Nama: Dewi Rahmawati

NIM: L200170188

Kelas: F

Modul 8

#### PRAKTIKUM DATA WAREHOUSING DATA MINING

# **Tugas**

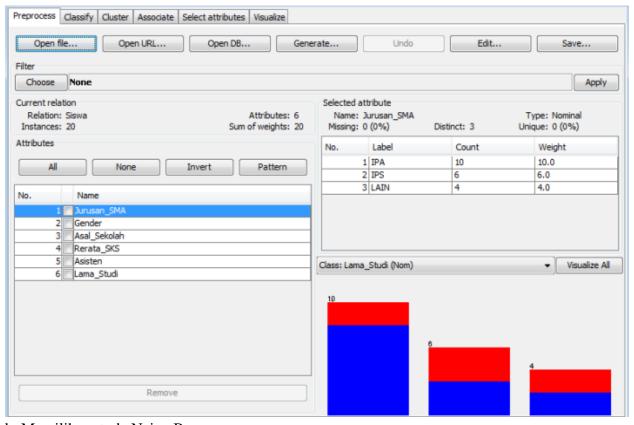
1. Berdasarkan tabel berikut, buatlah file dalam format Excel(.xls) dan format ARFF(.arff)! Data ini akan digunakan sebagai data testing

	F11	<b>→</b> (n)	f <sub>x</sub>			
	Α	В	С	D	Е	F
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	
2	LAIN	WANITA	SURAKARTA	18	TIDAK	
3	IPA	PRIA	SURAKARTA	19	YA	
4	LAIN	PRIA	SURAKARTA	19	TIDAK	
5	IPS	PRIA	LUAR	17	TIDAK	
6	LAIN	WANITA	SURAKARTA	17	TIDAK	
7	IPA	WANITA	LUAR	18	YA	
8	IPA	PRIA	SURAKARTA	18	TIDAK	
9	IPA	PRIA	SURAKARTA	19	TIDAK	
10	IPS	PRIA	LUAR	18	TIDAK	
11	LAIN	WANITA	SURAKARTA	18	TIDAK	
12						

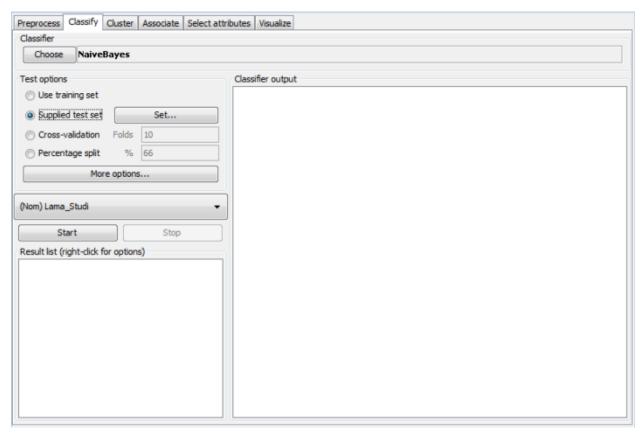
```
File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?
 data_testing arff 🖸
       Grelation Siswa
     Grelation Siswa
Gattribute Jurusan SMA (IPA, IPS, LAIN)
Gattribute Gender (FRIA, WANITA)
Gattribute Asal Sekolah (SURAKARTA, LUAR)
Gattribute Rerata_SKS Real
Gattribute Asisten (YA, TIDAK)
Gattribute Lama_Studi (TEPAT, TERLAMBAT)
                    WANITA, SURAKARTA, 18, TIDAK,
PRIA, SURAKARTA, 19, YA,
  10 LAIN,
11 IPA,
                     PRIA, SURAKARTA,
PRIA, LUAR,
WANITA, SURAKARTA,
                                                       17, TIDAK,
17, TIDAK,
      IPS,
LAIN,
                                 LUAR, 18, YA,
SURAKARTA, 18, TIDAK,
SURAKARTA, 19, TIDAK,
LUAR, 18, TIDAK,
                     WANITA, LUAR,
PRIA, SURAKARTA,
        IPA,
       IPA,
       IPA,
        IPS, PRIA, LUAR, 18, TIDAK,
LAIN, WANITA, SURAKARTA, 18, TIDAK,
```

2. Gunakan file ARFF yang dikerjakan pada Tugas nomor 1 dalam Modul 7 sebagai data

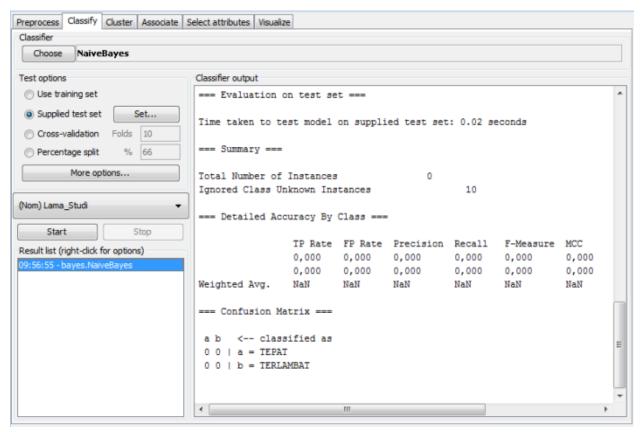
training. Lakukan prediksi terhadap data testing (ARFF) di atas menggunakan WEKA! a. Membuka file Training



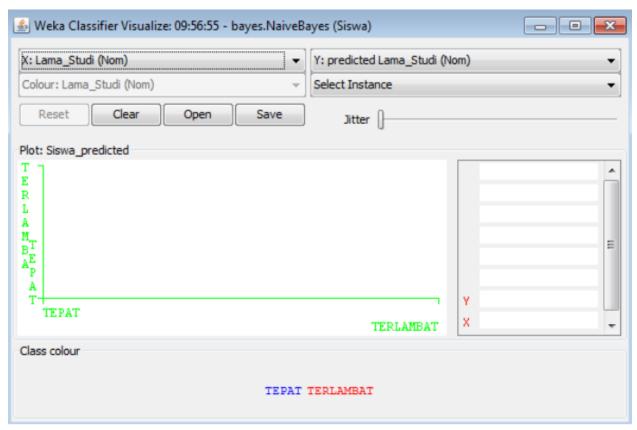
b. Memilih metode Naive Bayes



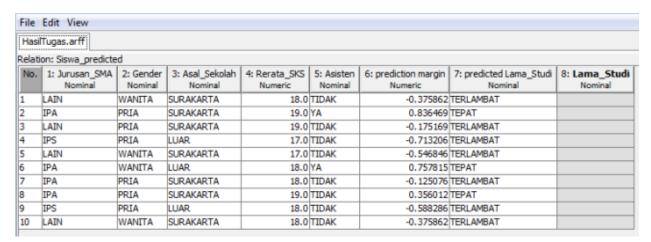
c. Membuka file Testing lalu klik kanan pilih Visualize classifier errors



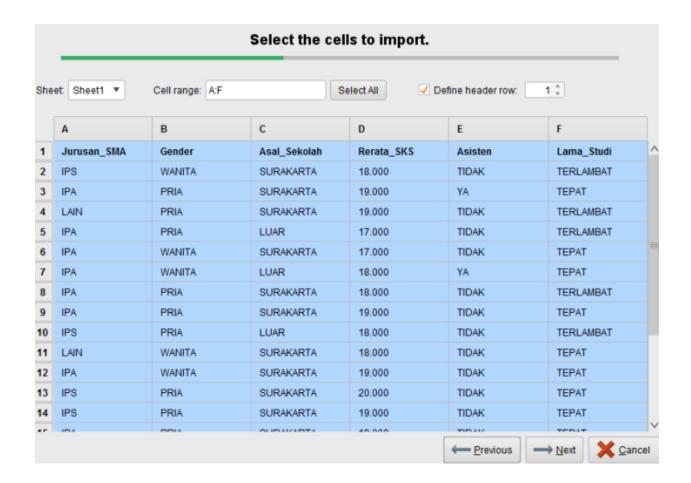
d. Menyimpan file hasil

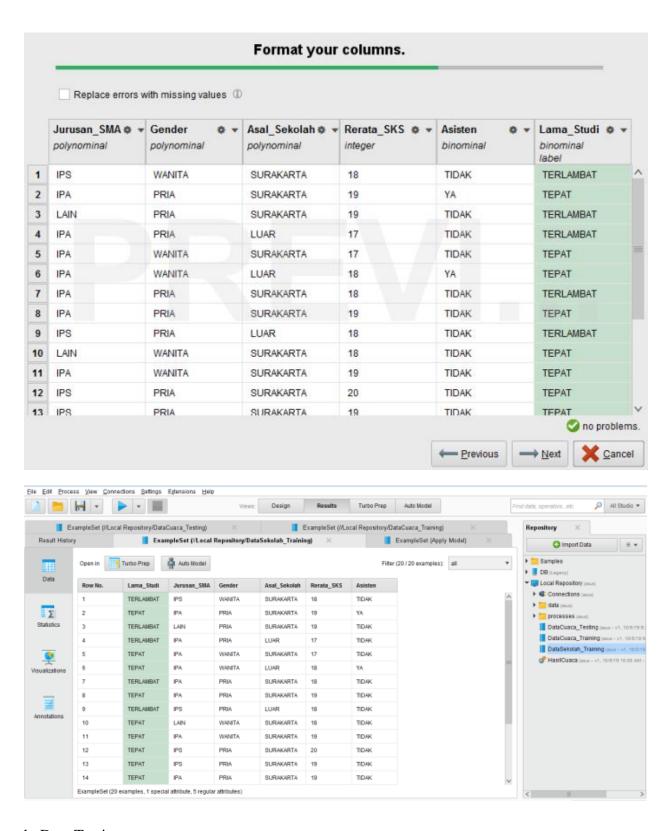


e. Hasil Prediksi

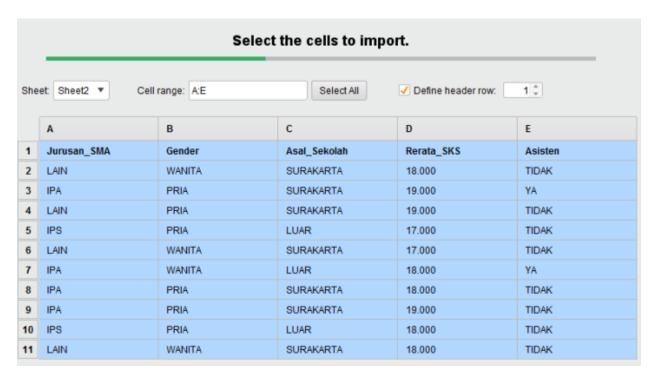


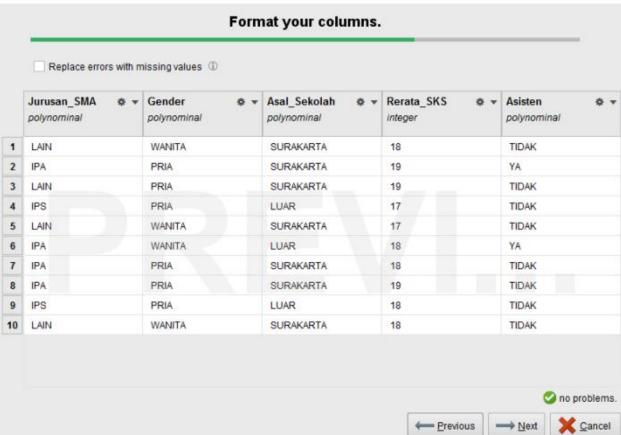
3. Gunakan file Excel yang dikerjakan pada Tugas nomor 1 dalam Modul 6 sebagai data training. Laukan prediksi terharap data testing (Excel) di atas menggunakan RapidMiner! a. Data Training

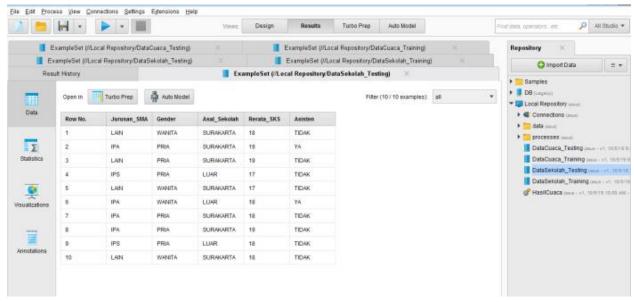




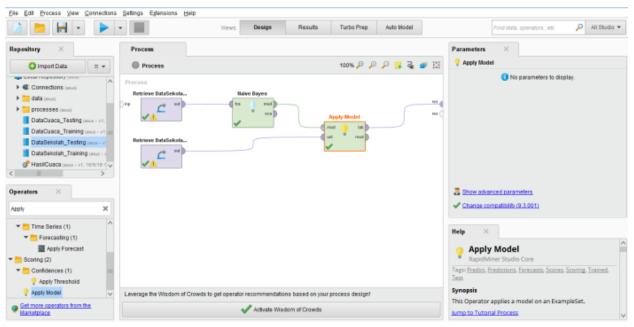
b. Data Testing



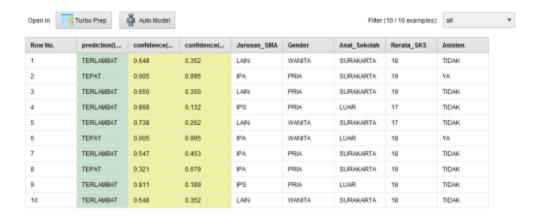


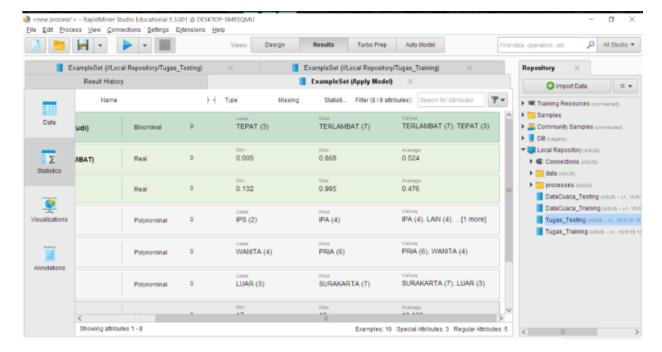


#### c. Implementasi Naive Bayes

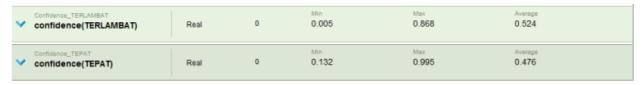


d. Hasil





4. Dari hasil percobaan Tugas nomor 3 di atas, berapakah nilai rerata confidence untuk atribut Lama\_Studi dengan nilai TEPAT? Berapakah nilai rerata confidence untuk atribut Lama\_Studi dengan nilai TERLAMBAT?



Jadi nilai rerata confidence untuk atribut Lama\_Studi dengan nilai TEPAT adalah **0,524** dan nilai rerata confidence untuk atribut Lama\_Studi dengan nilai TERLAMBAT adalah **0,476** 5. Dari hasil percobaan Tugas nomor 3 di atas, berapa orang yang akan lulus TEPAT, dan berapa orang yang yang akan lulus TERLAMBAT?

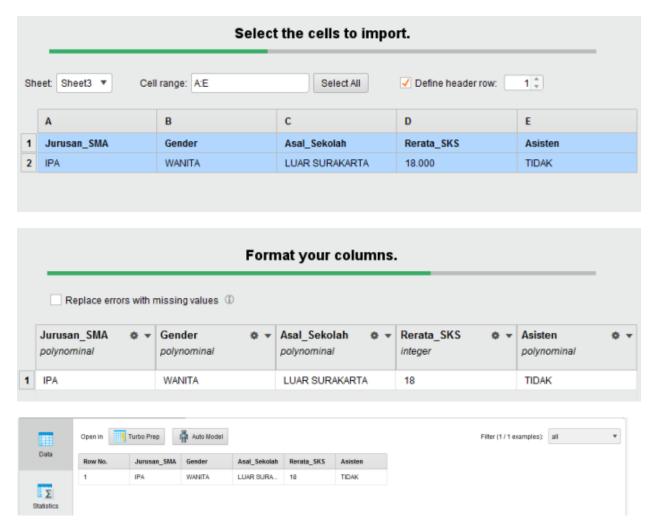


Jadi yang lulus dengan tepat sebanyak 3 orang sedangkan yang lulus terlambat sebanyak 7 orang.

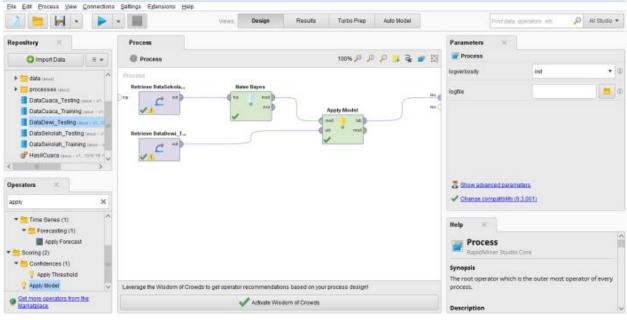
6. Prekdisikan ketepatan lama studi si Dewi, jika Dewi adalah seorang WANITA yang berasal dari jurusan IPA pada saat SMA, asal sekolah dari LUAR SURAKARTA, mengambil SKS dengan

rata-rata sebanyak 18 SKS tiap semester, dan tidak pernah menjadi Asisten selama kuliah.

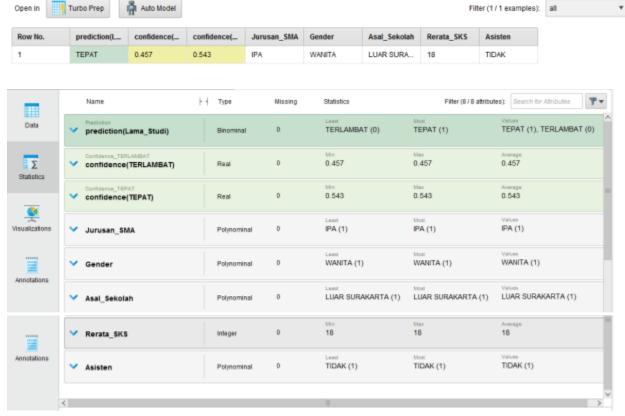
a. Mengimport data Dewi



b. Mengimplementasikan Naive Bayes



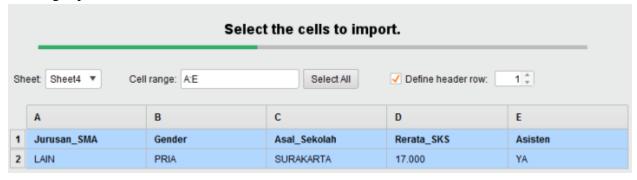
c. Hasil

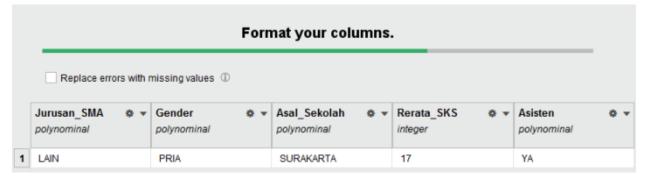


Berdasarkan hasil di atas maka prediksi untuk Dewi dia akan lulus dengan tepat

7. Prekdisikan ketepatan lama studi si Jono, jika Jono adalah seorang PRIA yang berasal dari jurusan IPA pada saat SMA, asal sekolah dari SURAKARTA, mengambil SKS dengan rata-rata sebanyak 17 SKS tiap semester, dan pernah menjadi Asisten selama kuliah.

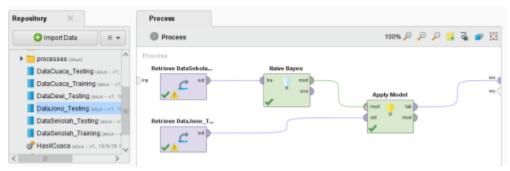
## a. Mengimport Data Jono



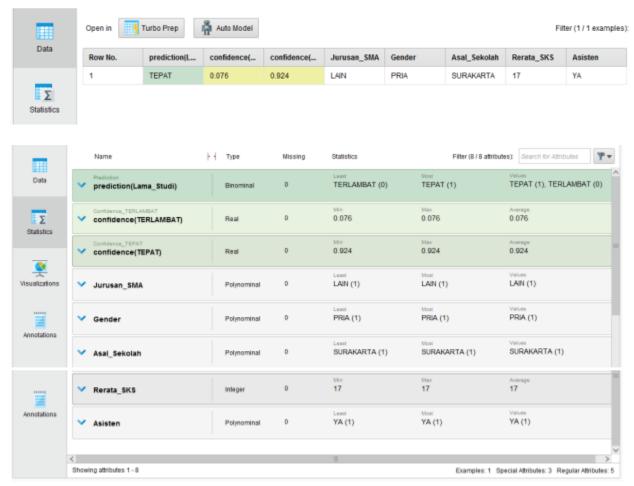




### b. Mengimplementasikan Naive Bayes



c. Hasil



Berdasarkan hasil di atas maka prediksi untuk Jono dia akan lulus dengan tepat