**LAPORAN PRAKTIKUM DATA WAREHOUSING DAN DATA**

**MINING**

**MODUL 8**

**“KLASIFIKASI: NAIVE BAYES”**



**Oleh:**

**NAMA : Daffa Putra Alwansyah NIM : L200190031**

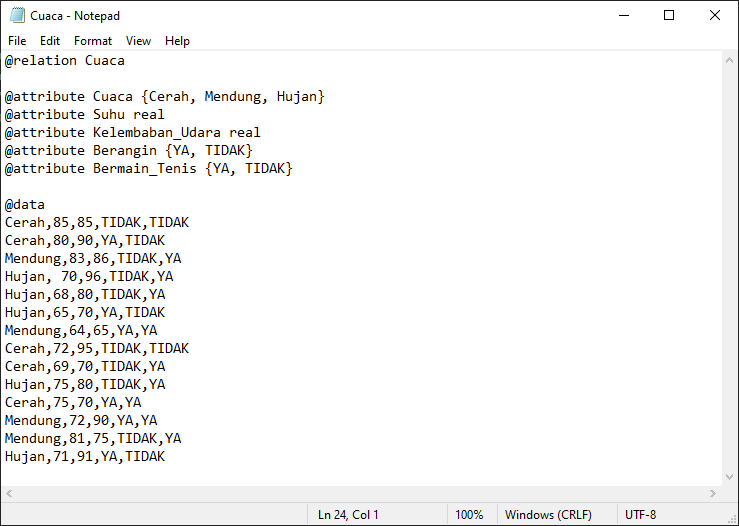
**KELAS : B**

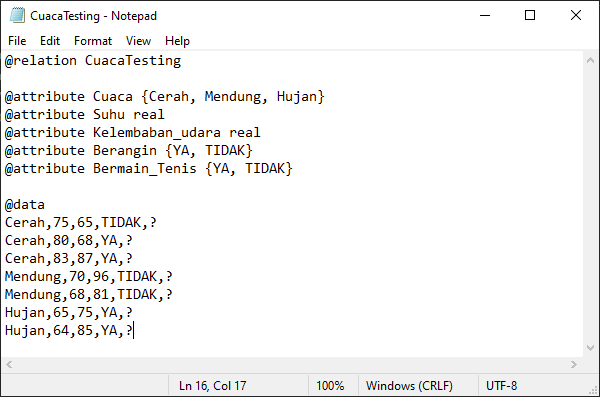
**PRODI : INFORMATIKA**

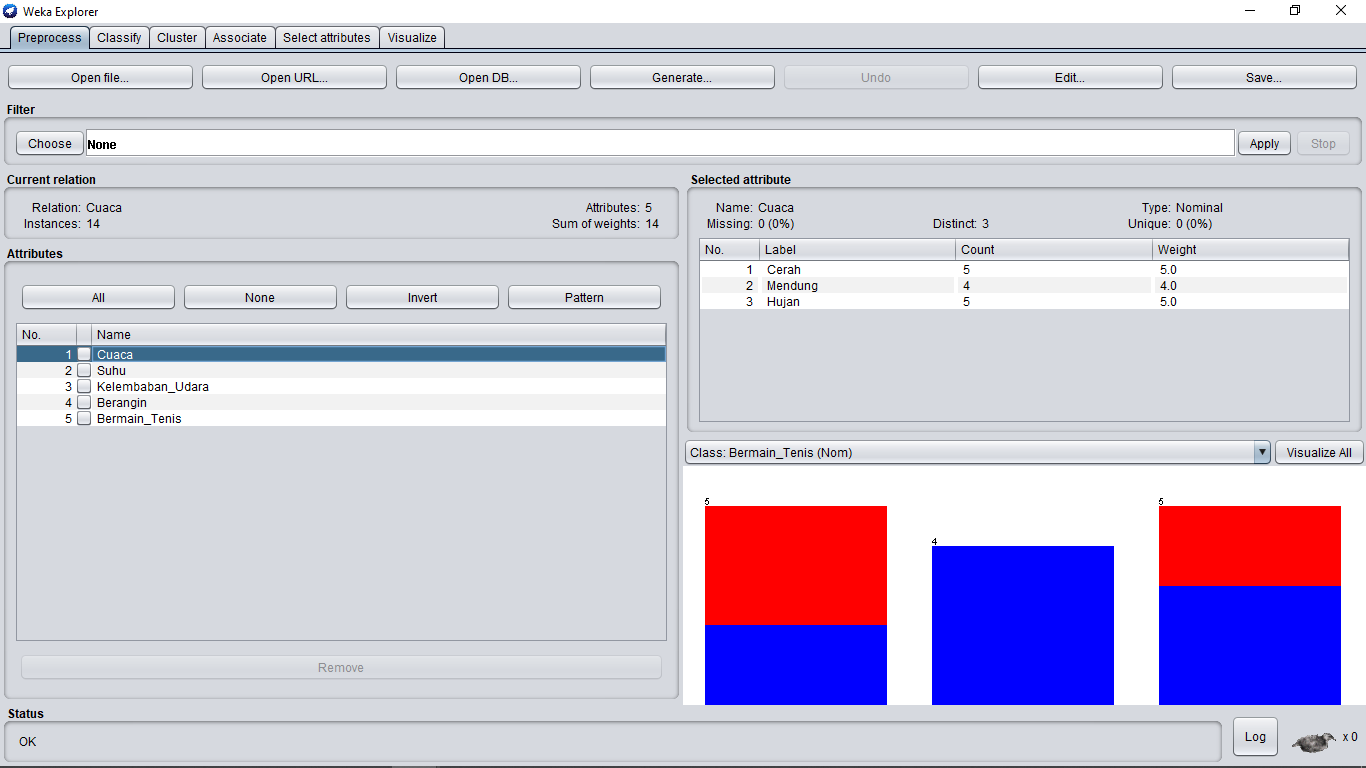
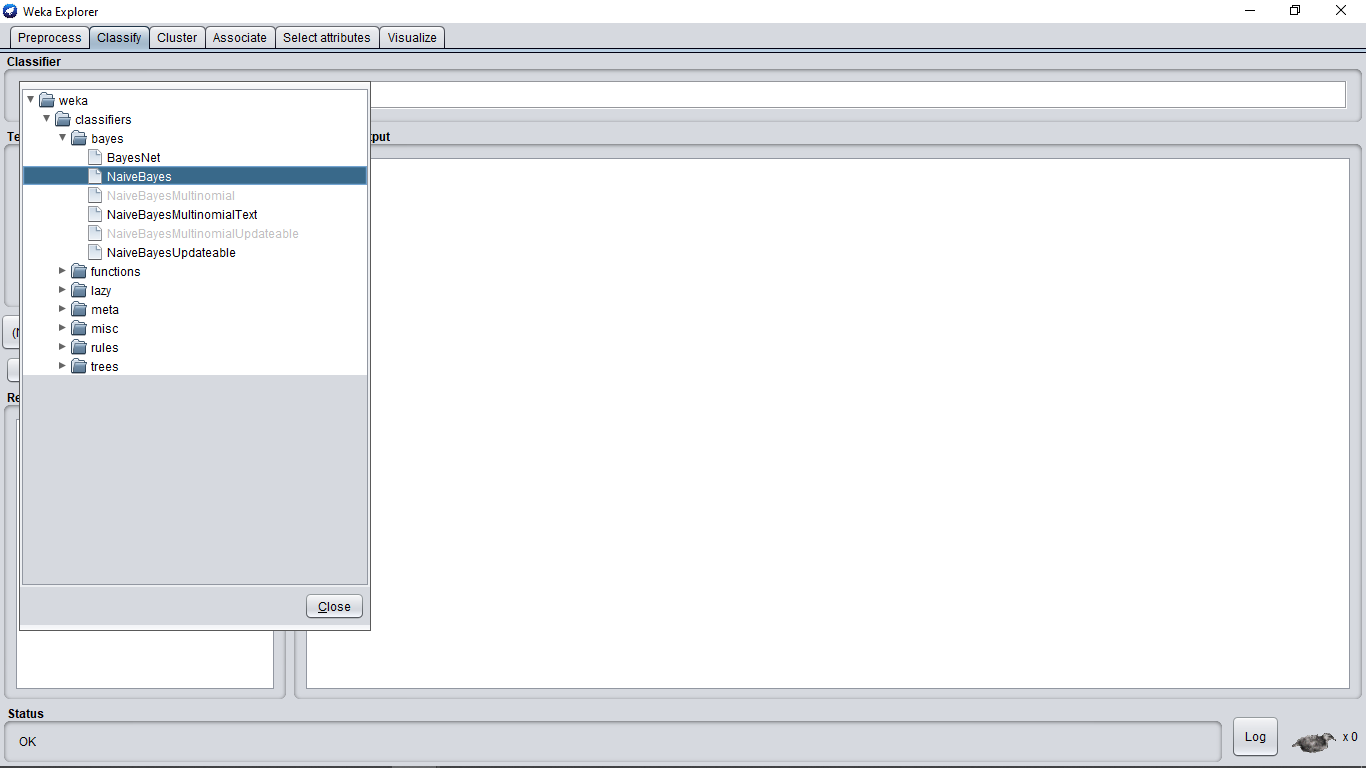
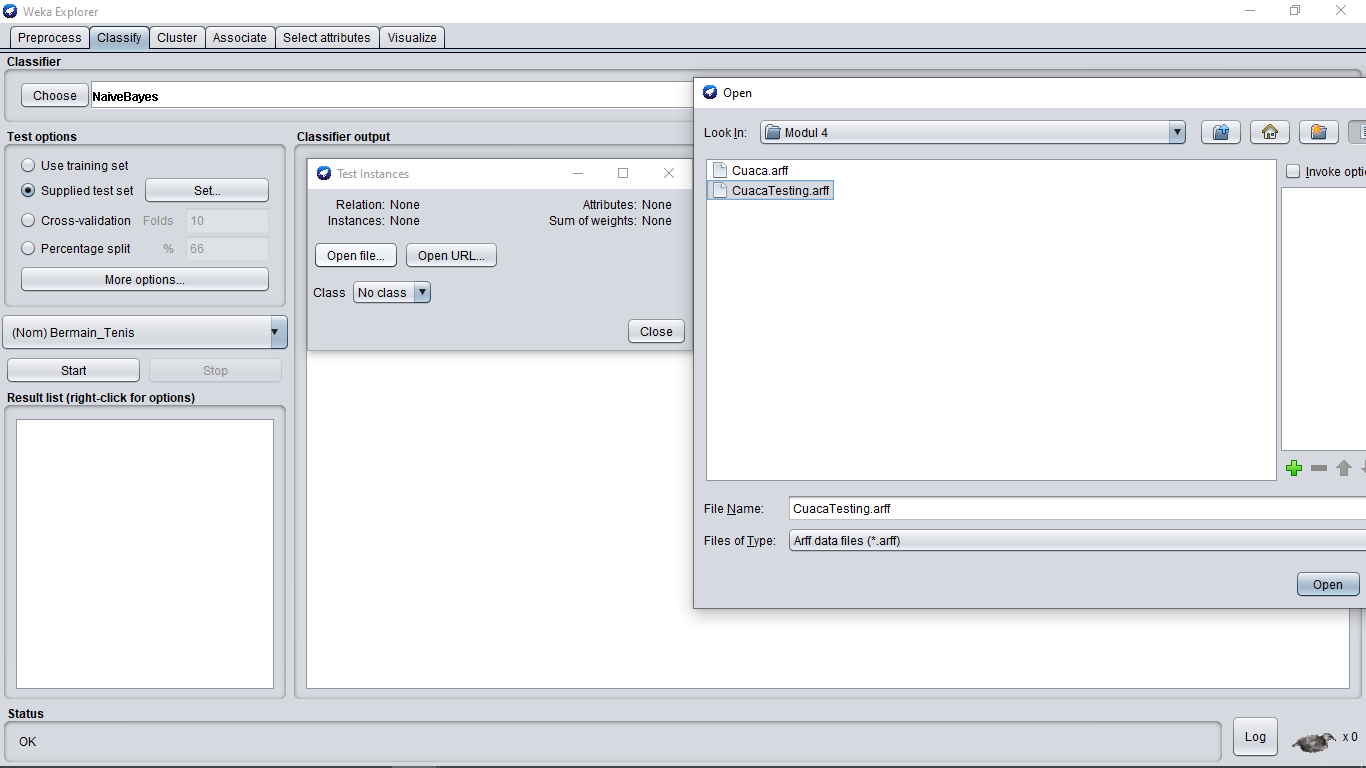
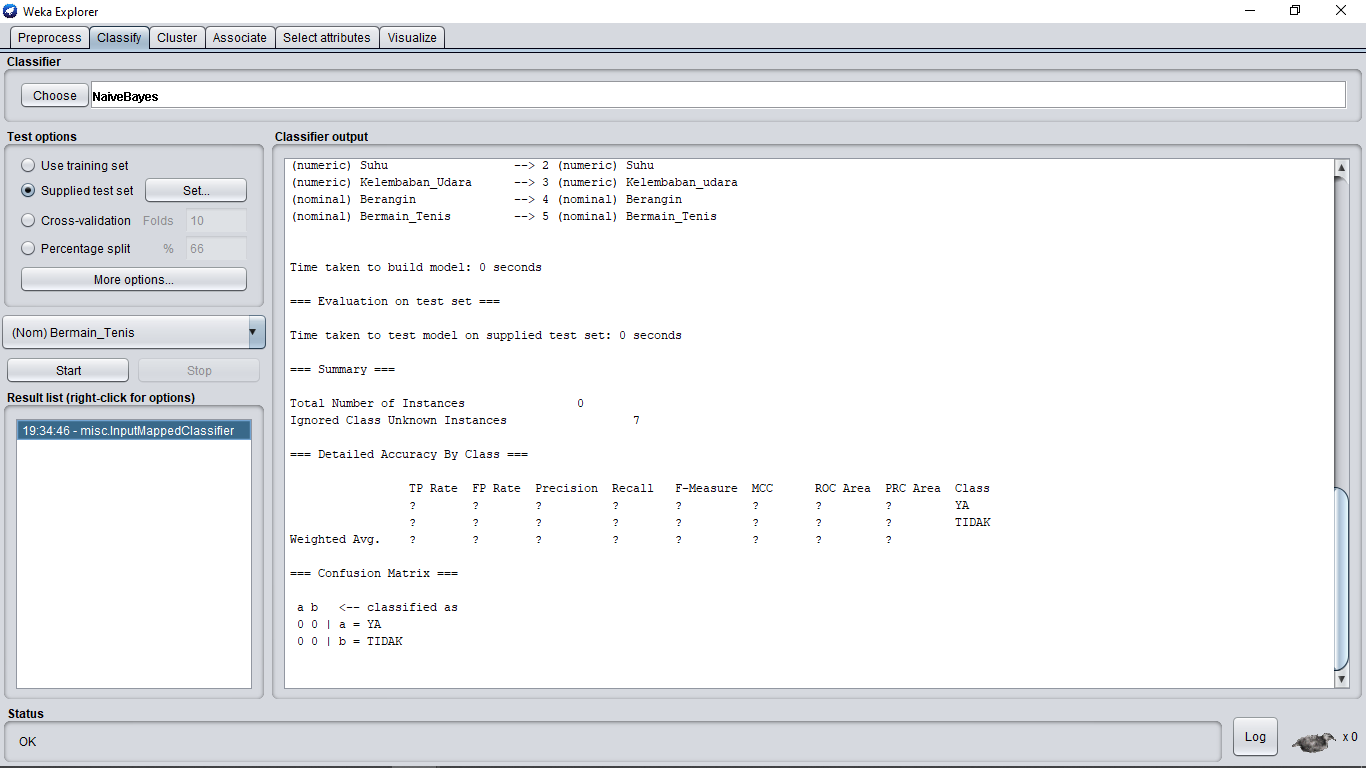
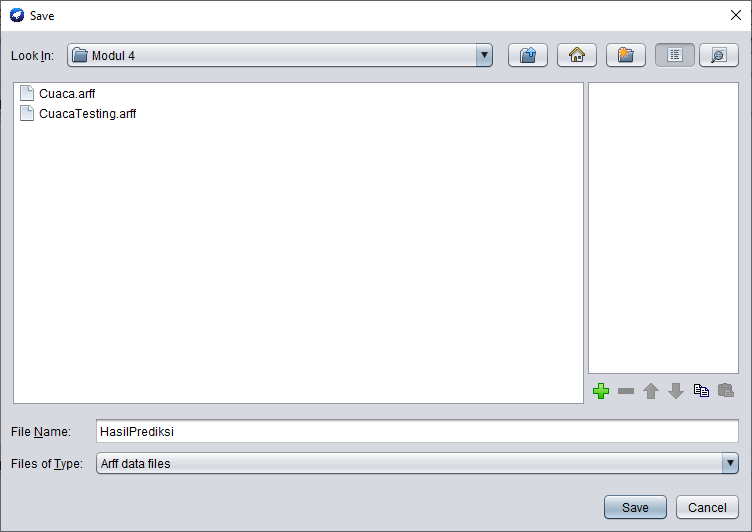
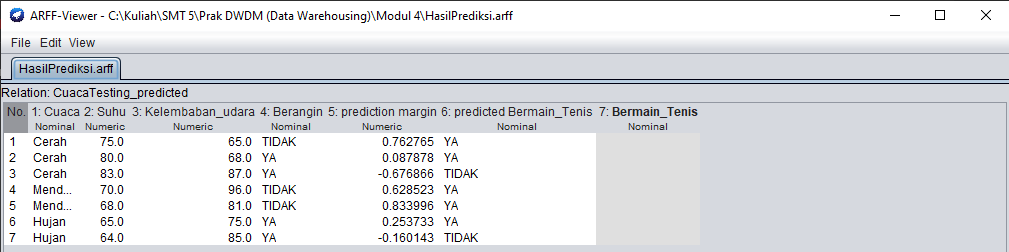
**Fakultas Komunikasi dan Informatika   
Universitas Muhammadiyah Surakarta**

**Langkah-Langkah Praktikum**

* **Implementasi Naïve Bayes dengan Weka**

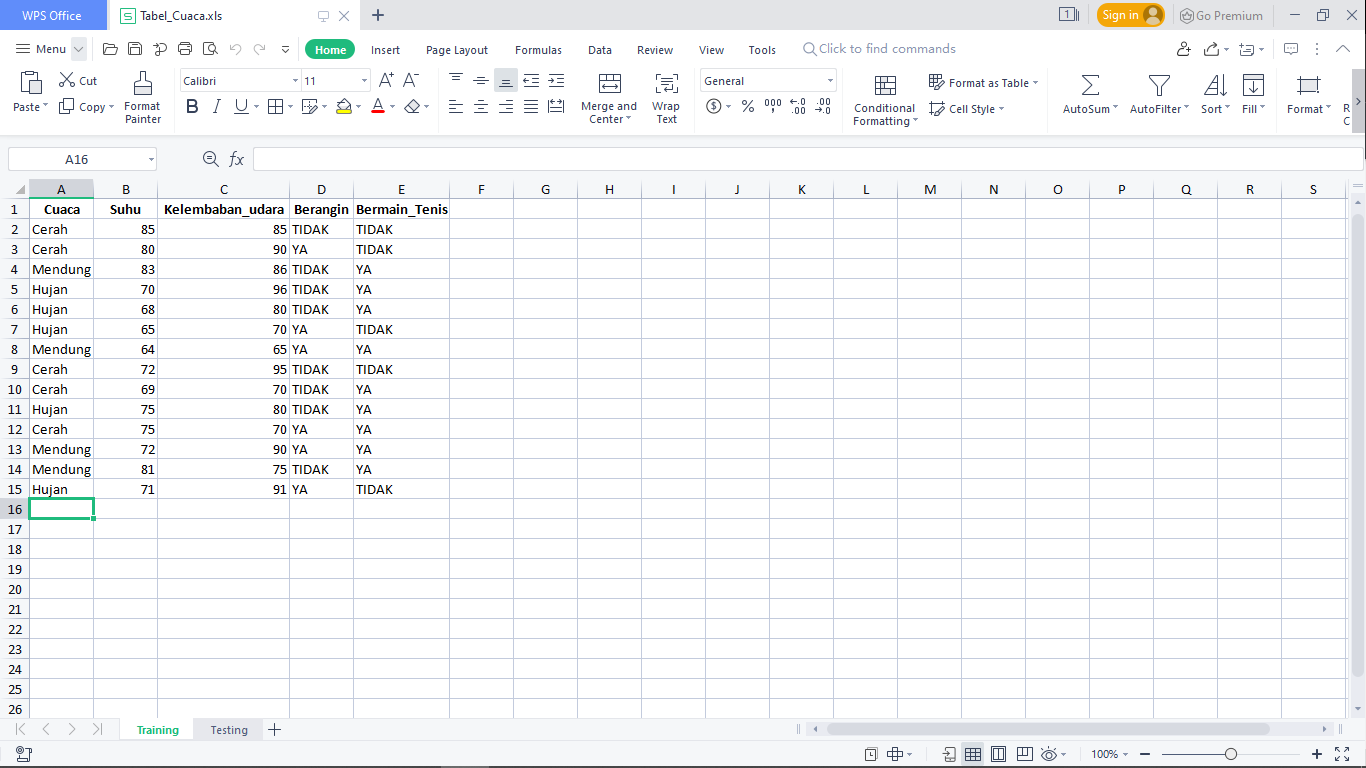
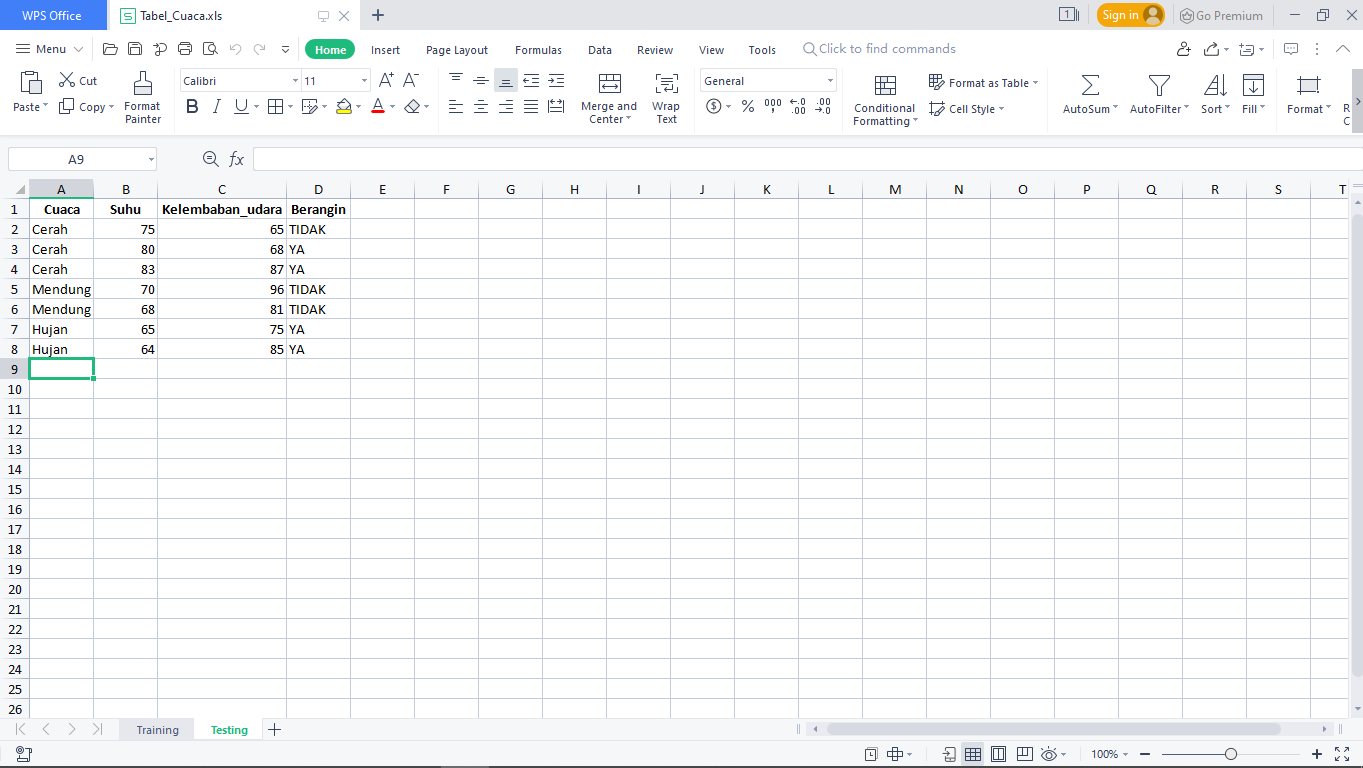
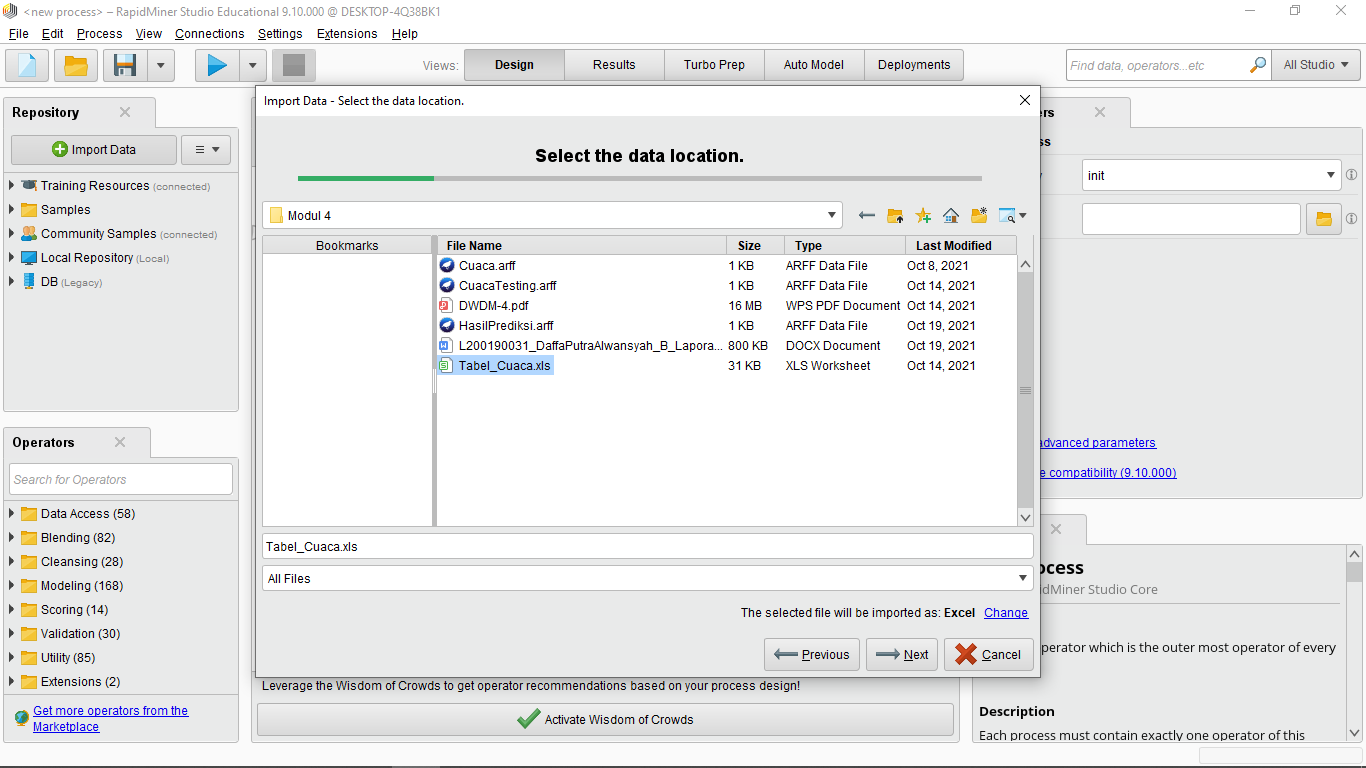
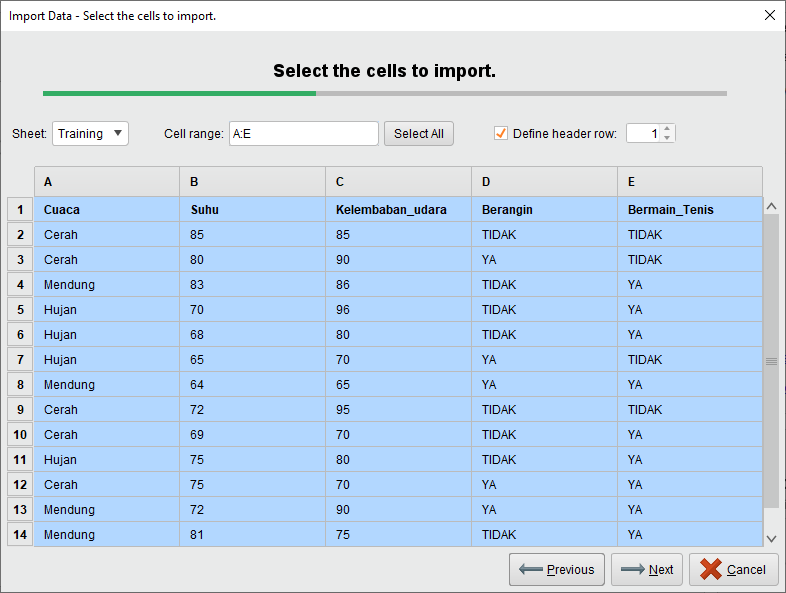
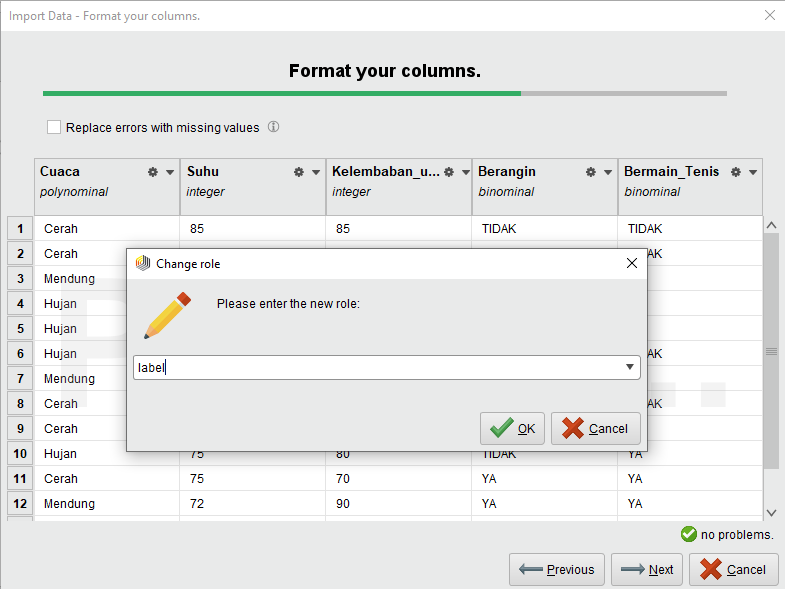
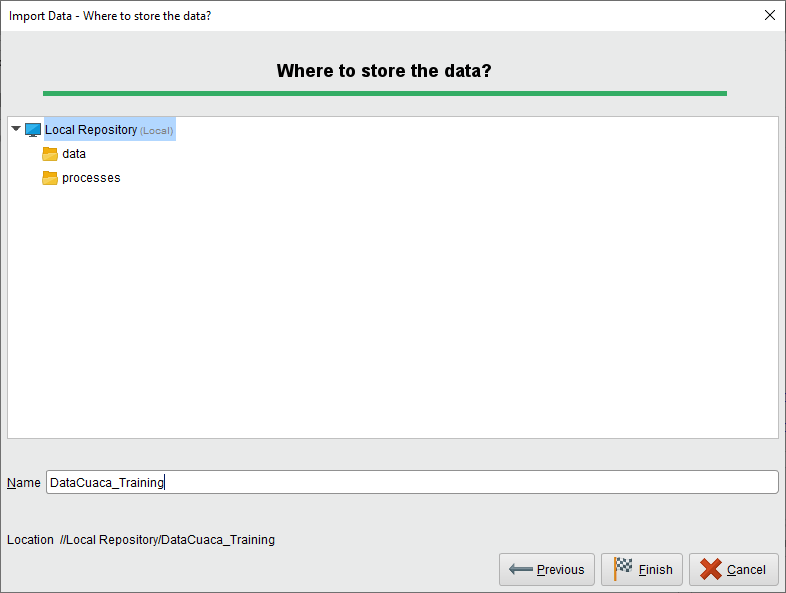
1. Menyiapkan file Cuaca.arff yang akan kita gunakan sebagai data training.  


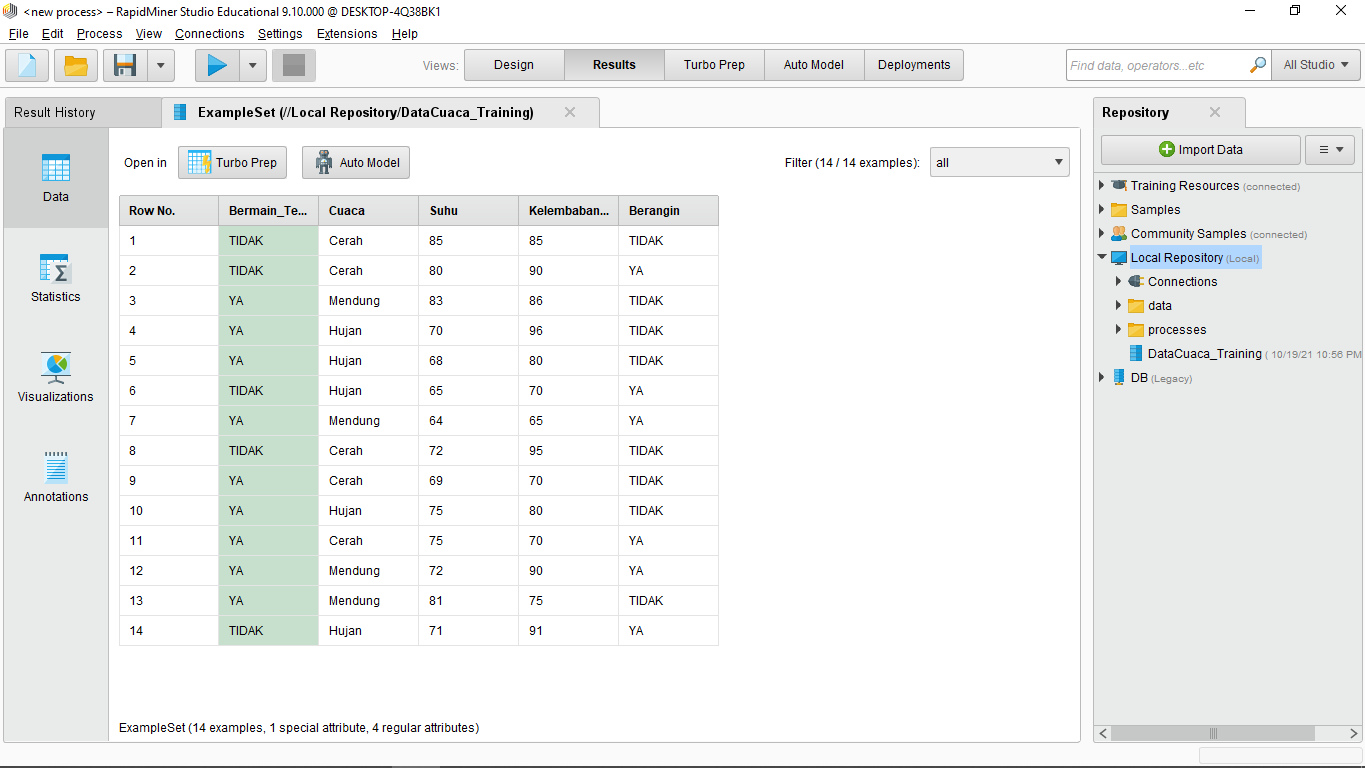
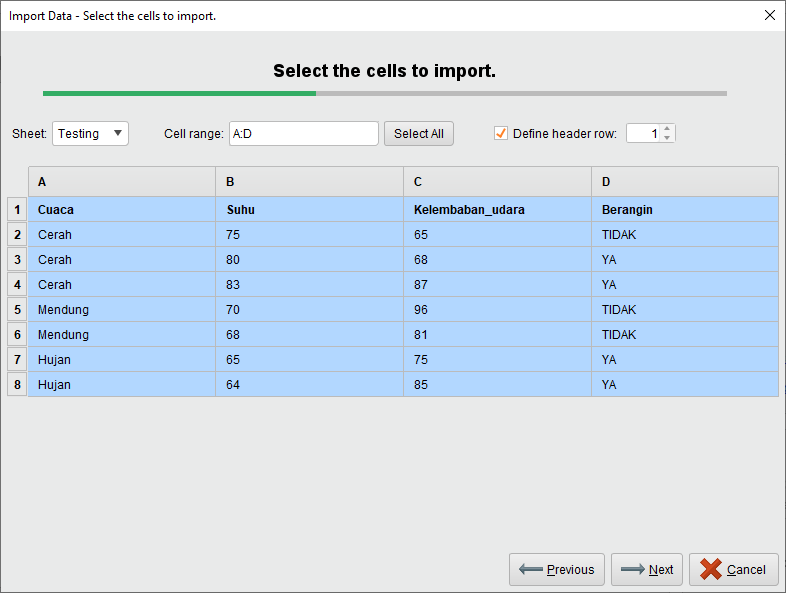
2. Membuat data testing sebagai data uji yang akan diprediksi dengan memiliki variable-variabel independent dan variable dependen yang sama. Diberi nama CuacaTesting.arff  


3. Buka kembali file Cuaca.arff dengan aplikasi weka  
  
  
4. Setelah itu pilih tab Classify, kemudian klik Choose dan pilih NaiveBayes.  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
5. Pilih Supplied test set, klik Open file dan pilih CuacaTesting.arff.  
  
  
6. Class yang dipilih adalah Bermain\_Tenis, setelah itu klik Start untuk memulai proses naïve bayes, setelah itu klik kanan pada result, dan pilih ‘visualize classifier errors’.  
  
  
  
  
  
  
  
  
  
  
  
  
7. Setelah itu, beri dengan nama HasilPrediksi.arff dan klik save.  
  
  
8. Tutup Weka Exploer, kembali ke menu utama, pilih Tool > ArffViewer, buka file HasilPRediksi.arff yang sudah disimpan tadi.  


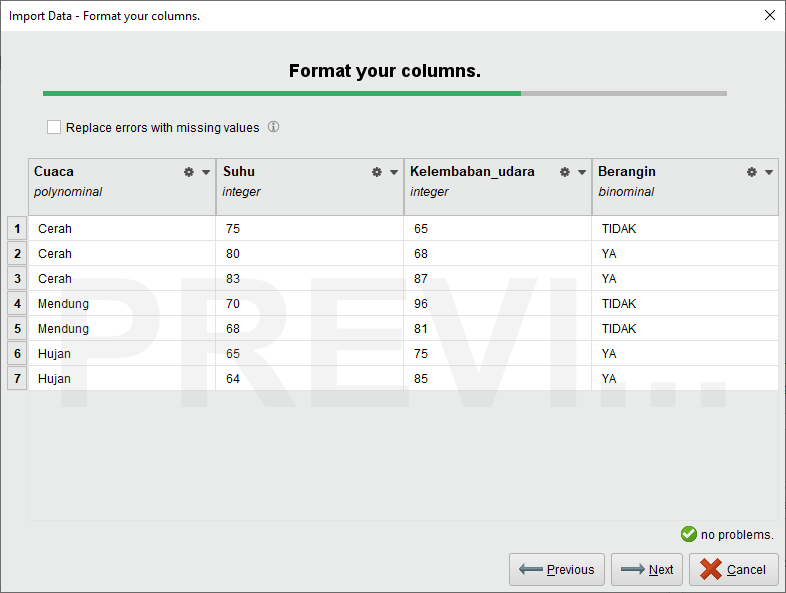
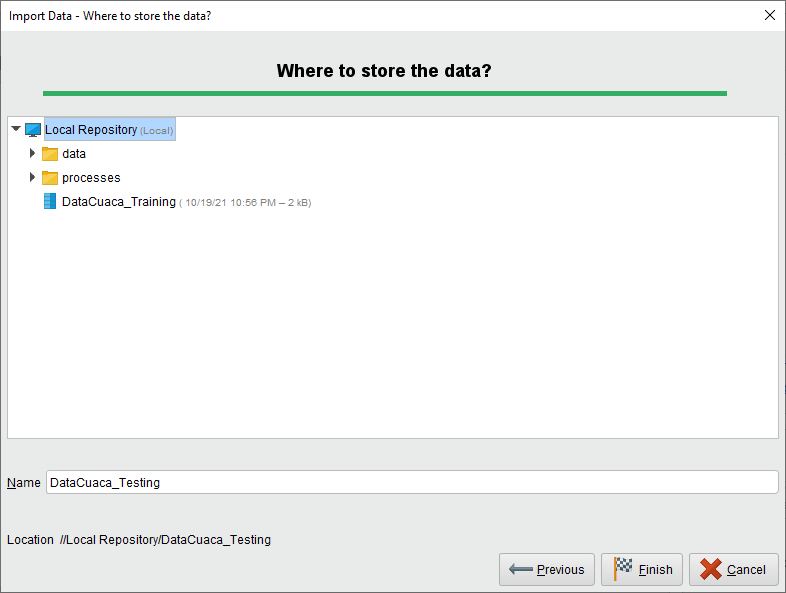
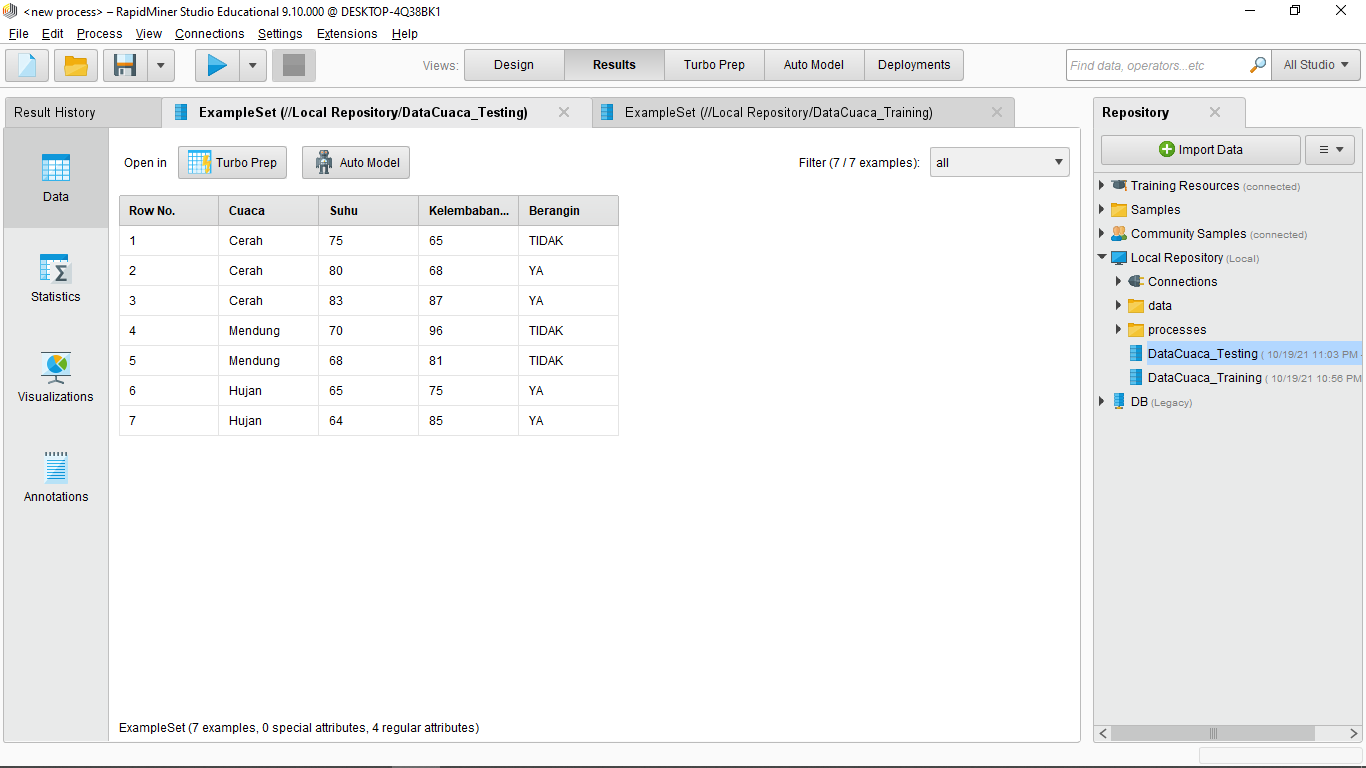
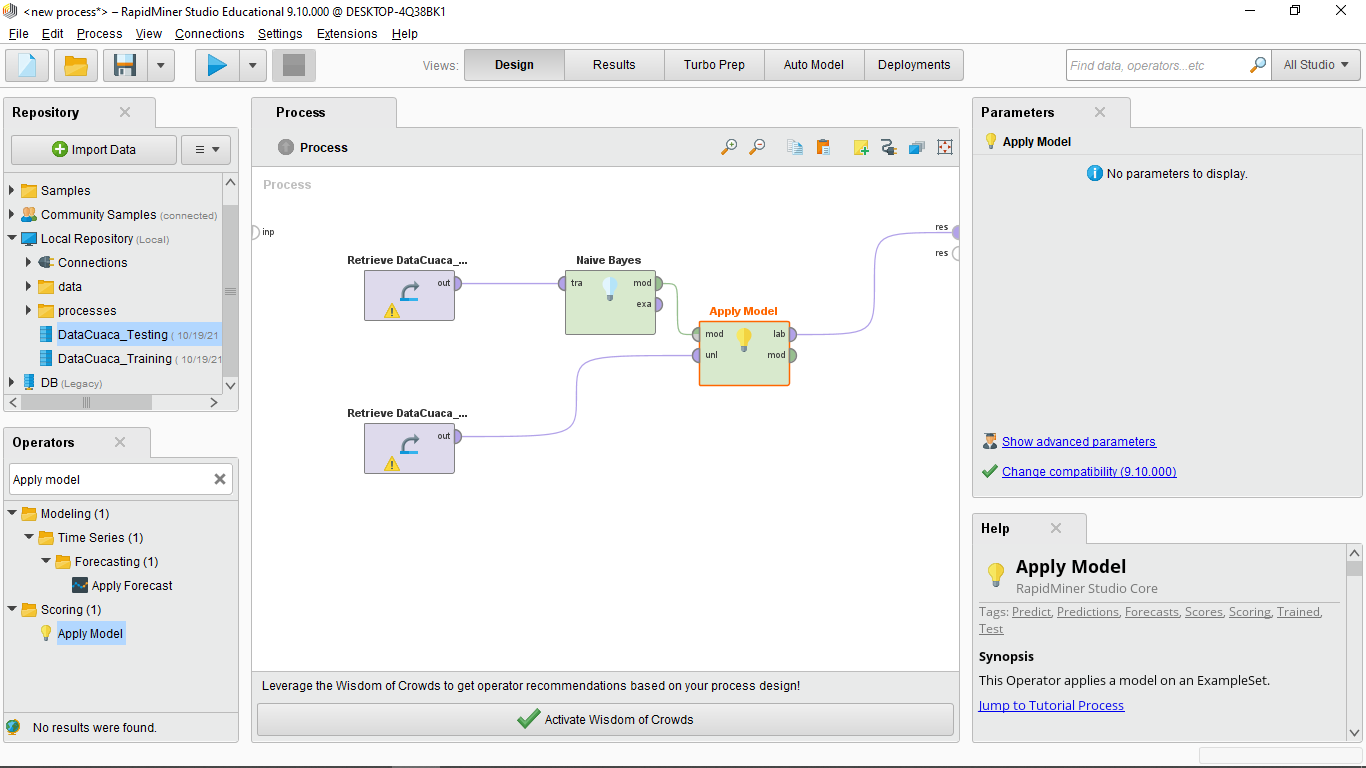
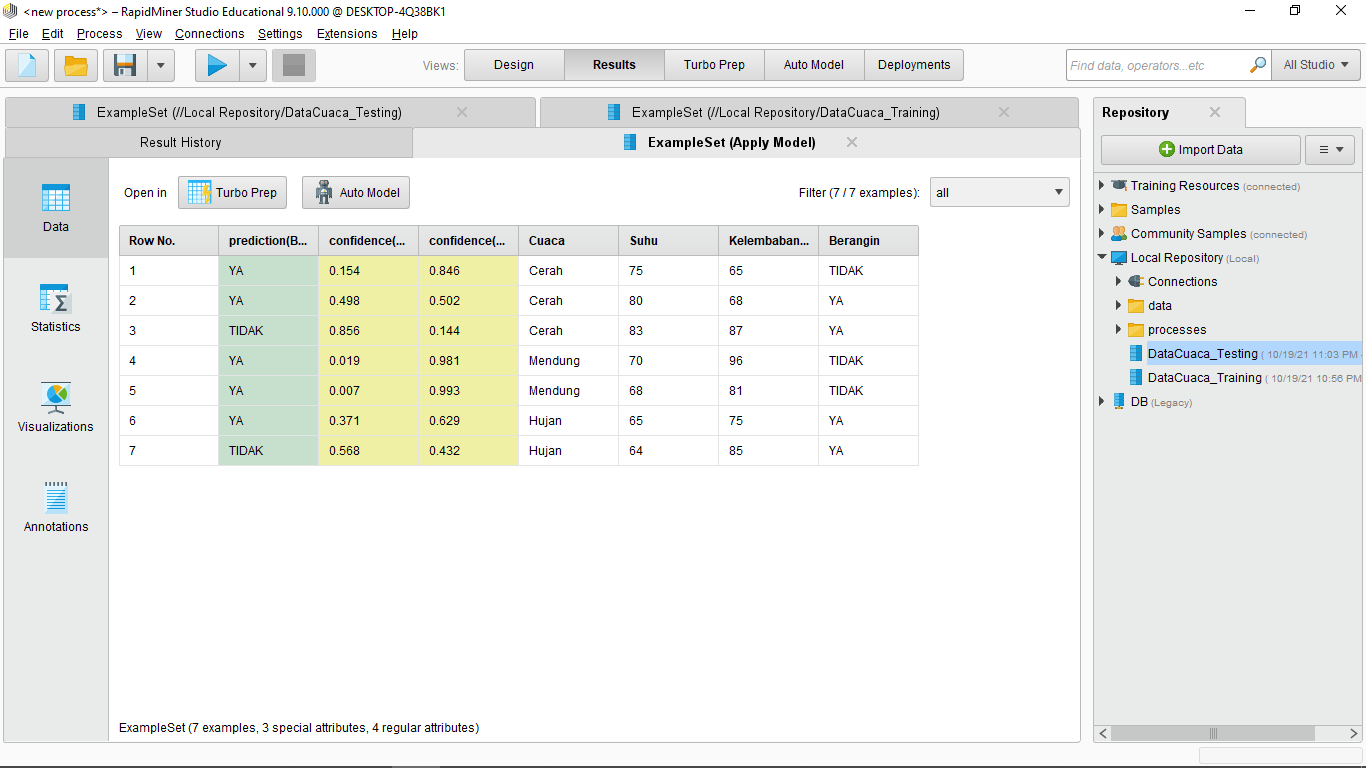
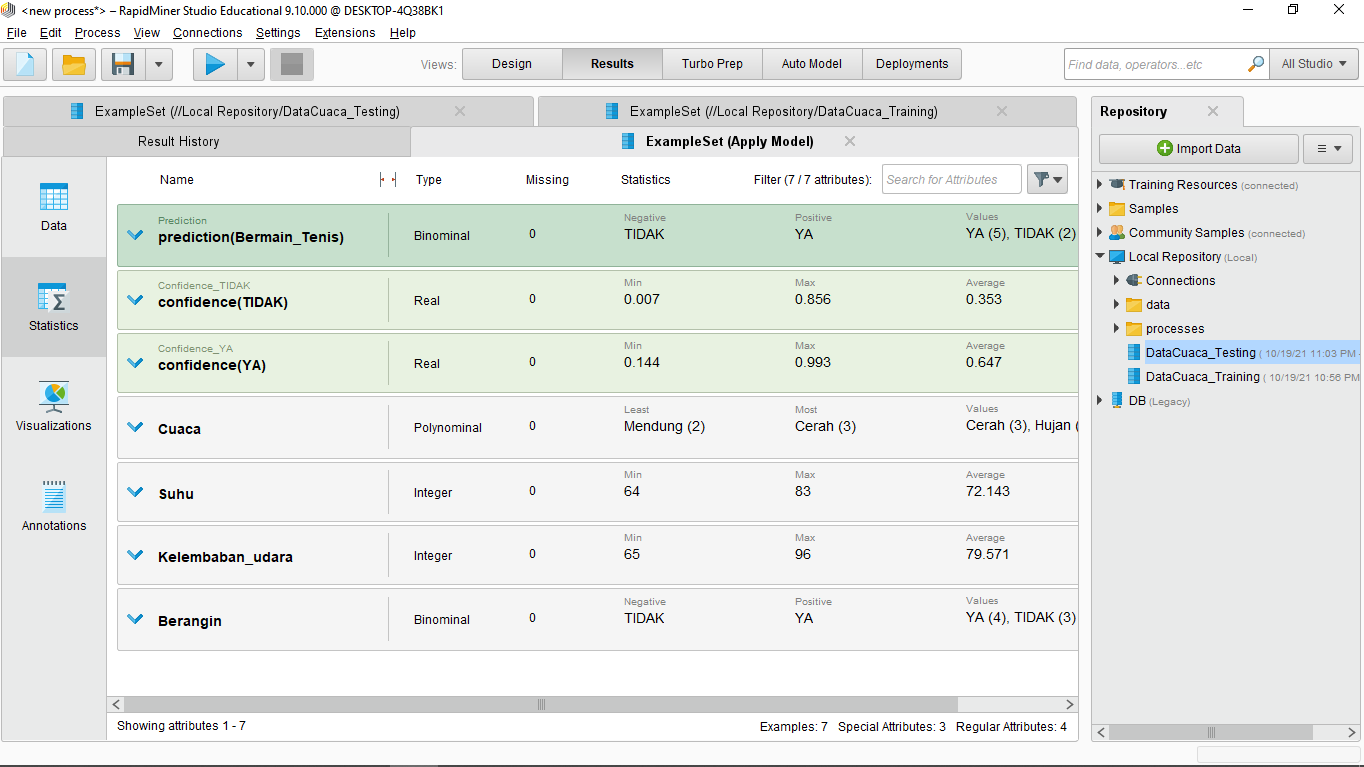
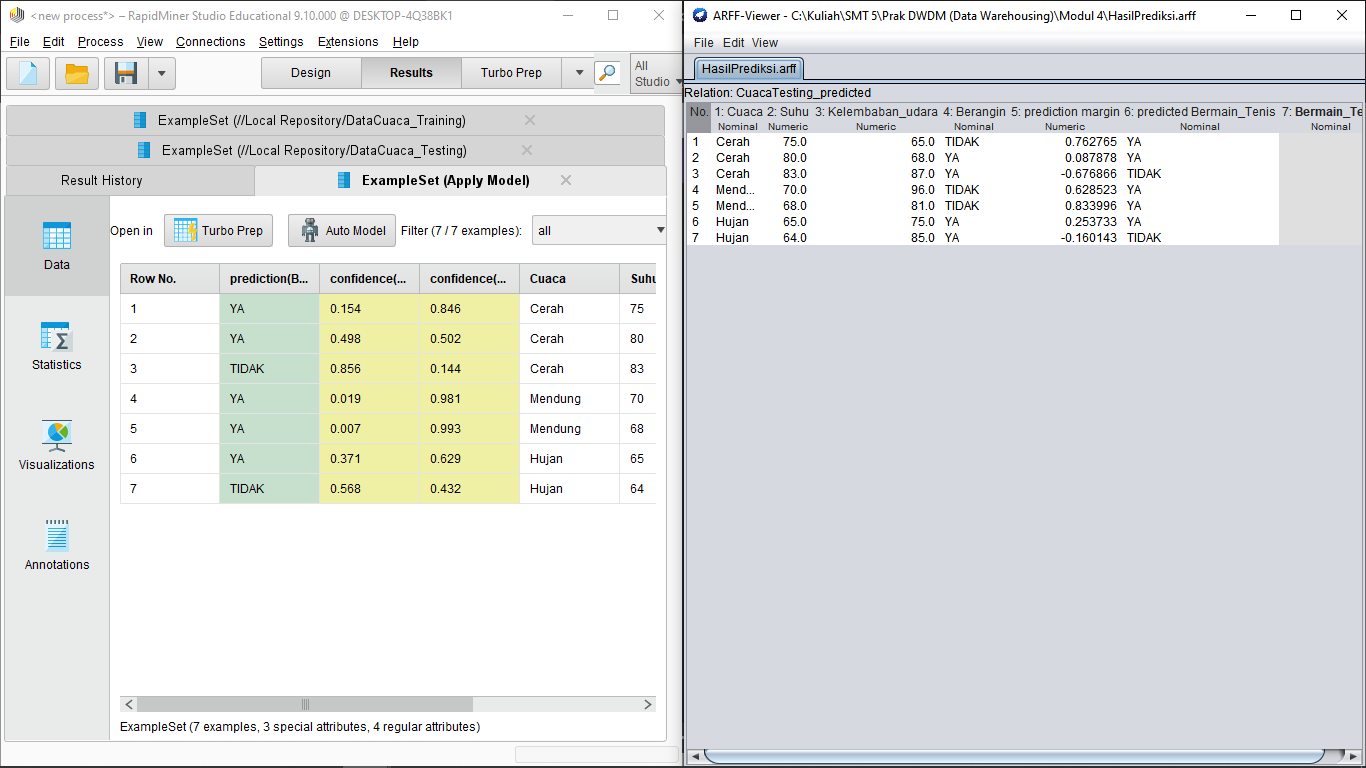
* **Implementasi Naïve Bayes dengan RapidMiner**

1. Mempersiapkan file Tabel\_Cuaca.xls yang terdiri dari 2 sheet, sheet pertama berupa Training dan sheet kedua berupa Testing. Sheet 1 digunakan sebagai data training dan sheet2 digunakan sebagai data uji.

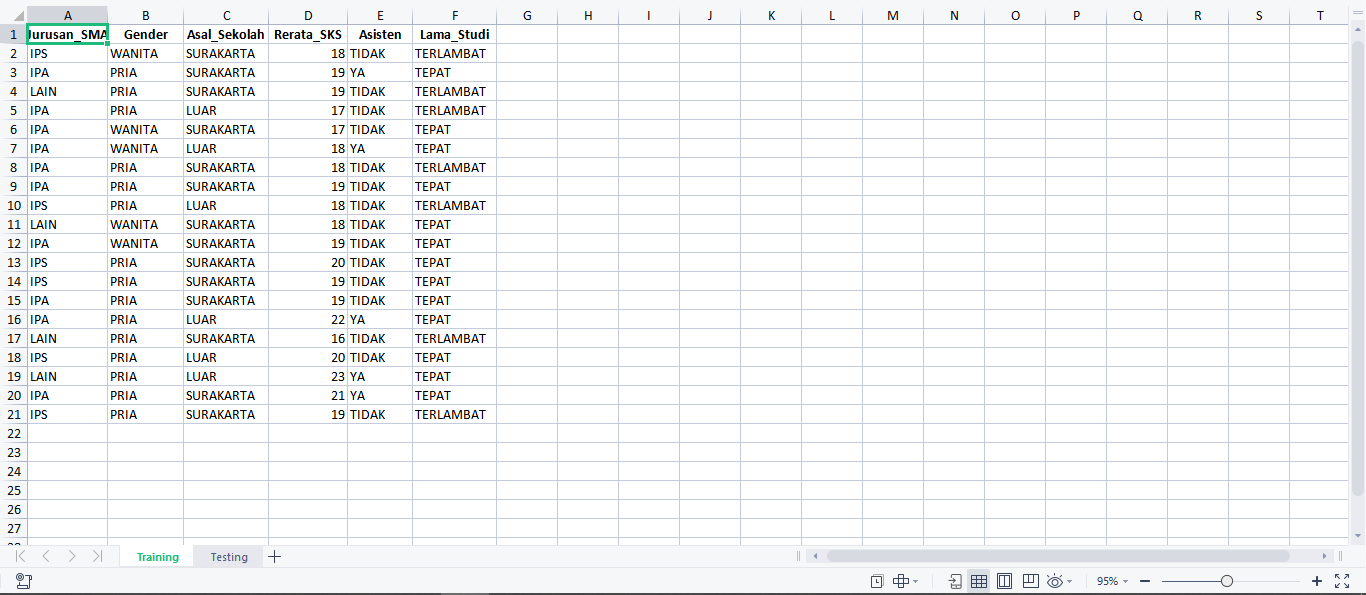
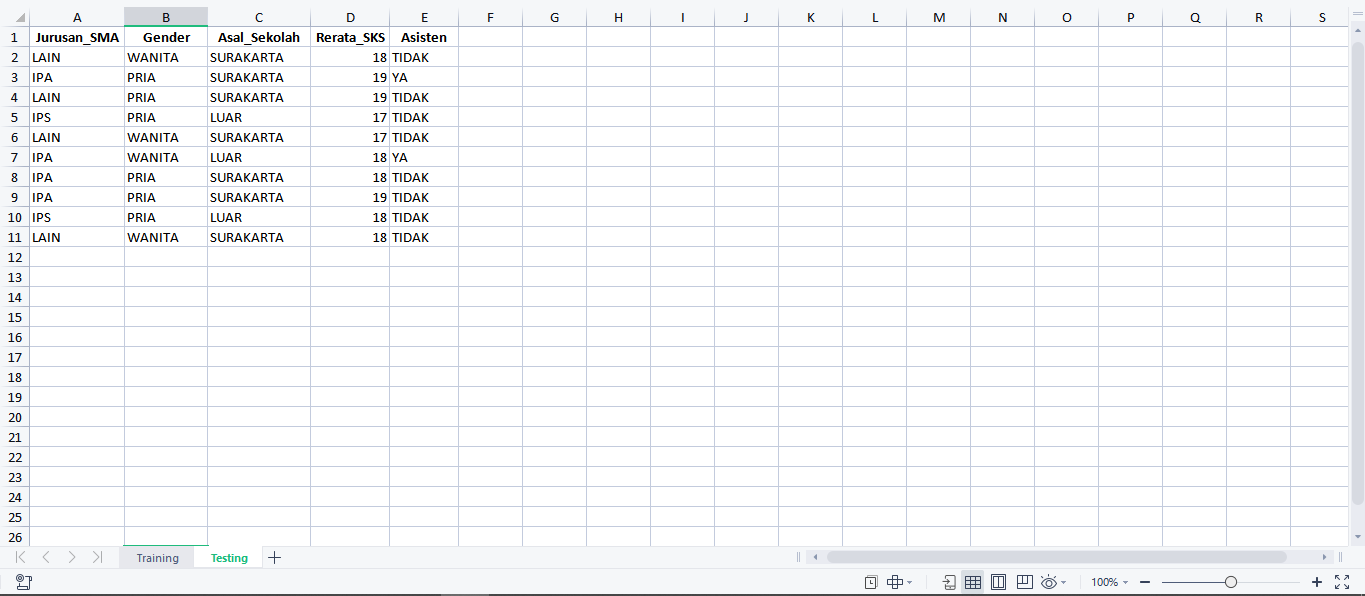
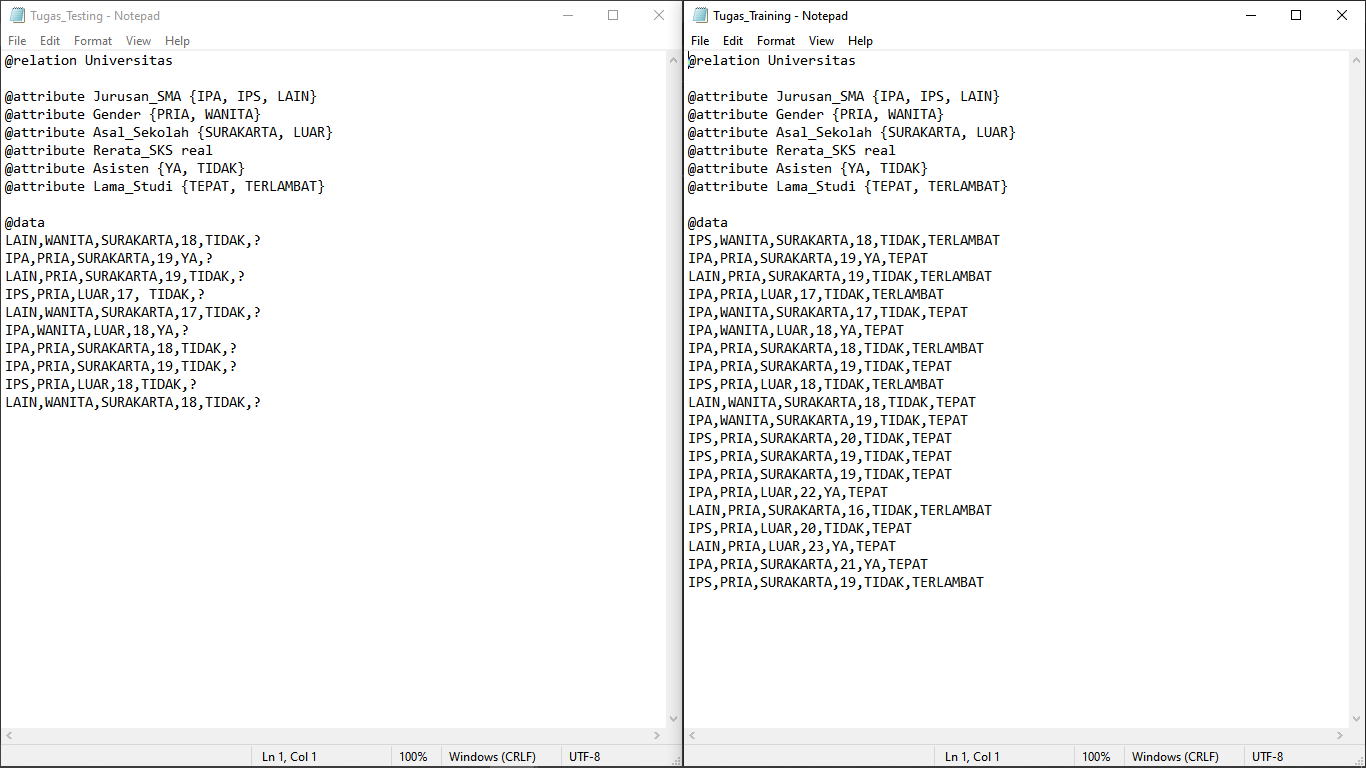
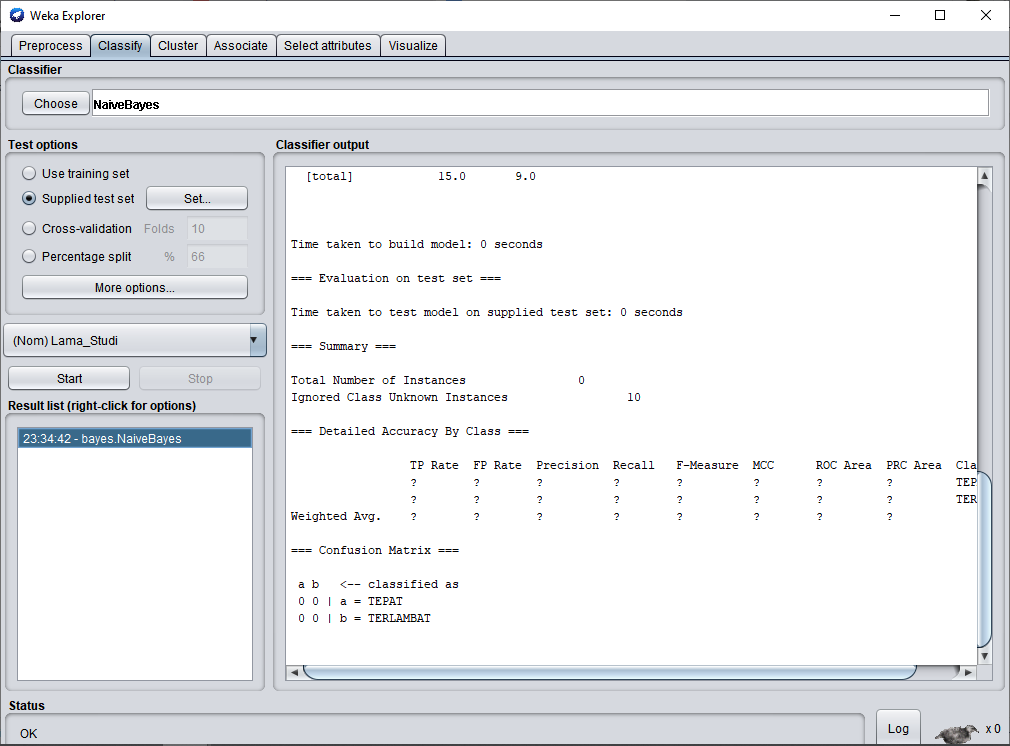
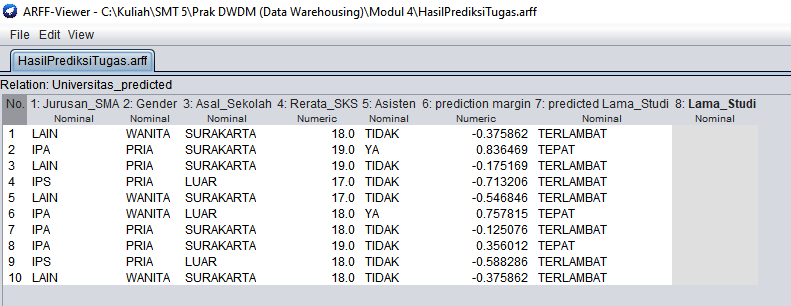
  
  
  
  
  
  
  
  
  
  
  
  
  
2. Buka aplikasi RapidMiner, lalu import data Tabel\_Cuaca.xls  
  
  
3. Memastikan sel excel sesuai di langkah Select the cells to import  
  
  
  
  
  
  
  
  
4. Setelah itu ubah type Berangin dan Bermain\_Tenis menjadi ‘binominal’ dan ubah role Bermain\_Tenis menjadi label.  
  
  
5. Simpan dengan nama DataCuaca\_Training.  


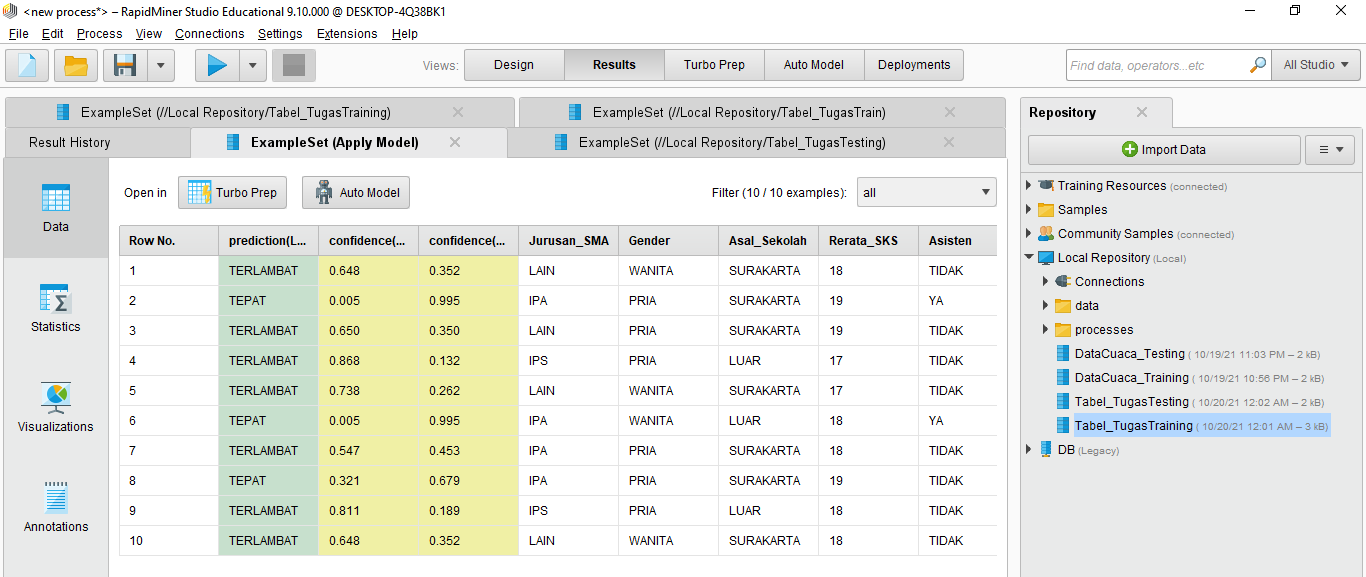
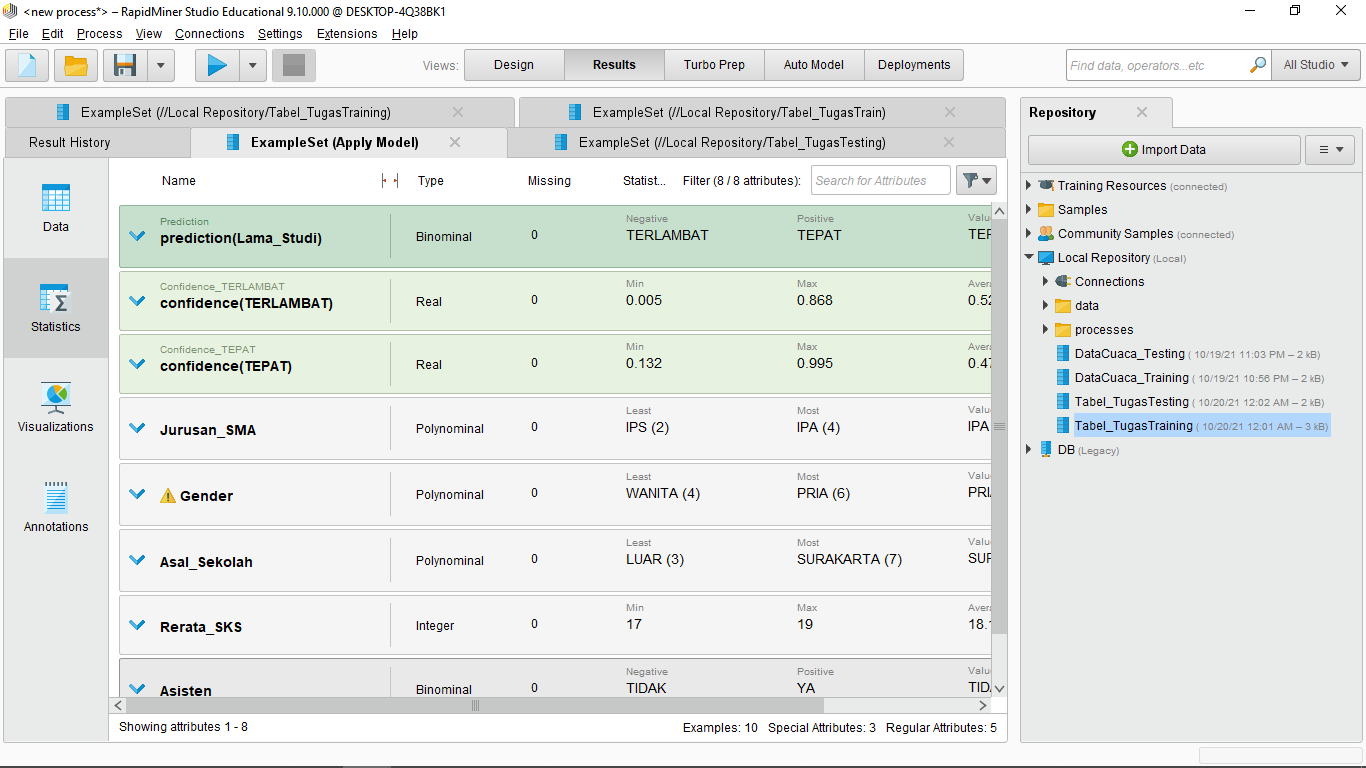
6. Menampilkan hasil import file Tabel\_Cuaca.xls pada sheet1t file Tabel\_Cuaca.xls pada sheet1(Training).  
  
  
7. Ulangi hal yang sama pada sheet1 tadi, tapi disini Sheet menggunakan Testing.  


8. Ubah Berangin menjadi type ‘binominal’, dan simpan dengan nama DataCuaca\_Testing.

  
  
  
  
9. Menampilkan hasil import file Tabel\_Cuaca.xls pada sheet2(Testing).  
  
  
10. Membuat desain Naïve Bayes dengan mendrag DataCuaca\_Training dan DataCuaca\_Testing serta memasukkan operator Naïve Bayes dan Aplly Model kedalam jendela process view.  
  
  
  
  
  
  
  
  
  
  
  
  
11. Menjalankan proses Naïve Bayes dengan klik tombol panah biru/Run. Nilai kelas pada variable Y (Bermain\_Tenis) rerata nilai confidence sebesar 0,856 untuk nilai TIDAK, dan 0,993 untuk nilai YA.  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
12. Membandingkan hasil RapidMiner dengan Weka.  


* **TUGAS**

1. Membuat file xls dan arff sesuai modul.  
  
  
  
2. Membuat file arff Testing dan Training.  
  
  
3. Setelah itu buka Weka Exploler, open file data Training, pilih tab Classify, set ke data Testing lalu klik Start.  
  
  
  
  
  
  
  
  
  
  
5. Setelah itu save dengna nama HasilPrediksiTugas.arff, berikut adalah hasil prediksi Weka.  
  
  
6. Lakuan langkah seperti di awal tadi untuk mendapatkan prediksi RapidMiner, Hasil prediksi data testing (Excel) menggunakan RapidMiner sebagai berikut.

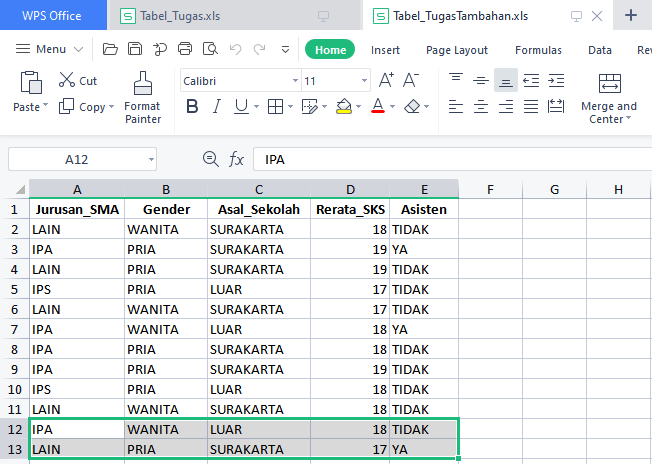
  
  
  
  
  
7. Nilai Rerata Confidance untuk atribut Lama\_Studi dengan nilai TEPAT dan TERLAMBAT

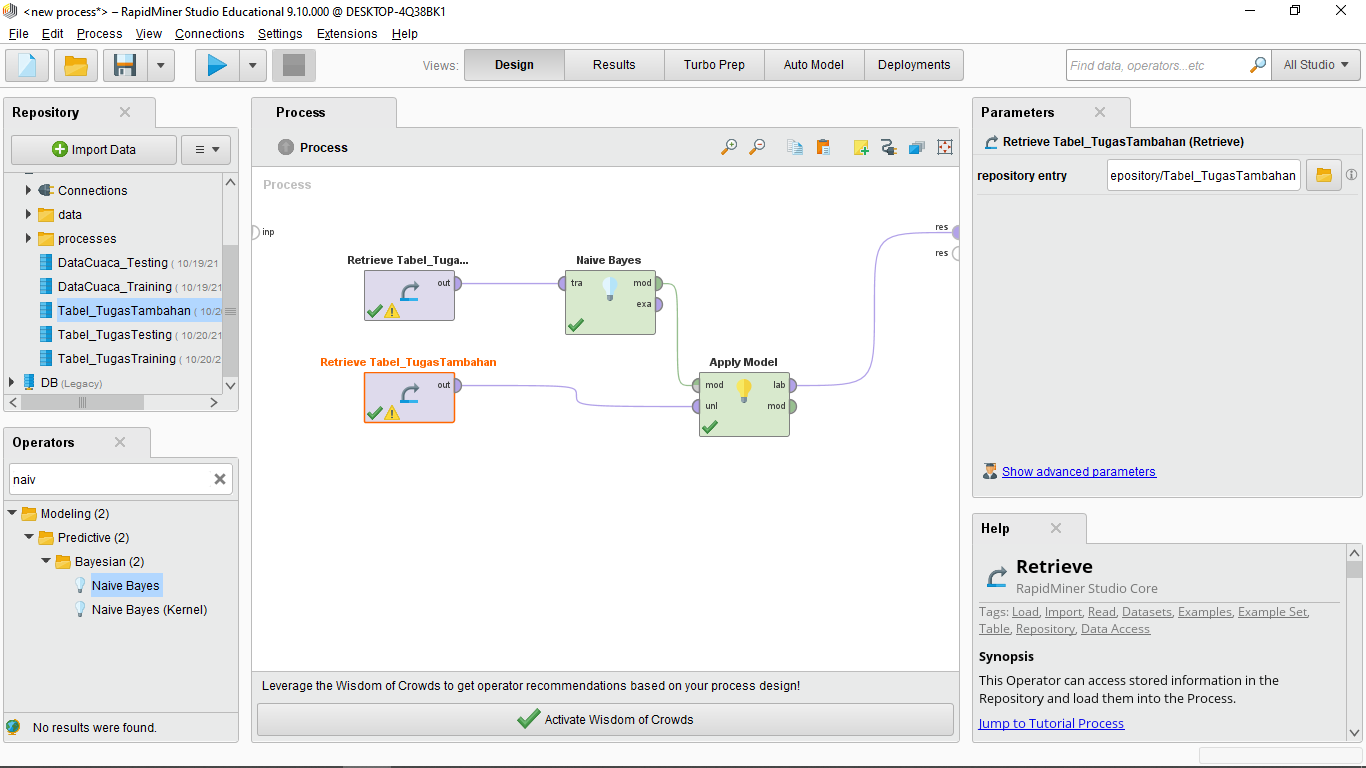
a. Nilai TEPAT : 0,995

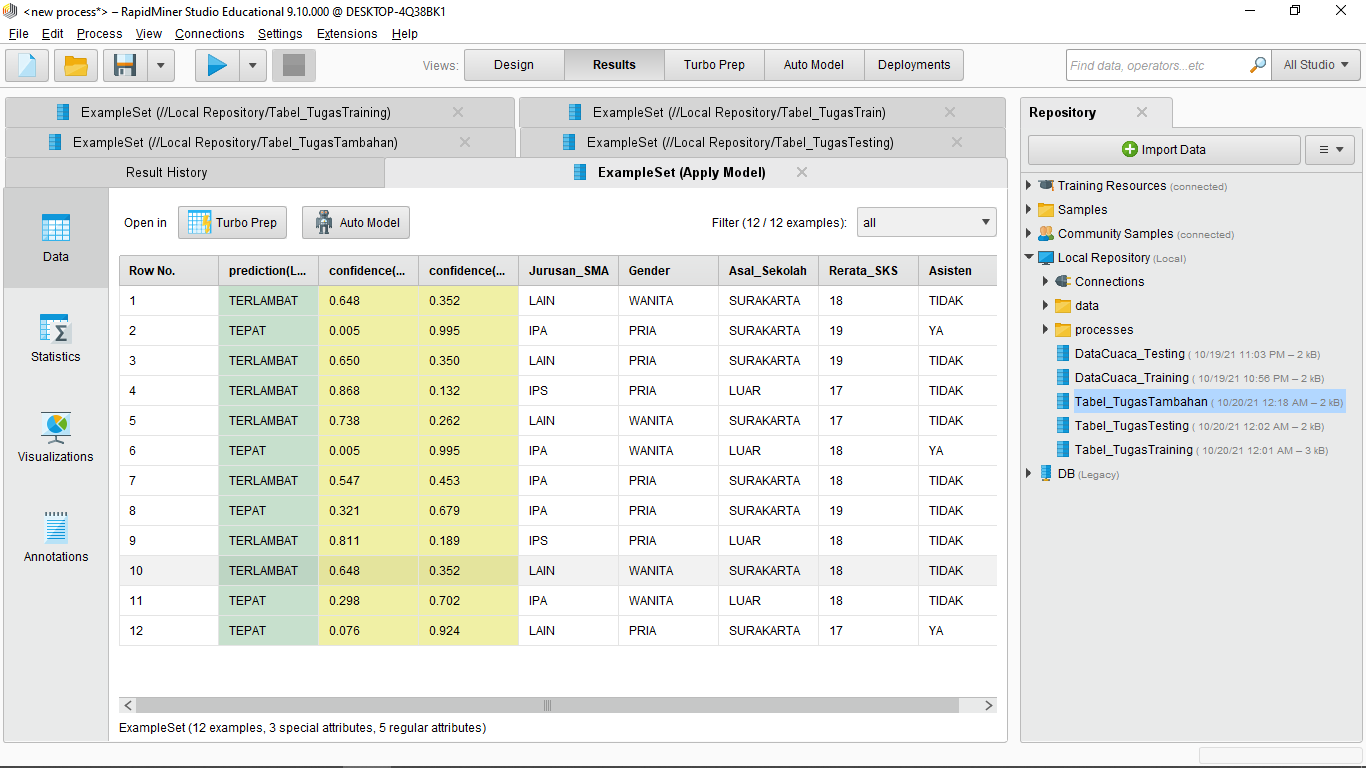
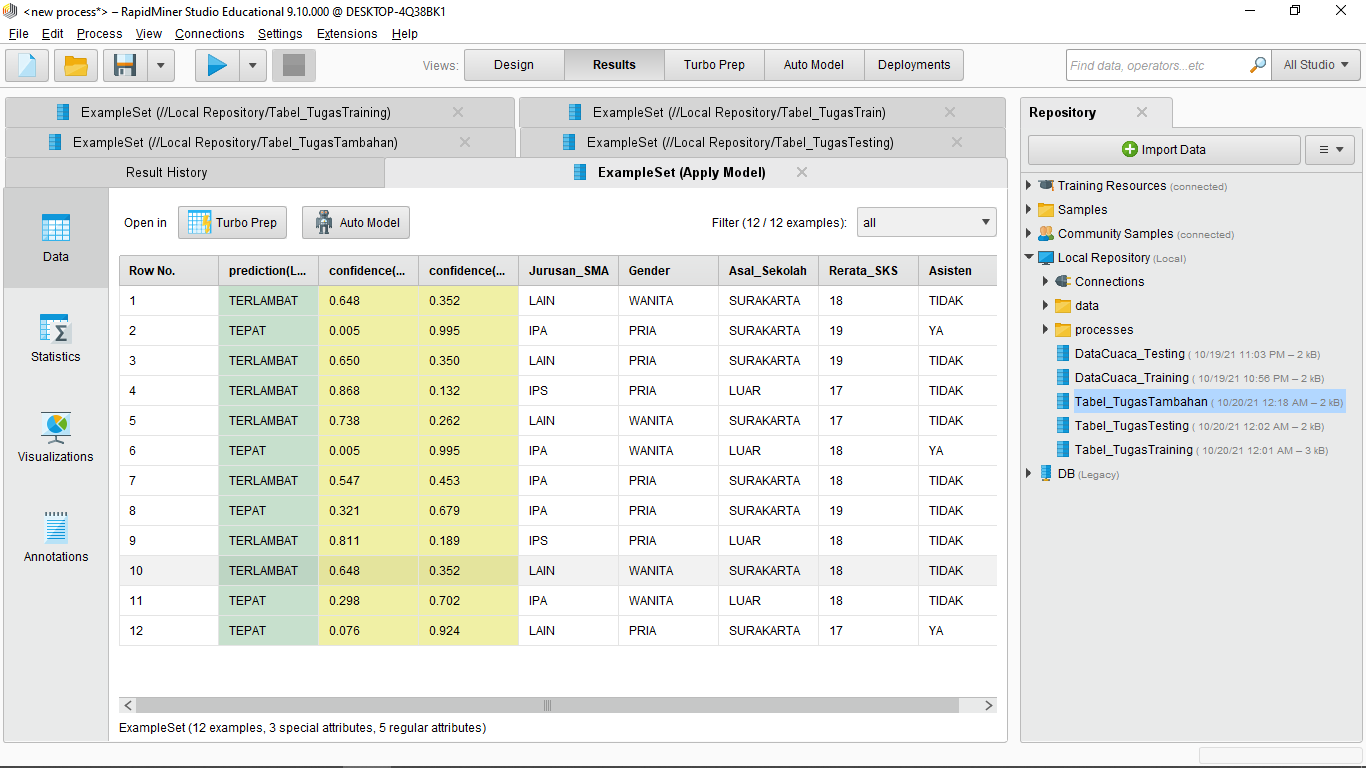
b. Nilai TERLAMBAT : 0,868

8. Jumlah orang yang akan lulus TEPAT dan TERLAMBAT:

a. Tepat : 3 orang

b. Terlambat : 7 orang  
  
9. Menambahkan kondisi pada data Testing yaitu Dewi dan Jono.  


10. Hapus Testing dengan data kondisi baru.  


11. Prediksi ketepatan lama studi si Dewi : TEPAT  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
12. Prediksi ketepatan lama studi si Jono : TEPAT  
  
  
13. Statistics  
