebagai data apakah karyawan telah menikah atau belum dan tabel *Dim_Year* berisi data tahun masuk kerja karyawan sedangkan tabel *Fact_Employment* sebagai penghubung antar tabel agar menjadi satu. Setelah selesai pilih *add* hingga muncul seperti gambar 4.6.



Gambar 4.6 Antar muka Membuat View Baru

Setelah tampak tabel – tabel seperti gambar diatas centang yang diperlukan. Dimana peneliti memilih *NIDN*, *Employee_Title_Name*, *Address*, *Date_Of_Birth* dari tabel *Dim_Employee*, *Marital_Status* dari tabel *Dim_Marital_Status*, *Year_Name* dari tabel *Dim_Year* dan *Work_Unit_Name* dari tabel *Dim_Emp_Work_unit*. Untuk mencentang tabel – tabel yang dipilih harus berurutan karena hasilnya akan muncul seperti urutan yang dicentang tadi misalnya jika *NIDN* dicentang pertama maka *NIDN* akan berada pada urutan pertama begitu seterusnya. setelah selesai lalu langkah selanjutnya adalah eksekusi seperti tanda ke 2 pada gambar 4.6 hingga muncul hasilnya seperti gambar 4.7.

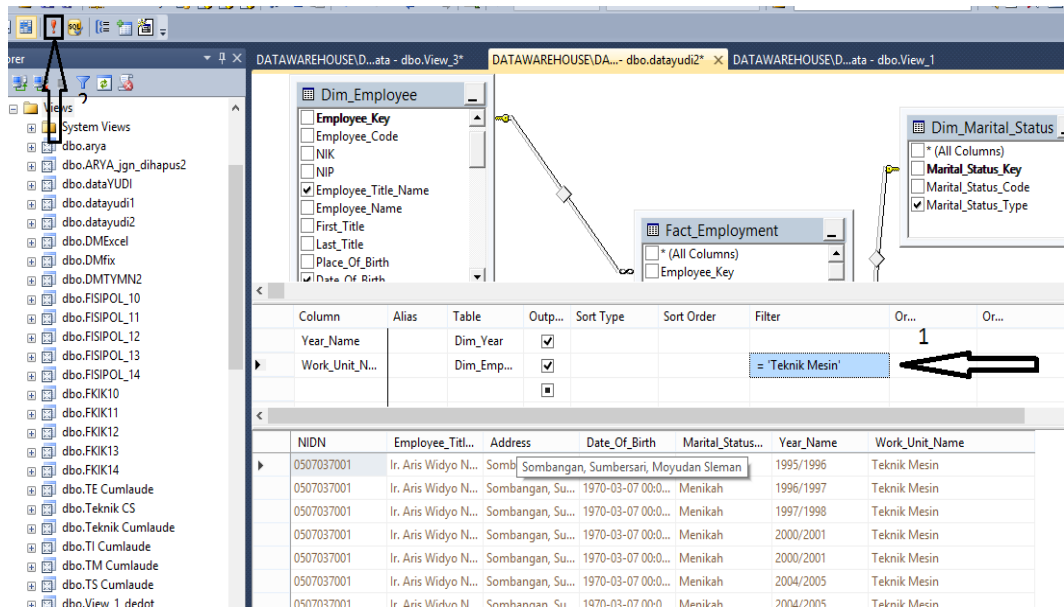
Column	Alias	Table	Order	Sort Type	Sort Order	Filter	Or	Or	Or
	NIDN	Employee_Titl...	Address	Date_Of_Birth	Marital_Status...	Year_Name	Work_Unit_Name		
▶	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	1988/1989	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	1988/1989	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	1999/2000	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	2002/2003	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	2002/2003	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	2005/2006	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	2008/2009	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	2008/2009	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	2009/2010	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	2009/2010	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	2009/2010	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	2012/2013	Pendidikan Agama Islam (PAI)		
	504036101	Dr. Abd. Madjid...	Godegan RT 07...	1961-03-04 00:0...	Menikah	2014/2015	Pendidikan Agama Islam (PAI)		

1 of 7664 | Cell is Read Only.

Gambar 4.7 Hasil Dari Pembuatan Data Baru

Pada gambar 4.7 hasil telah didapatkan dan hasil dari seluruh dosen Universitas Muhammadiyah Yogyakarta adalah 7664 orang data ini berbanding terbalik dari data dikti yang hanya berjumlah 600 orang setelah peneliti cermati ternyata banyak kesamaan data dari data dosen tersebut. Pada saat peneliti mencoba melakukan *distinct* (menghapus kesamaan data) ternyata data tidak berubah hal ini karena *Year_Name* yang berbeda beda tidak masalah sebenarnya karena data yang peneliti butuhkan bukanlah keseluruhan dari

data dosen melainkan hanya mengambil dari fakultas teknik saja. Untuk langkah selanjutnya silahkan lihat gambar 4.8.



Gambar 4.8 Filter Data

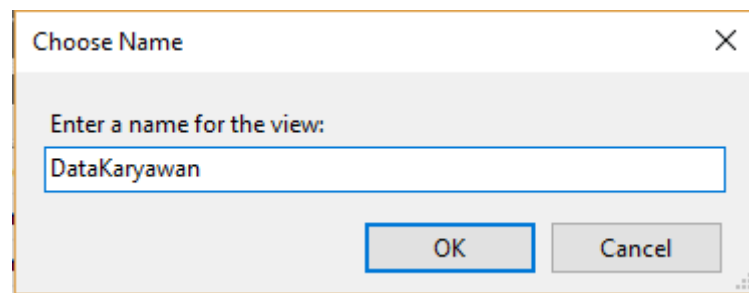
Langkah selanjutnya yaitu menyaring (*filter*), ini adalah teknik menyatukan semua data dari data yang sama. Peneliti menggunakan teknik ini agar tidak memilih data satu persatu karena hal ini sangat membuang waktu dan langkah – langkahnya yaitu pilih *filter* pada *column work_unit_name* letaknya ada diatas hasil, dan tuliskan Teknik Mesin pada *filter* tersebut lalu eksekusi dan hasilnya bisa dilihat pada gambar 4.9.

NIDN	Employee_Titl...	Date_Of_Birth	Marital_Status...	Year_Name	Work_Unit_Name
0514087501	Dr. Eng. Agus S...	1975-08-14 00:0...	Tidak Menikah	1999/2000	Teknik Sipil
0514087501	Dr. Eng. Agus S...	1975-08-14 00:0...	Tidak Menikah	2001/2002	Teknik Sipil
0514087501	Dr. Eng. Agus S...	1975-08-14 00:0...	Tidak Menikah	2004/2005	Teknik Sipil
0514087501	Dr. Eng. Agus S...	1975-08-14 00:0...	Tidak Menikah	2004/2005	Teknik Sipil
0514087501	Dr. Eng. Agus S...	1975-08-14 00:0...	Tidak Menikah	2004/2005	Teknik Sipil
0514087501	Dr. Eng. Agus S...	1975-08-14 00:0...	Tidak Menikah	2006/2007	Teknik Sipil
0514087501	Dr. Eng. Agus S...	1975-08-14 00:0...	Tidak Menikah	2011/2012	Teknik Sipil
0514087501	Dr. Eng. Agus S...	1975-08-14 00:0...	Tidak Menikah	2013/2014	Teknik Sipil
0514087501	Dr. Eng. Agus S...	1975-08-14 00:0...	Tidak Menikah	2015/2016	Teknik Sipil

Gambar 4.9 Hasil Dari Filter Data

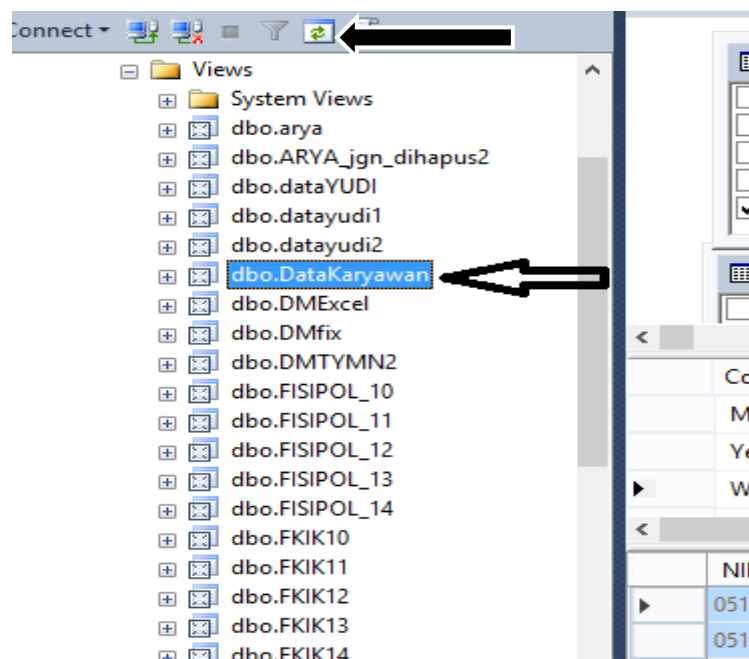
Copy kan semua data ke *excel* dan ulangi langkah tersebut untuk mengambil data dari Teknik Elektro, Teknik Sipil dan Teknologi Informasi.

Setelah semua data tadi diambil yaitu menyimpan data baru tersebut agar jika sewaktu – waktu dibutuhkan tinggal mencarinya di *folder view*. Peneliti menyimpan data baru dengan diberi nama DataKaryawan seperti gambar dibawah ini.



Gambar 4.10 Menyimpan Data

Dan jika DataKaryawan tadi belum muncul di *folder views refresh* seperti gambar 4.11. jika belum muncul juga *logout* dari *database server* lalu *login* lagi.



Gambar 4.11 Data DataKaryawan Sudah Ada difolder Views

4.2. Data Sampel

Data sampel ini adalah data yang akan digunakan dalam pengujian yang diambil dengan cara *copy* data dari *server* lalu *paste* di Office Excel 2013 untuk lebih jelasnya bisa dilihat gambar dibawah ini.

NIDN	Nama	Alamat	Usia	Status	Lama Kerja	Work unit name
514087501	Dr. Eng. Agus Setyo M, M.Eng.Sc.	Taman Sedayu I/1C RT. 44 Sedayu Bantul	1975-08-14 00:00:00.000	Tidak Menikah	1999/2000	Teknik Sipil
20076502	Ir. Anita Widianti, MT.	Banyumeneng RT 16 RW 04 No. 469 A Banyuraden Gamping, Sleman	1965-07-20 00:00:00.000	Tidak Menikah	1993/1994	Teknik Sipil
514046601	Ir. As'at Pujianto, M.T.	Kembang Tamantirto Kasihlan Bantul Griya Ketawang	1966-04-14 00:00:00.000	Tidak Menikah	1993/1994	Teknik Sipil
525046801	Burhan Barid, S.T., M.T.	Permai K 10 Ambarketawang	1970-09-07 00:00:00.000	Tidak Menikah	1996/1997	Teknik Sipil
507077302	Edi Hartono, S.T., M.T.	Jumus 02/08 Ngluwar Magelang Jawa	1973-07-07 00:00:00.000	Menikah	1999/2000	Teknik Sipil
507077302	Edi Hartono, S.T., M.T.	Jumus 02/08 Ngluwar Magelang Jawa	1973-07-07 00:00:00.000	Menikah	2001/2002	Teknik Sipil
507077302	Edi Hartono, S.T., M.T.	Jumus 02/08 Ngluwar Magelang Jawa	1973-07-07 00:00:00.000	Menikah	2003/2004	Teknik Sipil
507077302	Edi Hartono, S.T., M.T.	Jumus 02/08 Ngluwar Magelang Jawa	1973-07-07 00:00:00.000	Menikah	2008/2009	Teknik Sipil
507077302	Edi Hartono, S.T., M.T.	Jumus 02/08 Ngluwar Magelang Jawa	1973-07-07 00:00:00.000	Menikah	2012/2013	Teknik Sipil
507077302	Edi Hartono, S.T., M.T.	Jumus 02/08 Ngluwar Magelang Jawa	1973-07-07 00:00:00.000	Menikah	2013/2014	Teknik Sipil
507077302	Edi Hartono, S.T., M.T.	Jumus 02/08 Ngluwar Magelang Jawa	1973-07-07 00:00:00.000	Menikah	2015/2016	Teknik Sipil
10125702	Ir. Fadillawaty S, M.T.	Jl. Palagan Tentara Pelajar, Sedan No.80A RT.06 RW.34 Ngaglik,	1957-12-10 00:00:00.000	Menikah	1990/1991	Teknik Sipil

Gambar 4.12 Sebelum Pengolahan Data

Dari gambar diatas *dataset* masih harus diolah lagi menjadi data informasi yang bisa digunakan nantinya dipengujian. Seperti kotak hitam pada gambar 4.12 yaitu atribut nama, banyak duplikat data yang seharusnya hanya ada satu data saja tapi memang setelah diteliti ternyata dikarenakan tahun lama kerja yang membuat atribut nama banyak menduplikasi. Maka dari itu peneliti perlu menghapus atribut nama ganda dan peneliti hanya mengambil tahun pertama dosen masuk saja begitu juga atribut status, lama kerja dan alamat . Data sampel yang sudah diolah bisa dilihat pada gambar 4.13.

	A	B	C	D	E	F	G
	NIDN	Nama	Usia	Status	Lama Kerja	Work unit name	Jarak
2	514087501	Dr. Eng. Agus Setyo M, M.Eng.Sc.	41	Tidak Menikah	26 tahun	Teknik Sipil	dekat
3	20076502	Ir. Anita Widianti, MT.	51	Tidak Menikah	23 tahun	Teknik Sipil	jauh
4	514046601	Ir. Asat Pujiyanto, M.T.	51	Tidak Menikah	23 tahun	Teknik Sipil	dekat
5	525046801	Burhan Barid, S.T., M.T.	46	Tidak Menikah	20 tahun	Teknik Sipil	dekat
6	507077301	Edi Hartono, S.T., M.T.	43	Menikah	26 tahun	Teknik Sipil	jauh
7	10125702	Ir. Fadillawaty S, M.T.	59	Menikah	26 tahun	Teknik Sipil	jauh
8	25115502	Ir. M. Riang Endarto BS, M.S.	61	Menikah	36 tahun	Teknik Sipil	dekat
9	518025501	Ir. Mandiyo Priyo, M.T.	61	Tidak Menikah	22 tahun	Teknik Sipil	dekat
10	523125401	Ir. Sentot Hardwiyono, M.T., Ph.D.	62	Menikah	21 tahun	Teknik Sipil	dekat
11	524057201	Jazaul Ikhsan, S.T., M.T., Ph.D.	44	Tidak Menikah	18 tahun	Teknik Sipil	jauh
12	515026701	M. Heri Zulfiar, S.T., M.T.	49	Tidak Menikah	19 tahun	Teknik Sipil	dekat
13	515047801	Sri Atmaja PJNR, S.T., M.Sc.Eng., Ph.D.	38	Tidak Menikah	16 tahun	Teknik Sipil	dekat
14	501097202	Surya Budi Lesmana, S.T., M.T.	44	Tidak Menikah	16 tahun	Teknik Sipil	dekat
15	38126301	Ir. Wahyu Widodo, M.T.	53	Menikah	25 tahun	Teknik Sipil	dekat
16	522087401	Willis Diana, S.T., M.T.	42	Menikah	17 tahun	Teknik Sipil	jauh
17	529086601	Ir. Agus Jamal, M.Eng.	50	Menikah	21 tahun	Teknik Elektro	jauh
18	522126701	Ir. H. M. Fathul Qodir Ar.	49	Menikah	23 tahun	Teknik Elektro	dekat
19	502026801	Dr. Ir. Dwijoko Purbohadi, M.T.	48	Menikah	21 tahun	Teknologi Informasi	dekat
20	511116901	Haris Setvawan, S.T., M.Eng.	47	Tidak Menikah	16 tahun	Teknologi Informasi	dekat

Gambar 4.13 Sesudah Pengolahan Data

Dari gambar diatas terlihat data sudah jauh berbeda dengan gambar 4.12 pada atribut nama tidak ada lagi duplikasi data sementara pada atribut usia, alamat dan lama kerja terjadi perubahan. Atribut usia yang awalnya memakai tanggal lahir dirubah menjadi angka sedangkan atribut lama kerja yang awalnya (lihat gambar 4.13) diganti angka sama seperti atribut usia namun perbedaanya jika atribut usia hanya berupa angka atribut lama kerja memakai tahun dibelakang angka.

Perubahan mencolok terjadi pada atribut alamat, pada atribut alamat peneliti merubah alamat menjadi jarak sementara jauh dan dekat pada variabelnya. Bukan tanpa alasan peneliti merubah atribut alamat menjadi jarak dan jauh atau dekat pada variabelnya, jika tetap memakai alamat saat dipengujian RapidMiner akan menghitung satu persatu karena variabelnya yang banyak dan ini sangat mempengaruhi atribut yang lain. Peneliti tidak sembarangan menentukan varibael jauh dan dekat, untuk menentukan jauh dan dekat peneliti menggunakan *google maps* (kilometer) dimana alamat dosen dihubungkan dengan alamat Universitas Muhammadiyah Yogyakarta. Untuk lebih jelasnya silahkan lihat gambar 4.14.



Gambar 4.14 Menentukan Jarak

Pada gambar diatas untuk menentukan jauh dan dekat peneliti melihat jarak dan menit jika jarak hanya 1 sampai 10 kilometer berarti jarak dekat dan jarak jauh 11 kilometer lebih, berikut keterangan dalam tabel 4.1.

Tabel 4.1 Menentukan Jarak

No	Jarak	Nilai
1	1 - 10 km	Dekat
2	11 km lebih	Jauh

4.3. Pengujian Dengan WEKA

Setelah data sampel telah selesai diolah tahap selanjutnya adalah Pengujian menggunakan WEKA.



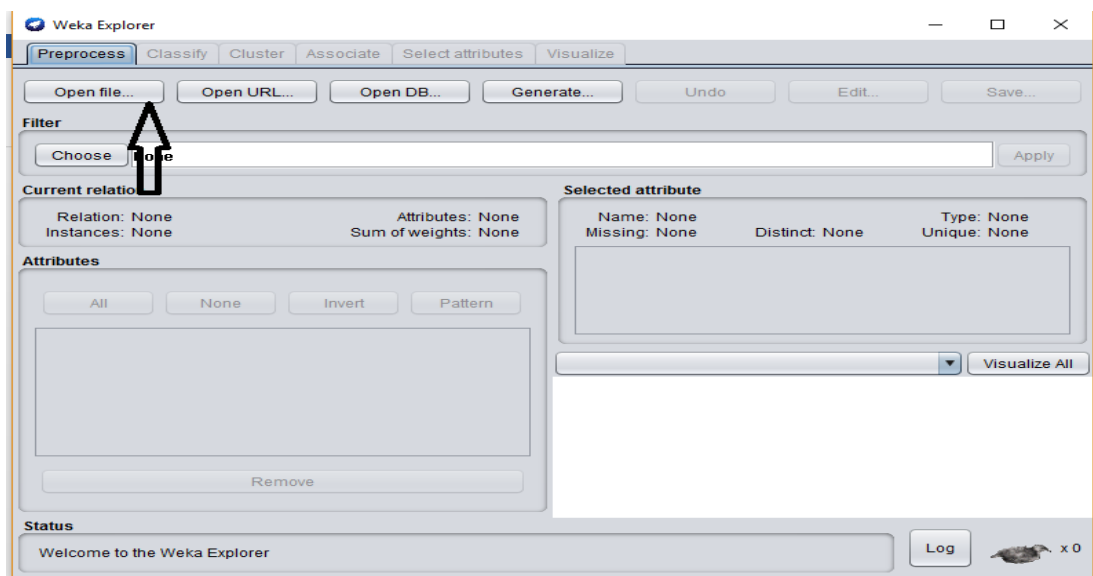
Gambar 4.15 Antar Muka Awal WEKA

Gambar 4.15 merupakan antar muka awal dari WEKA, untuk masuk ke antar muka selanjutnya yaitu antar muka WEKA dapat dilihat pada gambar 4.16. Pada gambar 4.16 merupakan antar muka WEKA versi 3.8 pada antar muka ini terdapat menu, gambar serta *tools*. Dan untuk memulai pengujian pilih *explorer* pada antar muka kanan atas seperti pada gambar 4.16.



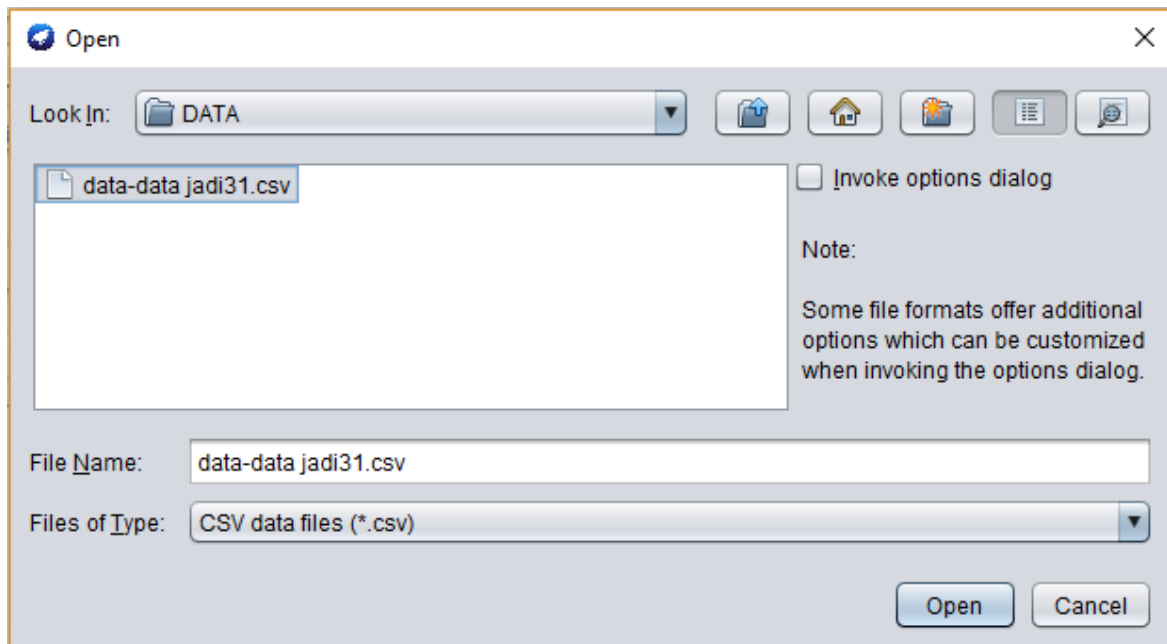
Gambar 4.16 Antar Muka WEKA

Langkah selanjutnya memilih *file* pada antar muka praproses WEKA seperti pada gambar 4.17.



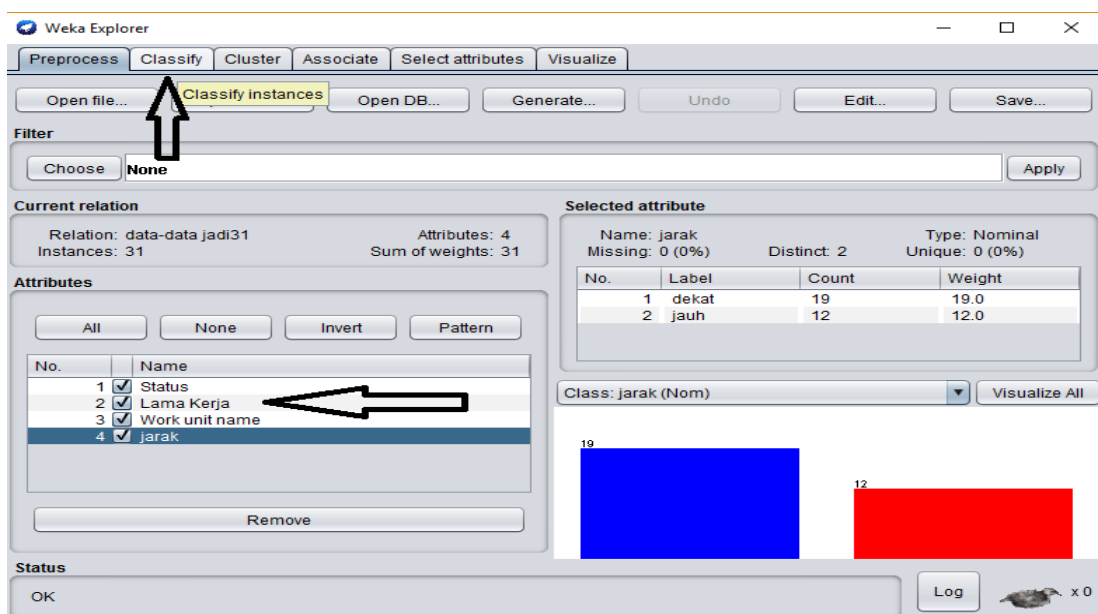
Gambar 4.17 Antar Muka Praproses

Langkah selanjutnya memilih data tersebut disimpan, untuk data sampel harus dirubah terlebih dahulu kedalam CSV (*Comma Separated Values*) karena WEKA tidak mendukung tipe *file excel*. Untuk lebih jelasnya dapat dilihat pada gambar 4.18.



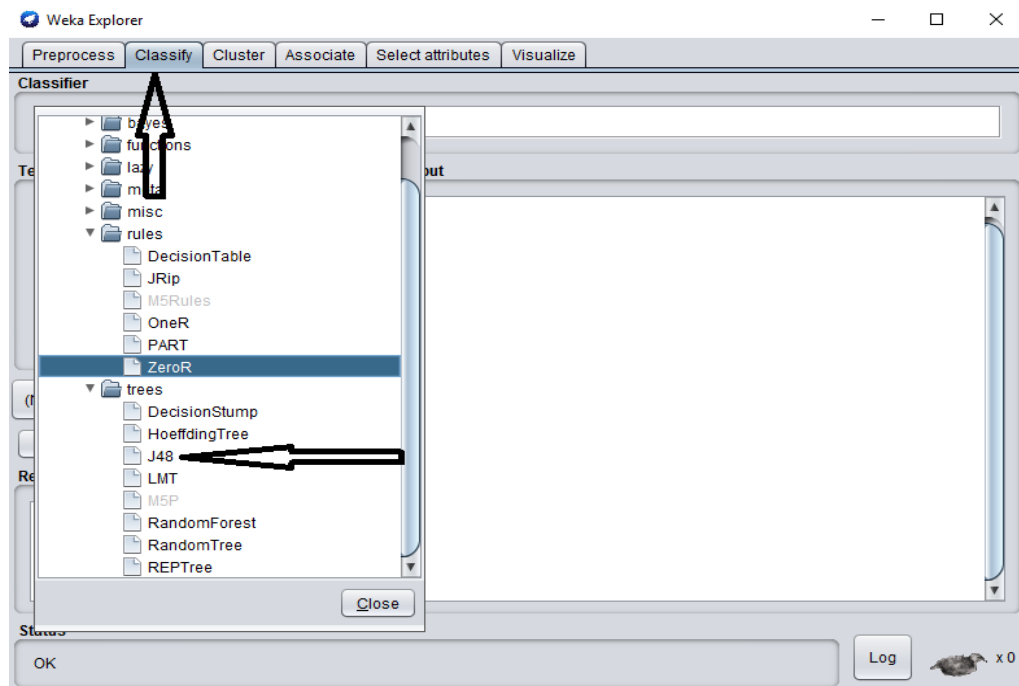
Gambar 4.18 Memilih *File*

Langkah selanjutnya yaitu centang atribut agar data bisa diuji seperti pada gambar 4.19 dan langkah selanjutnya pilih klasifikasi untuk memilih klasifikasi yang diinginkan seperti pada gambar 4.20.



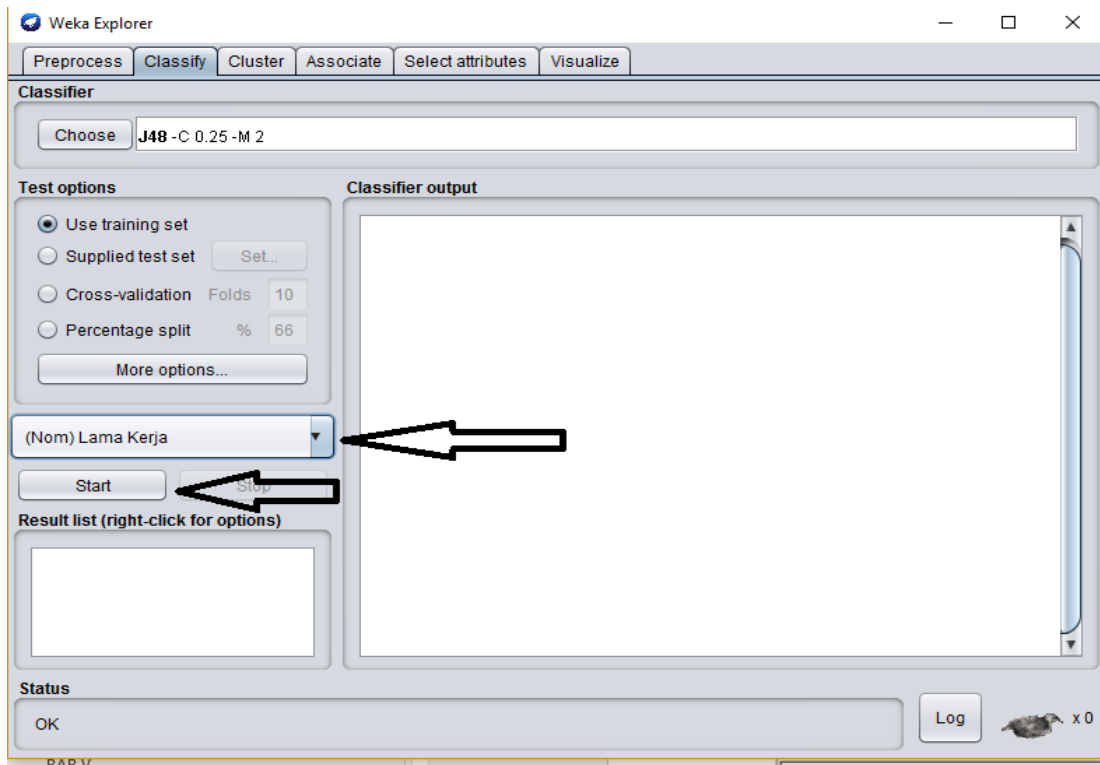
Gambar 4.19 Antar Muka Praproses

Pada antar muka klasifikasi terdapat banyak metode dari klasifikasi, karena peneliti menggunakan C.45 maka peneliti memilih C.45. untuk WEKA C.45 disebut J.48 untuk lebih jelasnya dapat dilihat pada gambar 4.20.



Gambar 4.20 Memilih Metode

Langkah selanjutnya setelah memilih metode klasifikasi yaitu memilih atribut target, dimana peneliti menggunakan lama kerja sebagai atribut target. Penjelasan dituangkan dalam gambar 4.21.



Gambar 4.21 Memilih Atribut Target

Setelah atribut diganti tahap selanjutnya yaitu menguji data sampel, dimana hasil akurasi dari klasifikasi ini 87.0968 % dan hasil bisa dilihat pada gambar 4.22.

```

=== Summary ===
Correctly Classified Instances      27
Incorrectly Classified Instances    4
Kappa statistic                    0.69
Mean absolute error                 0.143
Root mean squared error             0.2674
Relative absolute error             45.4016 %
Root relative squared error         68.6367 %
Total Number of Instances          31

=== Detailed Accuracy By Class ===

```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0,955	0,333	0,875	0,955	0,913	0,674	0,861	0,906	golongan B
	0,500	0,000	1,000	0,500	0,667	0,682	0,926	0,696	golongan A
	0,800	0,038	0,800	0,800	0,800	0,762	0,942	0,707	golongan C
Weighted Avg.	0,871	0,243	0,879	0,871	0,863	0,689	0,883	0,847	

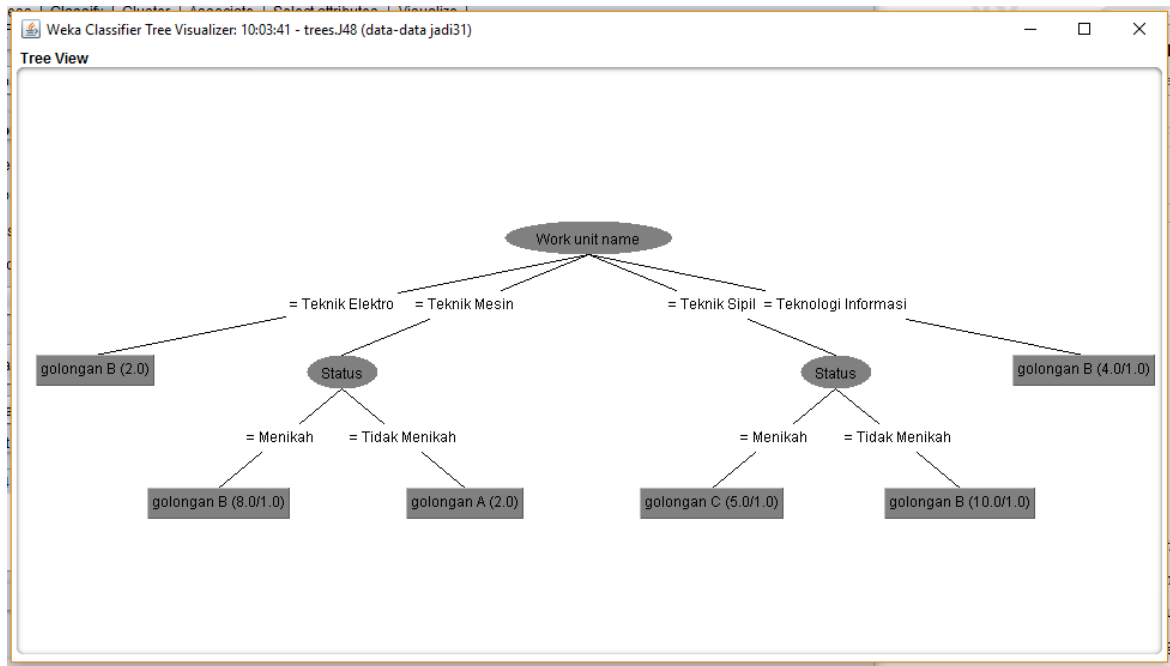
```

=== Confusion Matrix ===
 a  b  c  <-- classified as
21  0  1 | a = golongan B
 2  2  0 | b = golongan A
 1  0  4 | c = golongan C

```

Gambar 4.22 Hasil Pengujian WEKA

Hasil uji pohon keputusan yang dilakukan oleh algoritma C.45 (J48 di WEKA) di WEKA menunjukkan bahwa *work unit name* menjadi *root node* sementara status menjadi *child node*. Gambar pohon keputusan dapat dilihat seperti pada gambar 4.23.



Gambar 4.23 Hasil Pohon Keputusan Dari WEKA

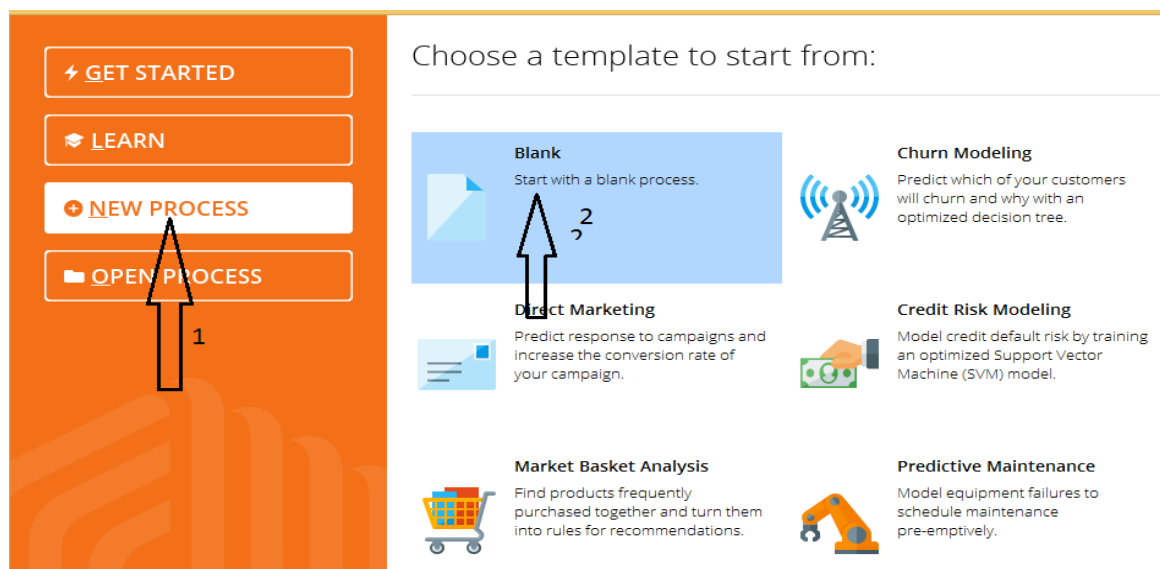
4.4. Pengujian Dengan RapidMiner

Setelah data sampel telah selesai diuji dengan WEKA tahap selanjutnya yaitu pengujian menggunakan RapidMiner versi 7.2 sebagai analisis, antar muka dari RapidMiner bisa dilihat pada gambar 4.24.



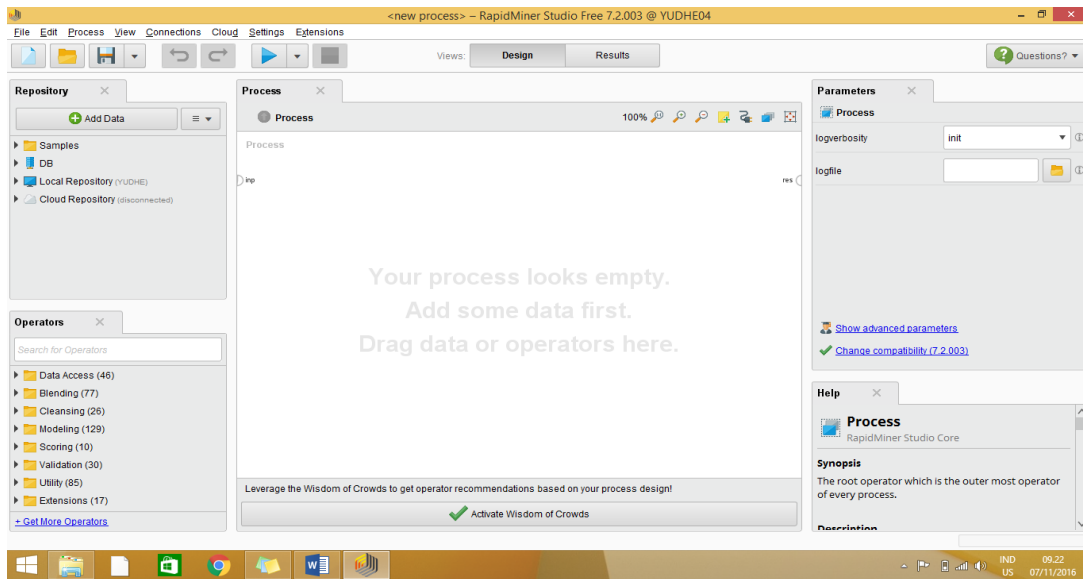
Gambar 4.24 Antar Muka RapidMiner Versi 7.2

RapidMiner versi 7.2 ini merupakan versi bukan terbaru yang dikembangkan oleh Rapidminer. Versi terbaru dari RapidMiner adalah versi 7.3 dan paling awal versi 5.3, untuk masuk ke antar muka awal sedikit lambat mungkin memakan waktu 1 sampai 3 menit. Setelah menunggu beberapa saat antar muka bisa dilihat pada gambar 4.25.



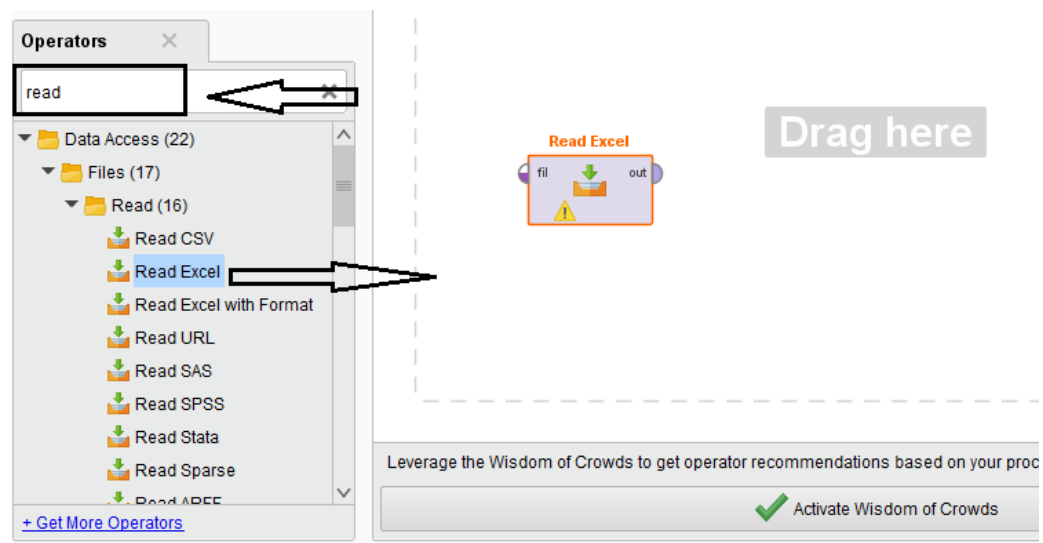
Gambar 4.25 Antar Muka Awal

Pada halaman awal pilih *new process* lalu pilih *blank* ini artinya akan melakukan proses analisa dari awal bukan yang sudah diproses sebelumnya tunggu beberapa saat hingga muncul tampilan seperti gambar 4.26.



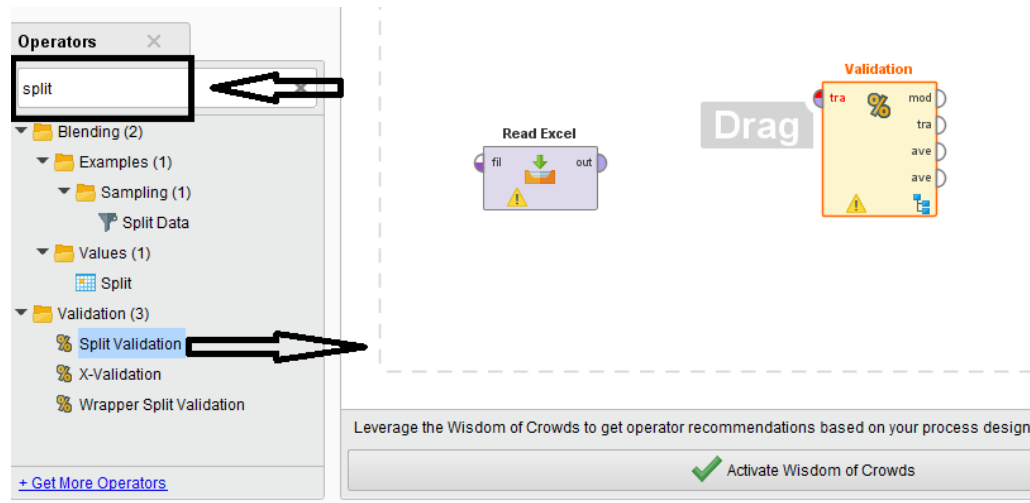
Gambar 4.26 Antar Muka Halaman Utama

Gambar 4.26 merupakan antar muka halaman utama dari RapidMiner yang terdapat menu – menu dari proses analisa dan prediksi. tahap selanjutnya menuju *panel operators* yang terletak dikiri bawah pada halaman utama, pada kotak pencarian *panel operators* ketik *read*. *Read* ini merupakan tipe *file* apa yang akan digunakan dalam penelitian dan dalam hal ini peneliti menggunakan *excel* sebagai tipe *file* selanjutnya *drop* dan *drag* ke *process*. Untuk lebih jelasnya silahkan lihat gambar 4.27.



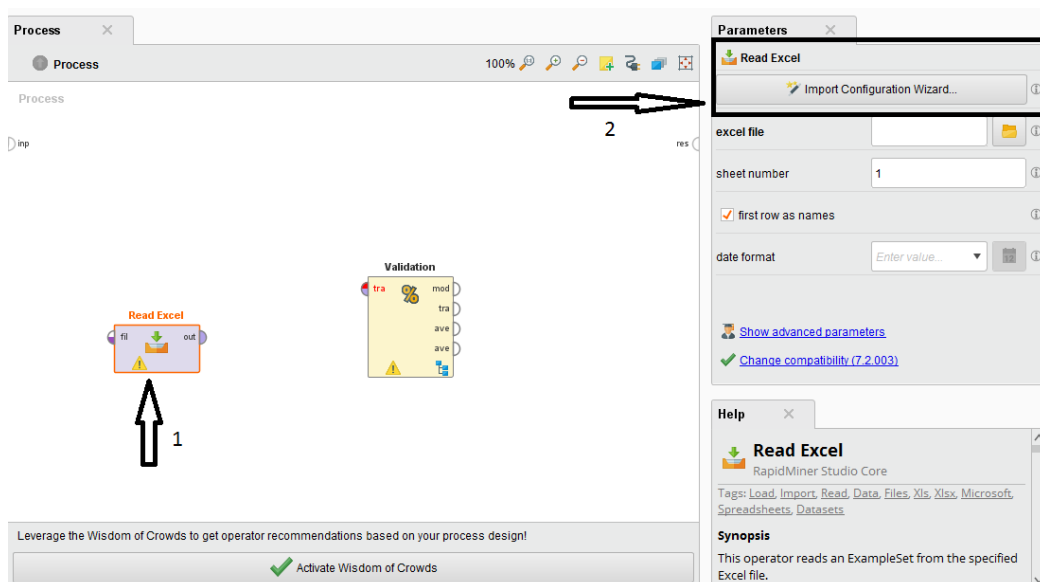
Gambar 4.27 Panel Operators

Setelah *read excel* di *drag* ke *process* ketik *split validation* dan langkahnya sama seperti *read excel* tadi, *split validation* merupakan operator yang mempunyai 2 subproses yaitu *training* dan *testing* . Coba lihat pada gambar 4.28.



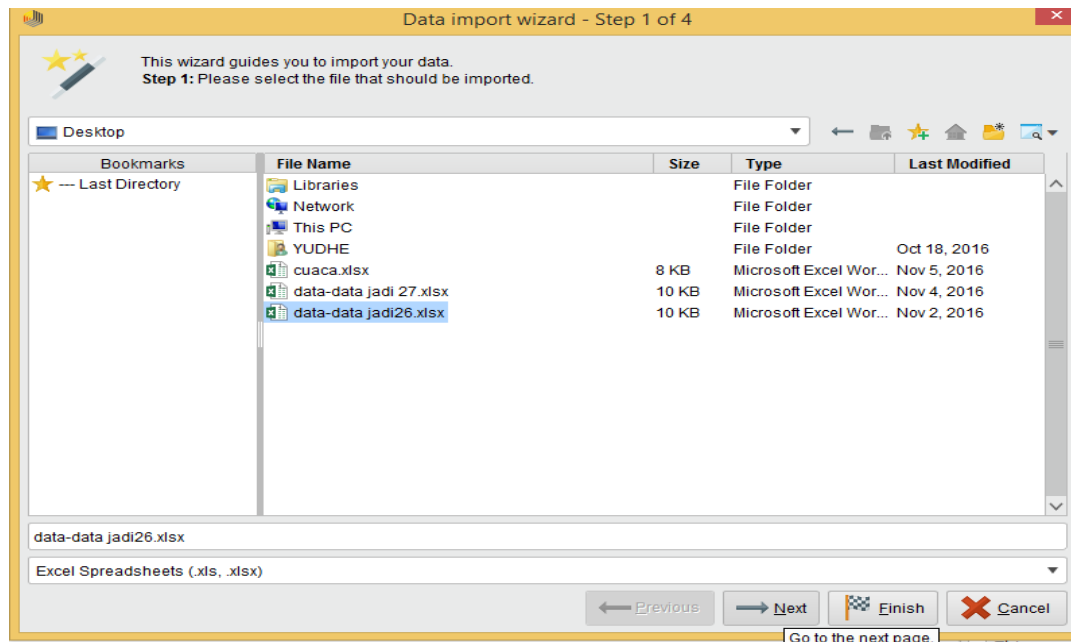
Gambar 4.28 Panel Operators

Setelah itu klik *read excel* pada *panel process* dan menuju ke *import configuration wizard* untuk mencari *file excel* seperti pada gambar 4.29.



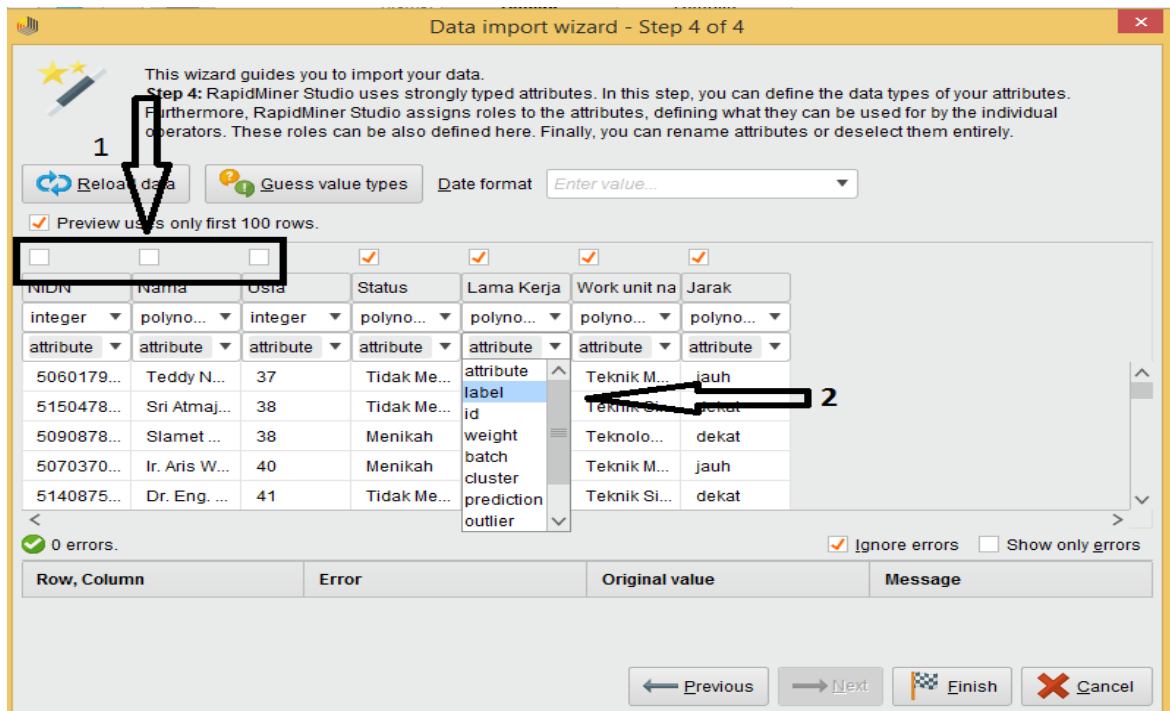
Gambar 4.29 Panel Process

Cari *file excel* yang sudah disimpan dan disiapkan seperti gambar 4.30 dan klik *next* dan abaikan 2 *next step* pada *data import wizard*.



Gambar 4.30 Pencarian Data

Di *step* yang terakhir peneliti menghilangkan centang pada atribut NIDN, nama dan usia karena ketiga atribut tersebut bukanlah bagian dari pengujian lalu di atribut lama kerja peneliti mengganti tipe atribut menjadi *label* hal ini karena lama kerja akan menjadi target objek nya klik *finish* dan jika kurang jelas lihat gambar 4.31.

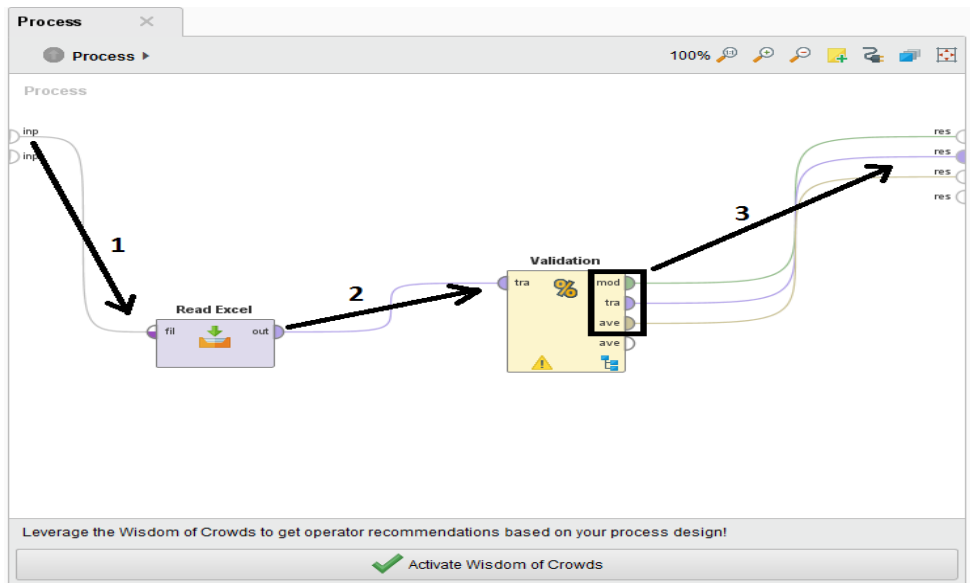


Gambar 4.31 Pengaturan Data *Import*

Setelah *import* data selesai tahap pengujian akan kembali ke *panel process* hubungkan *port inp (input)* ke *port fil (file)* yang ada pada *operator read excel*, *port out (output)* dihubungkan ke *port tra (training)* pada *split validation* dan tiga port yang ada di *split validation* dihubungkan ke *port res (result)*. Didalam setiap operator terdapat *port* yang ternyata mempunyai fungsinya masing – masing berikut penjelasan peneliti yang dituangkan dalam tabel 4.2.

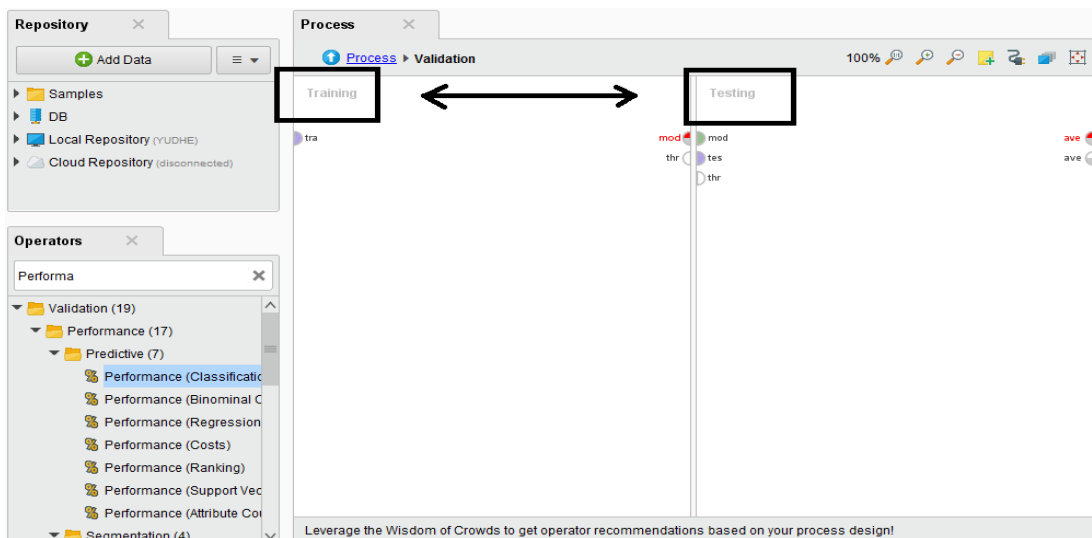
Tabel 4.2 Tabel Fungsi *Operator*

<i>Operator Read</i>		
No	Nama <i>Port</i>	Fungsi
1.	<i>Input Fil (File)</i>	Objek dari <i>file excel</i> yang mana nantinya akan dibaca pada operator lain.
2.	<i>Output Out (Output)</i>	Mengirimkan keluaran berupa <i>file</i> dalam bentuk tabel – tabel atau biasa disebut <i>ExampleSet</i> .
<i>Operator Validation</i>		
No	Nama <i>Port</i>	Fungsi
1.	<i>Input Tra (Training)</i>	Menerima <i>ExampleSet</i> dari <i>Output Operator Read</i> untuk dijadikan pelatihan data <i>training</i> .
2.	<i>Output Mod (Model)</i>	Mengambil <i>ExampleSet</i> yang sudah dilatih diinput <i>training</i> tadi dari subproses pengujian.
3.	<i>Output Tra (Training)</i>	Mengambil data <i>training</i> dari <i>input training</i> dan dilewatkan melalui <i>port</i> ini tanpa mengubah <i>output</i> .
4.	<i>Input Ave (Averagable)</i>	Mengembalikan <i>vector</i> kinerja yang telah diuji oleh subproses pengujian.



Gambar 4.32 Panel Process

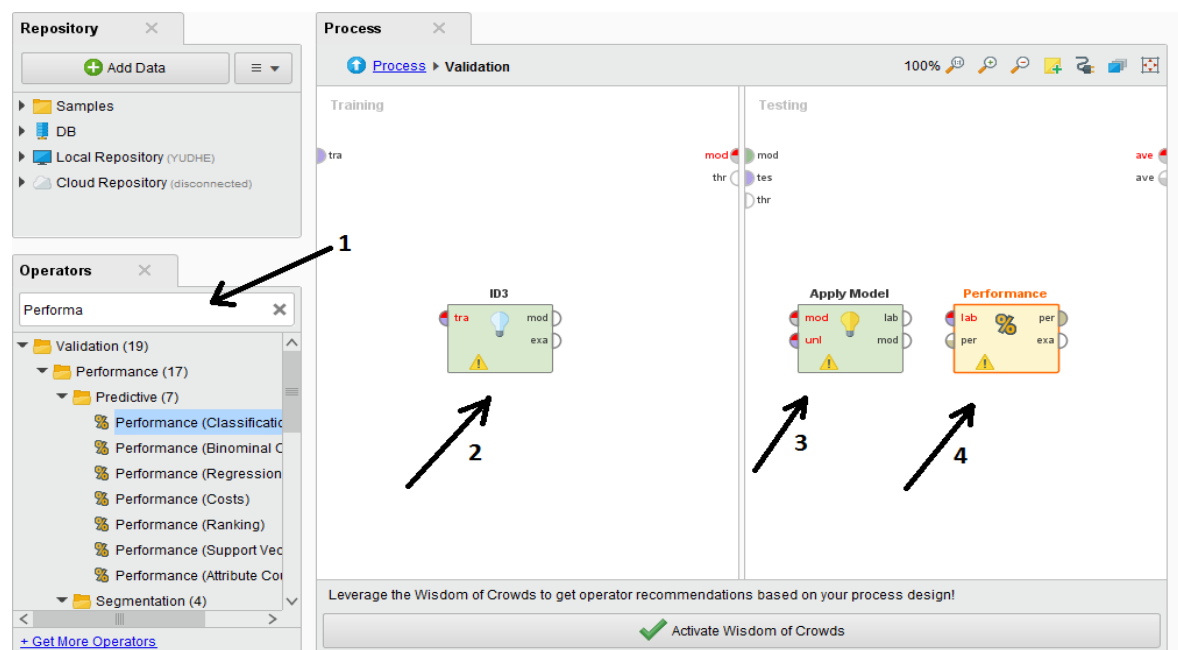
Selanjutnya *double* klik pada operator *validation* hingga muncul subproses *training* dan *testing* seperti gambar 4.33.



Gambar 4.33 Panel Process Validation

Pada gambar 4.33 merupakan antar muka subproses pengujian *training* dan *testing*. *ExampleSet* yang sudah dilatih sebelumnya akan dibangun dalam metode yang yang diperlukan dalam hal ini peneliti menggunakan algoritma C.45 namun algoritma C.45 tidak ada dalam RapidMiner jadi peneliti menggunakan ID3 sebagai penggantinya, perlu

diketahui algoritma C.45 merupakan pengembangan dari ID3 dan ID3 ini untuk membangun sebuah pohon keputusan atau *decision tree* pada subproses *testing* peneliti menggunakan *apply model* dan *performance*. untuk meletakkan ID3 pada subproses *training* caranya ketikkan ID3 pada pencarian *panel operators* lalu *drop and drag* begitu juga *apply model* dan *performance* namun *apply model* dan *performance* diletakan di subproses *testing* untuk lebih jelasnya silahkan lihat gambar 4.34.



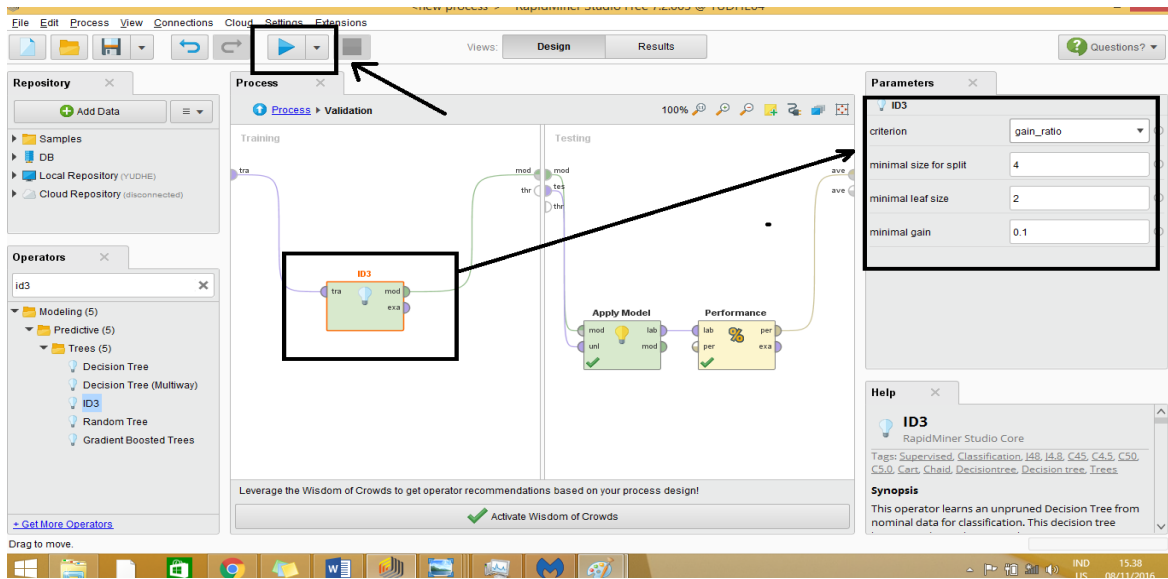
Gambar 4.34 Subproses *Training* dan *Testing*

Apply model fungsinya mempelajari informasi *ExampleSet* yang sudah dilatih dan biasanya untuk prediksi menggunakan model ini sedangkan *performance* ini digunakan untuk evaluasi statistik dari kinerja klasifikasi dan memberikan daftar nilai kriteria dari kinerja klasifikasi tersebut. Meletakkan algoritma ID3 pada subproses *training*, di *panel operators* ketik ID3 begitu juga dengan *apply model* dan *performance* setelah itu dihubungkan ke *port*. Dan dalam setiap operator terdapat *port* yang masing – masing fungsinya akan dijelaskan dalam tabel 4.3.

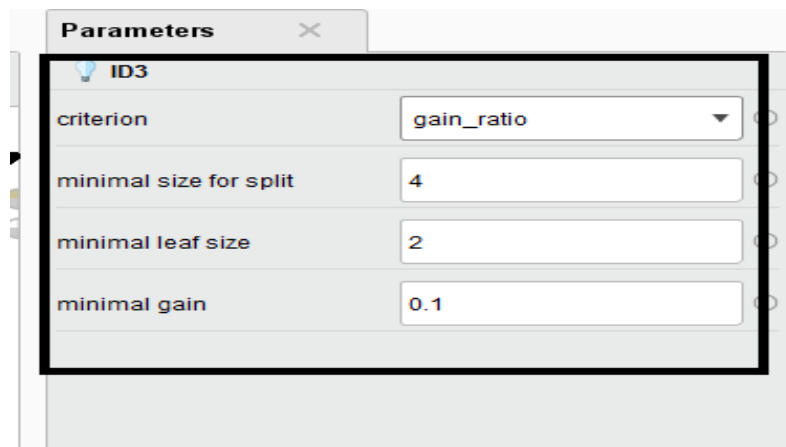
Tabel 4.3 Tabel Fungsi *Operator*

<i>Operator ID3</i>		
No	Nama <i>Port</i>	Fungsi
1.	<i>Input (training)</i>	Menerima <i>ExampleSet</i> dari <i>operator validation</i> .
2.	<i>Output model</i>	Mengirimkan <i>decision tree</i> ke <i>apply model</i> .
3.	<i>Output Exa</i>	mengirimkan <i>trainingset</i> dari <i>ExampleSet</i> yang sudah diolah di ID3 untuk diuji pada subproses <i>testing</i> .
<i>Operator Apply Model</i>		
No	Nama <i>Port</i>	Fungsi
1.	<i>Input model</i>	Menerima <i>Model</i> dari <i>Output Operator ID3</i> .
2.	<i>Input unlabelled</i>	Mengambil data <i>testingset</i> dan memastikan bahwa jumlah, jenis, urutan dan peran atribut dari data <i>testing</i> ini konsisten dengan <i>trainingset</i> .
3.	<i>output labelled</i>	mengirimkan semua olahan data yang diproses <i>apply model</i> lalu teruskan dioperator <i>performance</i> .
<i>Operator Performance</i>		
No	Nama <i>Port</i>	Fungsi
1	<i>Input labelled</i>	Menerima data dari <i>Output Operator Apply Model</i> .

Setelah *port – port* terhubung langkah selanjutnya klik operator ID3 untuk mengatur parameternya seperti tipe perhitungannya menggunakan tipe apa, peneliti menggunakan *Information_gain* untuk *criterion* untuk *minimal size for split* nya peneliti menggunakan pengaturan *default* yaitu 4 begitu juga *minimal leaf size* dan *minimal gain* seperti pada gambar 4.26 atau jika kurang jelas untuk parameternya lihat gambar 4.35.



Gambar 4.35 Subproses *Training* dan *Testing*



Gambar 4.36 Parameter ID3

Setelah langkah – langkah sudah dilakukan tahap terakhir yaitu memproses dengan klik proses seperti tanda panah pada seperti gambar 4.35 dan hasilnya seperti pada gambar 4.37.

PerformanceVector (Performance) | ExampleSet (Read Excel) | Tree (ID3)

Table View | Plot View

accuracy: 14.29%

	true 23 t...	true 21 t...	true 13 t...	true 20 t...	true 19 t...	true 18 t...	true 17 t...	true 26 t...	true 25 t...	true 36 t...	true 15 t...
pred. 23 ...	0	0	0	0	0	1	0	0	0	0	0
pred. 21 ...	1	0	0	0	0	0	0	0	0	0	0
pred. 13 ...	0	0	0	1	0	0	0	0	0	0	0
pred. 20 ...	0	0	0	0	0	0	0	0	0	0	0
pred. 19 ...	0	0	0	0	0	0	0	0	0	0	0
pred. 18 ...	0	0	0	0	0	0	0	0	0	0	0
pred. 17 ...	0	0	0	0	0	0	0	0	0	0	0
pred. 26 ...	0	0	0	0	0	0	0	0	0	0	0
pred. 25 ...	0	0	0	0	0	0	0	0	0	0	0
pred. 36 ...	0	0	0	0	0	0	0	0	0	0	0
pred. 15 ...	0	1	0	0	0	0	0	0	0	0	0
pred. 16 ...	0	0	0	0	1	0	1	0	0	0	0
pred. 22 ...	0	0	0	0	0	0	0	0	0	0	0
class rec...	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Gambar 4.37 Hasil Proses Analisa

Antar muka hasil dari proses analisa bisa dilihat pada gambar 4.37 diketahui akurasi dari *dataset* yang diuji menggunakan Algoritma ID3 adalah 14.29%. Jika tingkat akurasi menggunakan angka 1 sampai 100 persen ini artinya *vector* kinerja dari klasifikasi *dataset* belum maksimal, penyebab belum maksimalnya adalah banyaknya *atribut target* atau *class* dalam *dataset* tersebut disamping itu data yang digunakan untuk menguji hanya berjumlah 31 dan berbanding 13 *label*, dengan *label* sebanyak itu perhitungan *entropy* dan *gain* juga tidak maksimal seperti pada pada gambar 4.37.

Dengan hasil akurasi yang hanya 14,29% bukan berarti pengujian ini gagal hanya saja untuk menerapkan metode klasifikasi pada data sampel dengan *label* yang banyak ditambah dengan data yang sedikit ini tidak cocok. Tidak hanya itu banyak faktor yang mempengaruhi hasil akurasi pengujian (bukan penelitian penulis) misalnya dari faktor datanya yang tidak relevan seperti dalam menentukan kelulusan calon mahasiswa dengan target atribut lulus dan tidak lulus, tentu atribut yang digunakan adalah nilai UN (Ujian Nasional),usia, jurusan tujuan dan hasil ujian dari intitusi jika atribut yang digunakan hanya id, nama dan usia artinya data tidak relevan dengan atribut target.

Untuk mengatasi ini peneliti memangkas dan mengolah kembali data sampel dengan merubah atribut target atau label menjadi lebih sedikit dengan cara menggabungkan lama kerja menjadi 3 label. Jika rata – rata lama kerja dosen Fakultas Teknik UMY adalah 13 sampai 30 tahun maka lama kerja 1 sampai 15 tahun masuk kategori golongan A, 16 sampai 25 tahun golongan B, 26 sampai 40 tahun golongan C. Data sampel bisa dilihat pada gambar 4.38 sedangkan untuk daftar tabel golongan bisa dilihat pada tabel 4.4.

Tabel 4.4 Daftar Golongan Lama Kerja

No	Lama Kerja	Golongan
1	1-15 tahun	A
2	16-25 tahun	B
3	26-40 tahun	C

	A	B	C	D
1	Status	Lama Kerja	Work unit name	Jarak
2	Menikah	golongan B	Teknik Elektro	dekat
3	Menikah	golongan B	Teknik Elektro	jauh
4	Menikah	golongan A	Teknik Mesin	dekat
5	Menikah	golongan B	Teknik Mesin	dekat
6	Menikah	golongan B	Teknik Mesin	dekat
7	Menikah	golongan B	Teknik Mesin	dekat
8	Menikah	golongan B	Teknik Mesin	dekat
9	Menikah	golongan B	Teknik Mesin	jauh
10	Menikah	golongan B	Teknik Mesin	jauh
11	Menikah	golongan B	Teknik Mesin	jauh
12	Tidak Menikah	golongan A	Teknik Mesin	jauh
13	Tidak Menikah	golongan A	Teknik Mesin	jauh
14	Menikah	golongan B	Teknik Sipil	dekat
15	Menikah	golongan C	Teknik Sipil	dekat
16	Menikah	golongan C	Teknik Sipil	dekat
17	Tidak Menikah	golongan B	Teknik Sipil	dekat
18	Tidak Menikah	golongan B	Teknik Sipil	dekat
19	Tidak Menikah	golongan B	Teknik Sipil	dekat
20	Tidak Menikah	golongan B	Teknik Sipil	dekat
21	Tidak Menikah	golongan B	Teknik Sipil	dekat
22	Tidak Menikah	golongan B	Teknik Sipil	dekat
23	Tidak Menikah	golongan C	Teknik Sipil	dekat

Gambar 4.38 Data Sampel Baru

Setelah melakukan perubahan pada atribut label hasil akurasi yang didapatkan adalah 88.67 % ini menunjukkan bahwa klasifikasi sudah baik seperti pada gambar 4.30.

Table View
 Plot View

accuracy: 86.67%

	true golongan B	true golongan A	true golongan C	class precision
pred. golongan B	10	1	0	90.91%
pred. golongan A	0	1	0	100.00%
pred. golongan C	1	0	2	66.67%
class recall	90.91%	50.00%	100.00%	

Gambar 4.39 Hasil Akurasi

Hasil statistik menunjukkan bahwa lama kerja dengan golongan B adalah 22 dosen, lama kerja dengan golongan C adalah 5 sedangkan lama kerja dengan golongan A ada 4. Dari status menikah ada 18 dosen dan tidak menikah 13 dosen. Dari *work unit name* dari Teknik Sipil adalah terbanyak dengan 15 dosen diikuti dengan Teknik Mesin 10, Teknologi Informasi 4 dan dari Teknik Elektro sebanyak 2, dari jarak dekat ada 19 dan jauh 12. Hasil statistik bisa dilihat pada gambar 4.40.



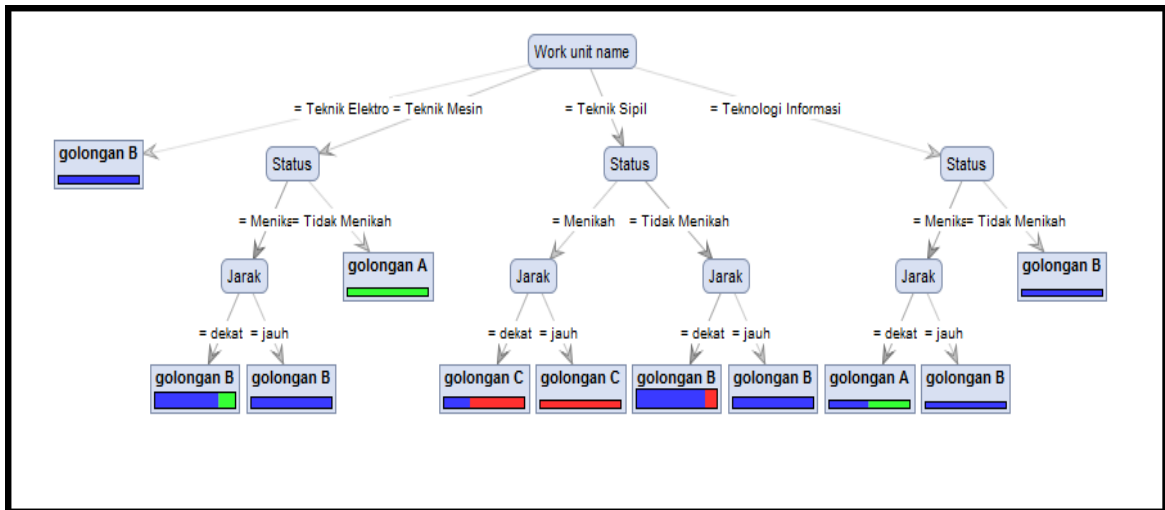
Gambar 4.40 Hasil Statistik

Hasil klasifikasi *decision tree* dari *ExampleSet* dimana yang menjadi *root node* adalah *work unit name* yang terbagi atas 4 cabang nilai atribut yaitu Teknik Elektro, Teknik Mesin, Teknik Sipil dan Teknologi Informasi. Bisa disimpulkan bahwa dosen dari Teknik Elektro masuk dalam lama kerja golongan B dengan warna biru mutlak yang berarti dosen dari Teknik Elektro telah mengajar di fakultas tersebut berkisar antara 16 sampai 25 tahun. Arti warna dari atribut menjelaskan bahwa tidak adanya golongan lama kerja yang lain.

Dari dosen teknik mesin yang berstatus menikah dan berjarak jauh masuk kategori lama kerja golongan B dengan biru mutlak sedangkan yang berstatus menikah dan berjarak dekat masuk golongan B, dari status tidak menikah masuk golongan A. kombinasi warna dari nilai atribut target menjelaskan dari hasil nilai yang ditampilkan terdapat golongan lama kerja yang lain.

Dari dosen teknik sipil yang berstatus menikah dan berjarak dekat lama kerjanya golongan C dengan kombinasi warna, biru dan merah. Dari yang berjarak jauh lama kerjanya masuk golongan C dengan warna merah mutlak sedangkan yang berstatus tidak menikah dan berjarak dekat lama kerjanya masuk golongan B dengan kombinasi warna, biru dan merah dan yang berjarak jauh lama kerjanya masuk golongan B dengan warna biru mutlak.

Dari dosen Teknologi Informasi yang berstatus menikah dan berjarak dekat lama kerjanya masuk golongan A dengan 2 kombinasi warna, biru dan hijau dan yang berjarak jauh lama kerjanya masuk golongan B dengan warna biru mutlak. Sedangkan yang berstatus tidak menikah lama kerjanya masuk golongan B dengan warna biru mutlak. Lihat gambar 4.41.



Gambar 4.41 Decision Tree Hasil Pengujian

4.5. Perbandingan Hasil Pengujian

Perbandingan hasil pengujian ini digunakan untuk membandingkan antara hasil pengujian yang dilakukan di WEKA menggunakan algoritma J48 dan RapidMiner menggunakan algoritma C.45, perbandingan antaran dua *software* ini meliputi tingkat akurasi, *class recall* dan *class precision*. Perbandingan hasil pengujian peneliti tuangkan kedalam tabel 4.5.

Tabel 4.5 Perbandingan Hasil Pengujian

Akurasi		
Algoritma C.45	87.0968%	
Algoritma ID3	86.67%	
<i>Class Precision</i> Menggunakan C.45	<i>Class Precision</i> Menggunakan ID3	
Golongan A	100 %	100 %
Golongan B	87.50 %	90.91 %
Golongan C	87.90 %	66.67 %

Class Recall Menggunakan C.45		Class Recall Menggunakan ID3
Golongan A	50 %	50 %
Golongan B	95.50 %	90.91 %
Golongan C	80 %	100 %

4.6. Perhitungan *Entropy* dan *Gain*

Dalam pembentukan sebuah pohon keputusan atau *decision tree* diperlukan perhitungan *entropy* dan *gain* dimana *gain* yang paling tertinggi akan menjadi *root node*. Karena peneliti menggunakan *software* sebagai analisa jadi peneliti perlu membuktikan mengapa *work unit name* menjadi *root node* dari *decisio tree*, perhitungan *entropy* ini menggunakan persamaan I sebagai berikut.

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

Keterangan:

S : himpunan kasus

n : jumlah partisi S

P_i : proporsi dari S_i terhadap S

Sementara itu, perhitungan nilai *gain* menggunakan persamaan II sebagai berikut.

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

Keterangan:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

$|S_i|$: jumlah kasus pada partisi ke-i

$|S|$: jumlah kasus dalam S

Berikut ini adalah penjelasan lebih terperinci mengenai tiap – tiap langkah dalam pembentukan *decision tree* dengan menggunakan algoritma C.45 yang dituangkan dalam tabel 4.3.

Tabel 4.5 Hasil Perhitungan Node 1

NODE			jumlah kasus	Gol A	Gol B	Gol C	Entropy	Gain
	total		31	4	22	5	1,15686934	
	WUN							0,5630108
		TS	15	0	10	5	0	
		TI	4	1	3	0	0	
		TE	2	0	2	0	0	
		TM	10	3	7	0	0	
	Status							0,0302392
		Menikah	18	2	12	4	1,22439445	
		Tdk Menikah	13	2	10	1	0,99126426	
	Jarak							0,0061495
		dekat	12	2	8	2	1,25162916	
		jauh	19	2	14	3	1,08698770	

Berdasarkan tabel 4.4 *entropy* total lama kerja adalah 1,15686934 langkah selanjutnya yaitu menghitung *entropy* dari tiap – tiap atribut untuk menentukan *node* dari sebuah pohon keputusan. Pada tabel 4.1 diketahui *entropy* dari variabel *work unit name* adalah *null* penyebabnya karena didalam perhitungan *entropy* jika salah satu nilai variabel adalah *null* maka hasilnya akan tetap *null*. Dari hasil perhitungan diatas diketahui *gain* yang paling besar adalah WUN (*work unit name*) dimana *atribut work unit name* mendapat hasil

0,5630108, hasil *gain* dari *work unit name* didapatkan dari hasil normalisasi dengan menggunakan *gain ratio*. sedangkan *gain* dari status 0,0302392 dan jarak mendapatkan hasil *gain* 0,0061495. Untuk lebih jelasnya peneliti telah menghitung dari *entropy* dan *gain* di lembar lampiran.