**LAPORAN PRAKTIKUM DATA WAREHOUSING DAN DATA**

**MINING**

**PERTEMUAN 6**

**“ALGORITMA KLASIFIKASI: K-NEAREST NEIGHBOR(K-NN)”**



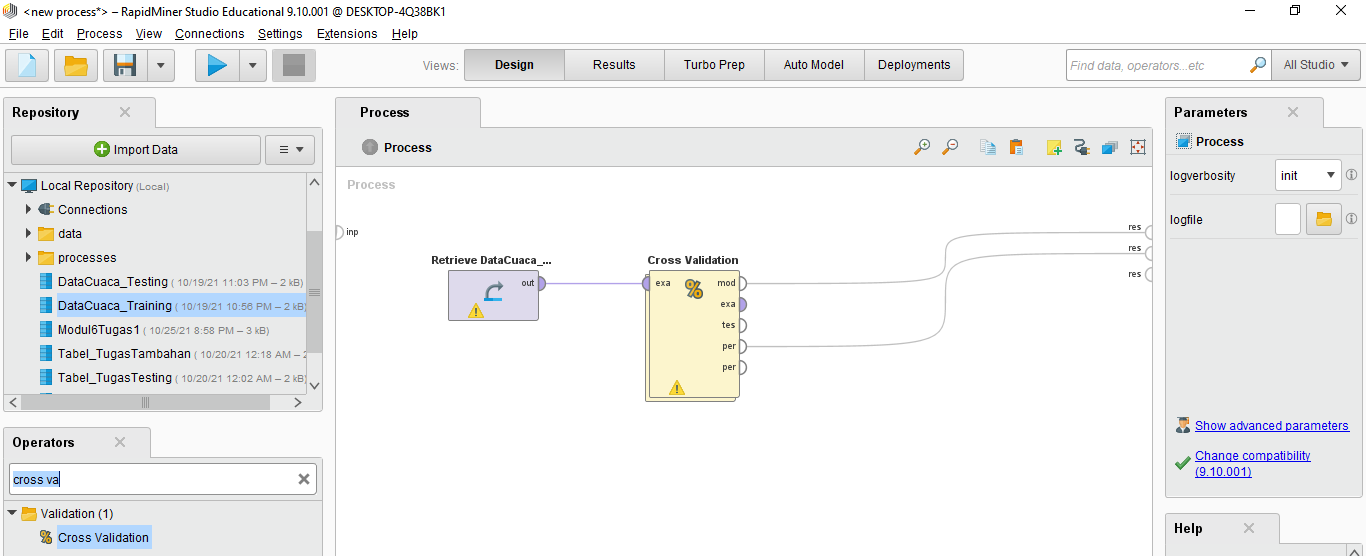
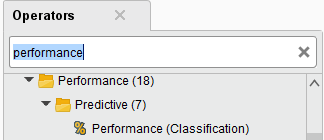
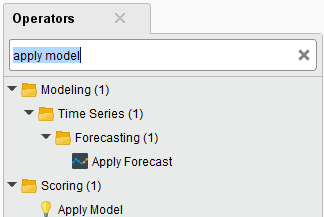
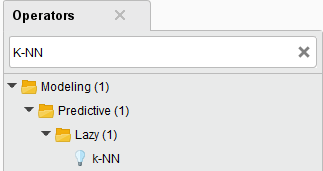
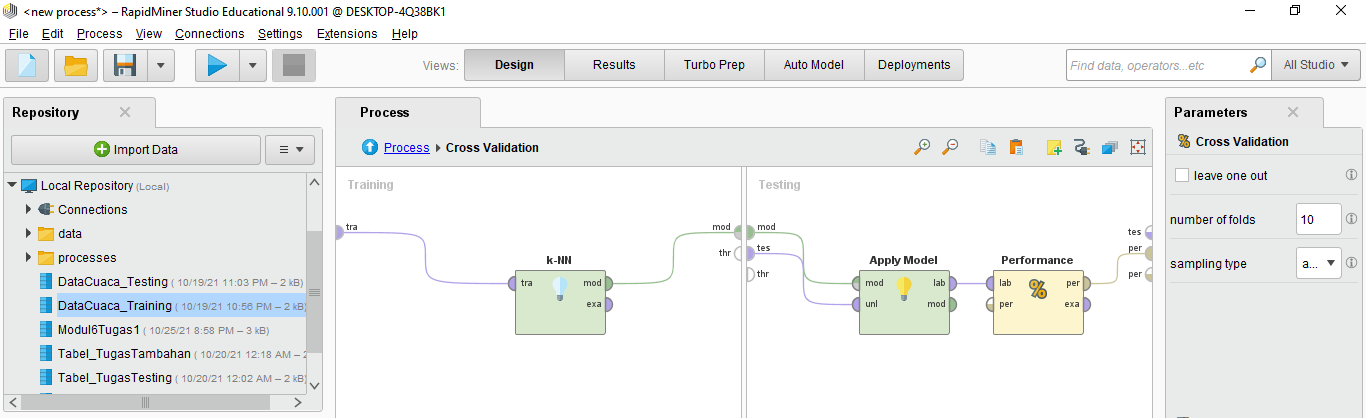
**Oleh:**

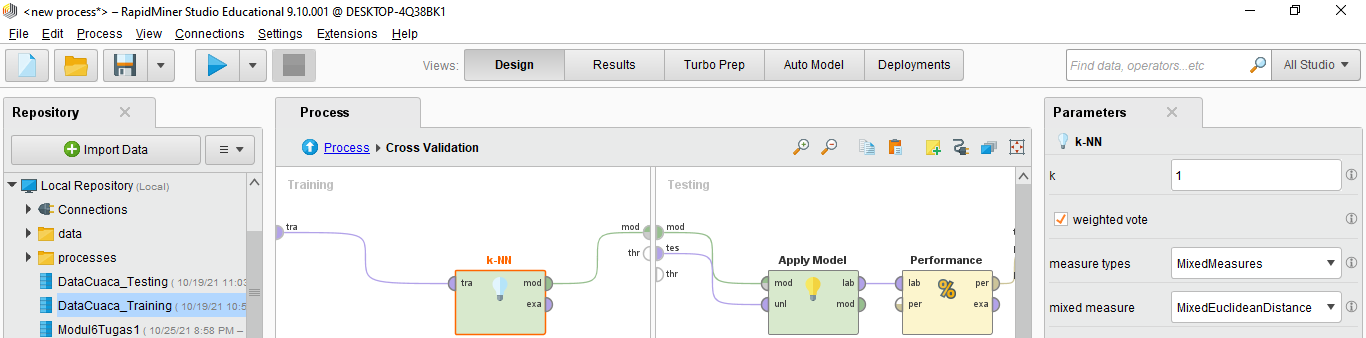
**NAMA : Daffa Putra Alwansyah   
 NIM : L200190031**

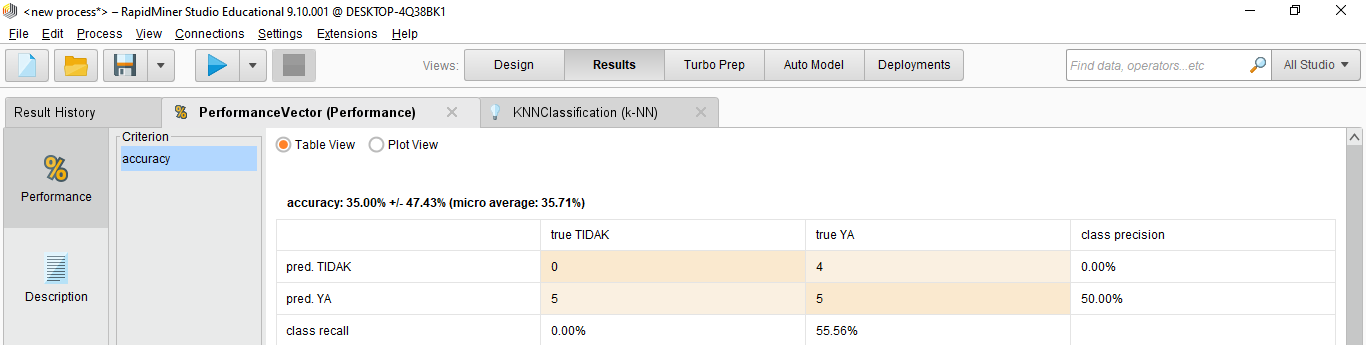
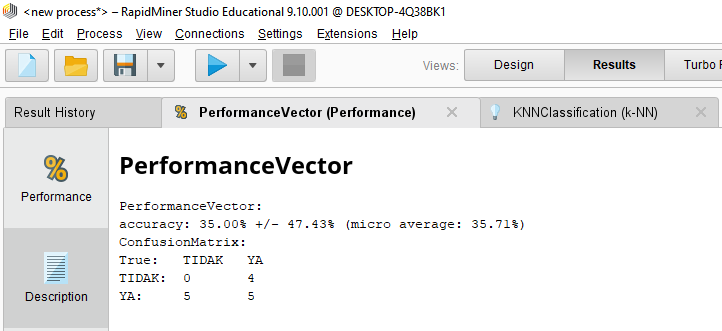
**KELAS : B**

**PRODI : INFORMATIKA**

**Fakultas Komunikasi dan Informatika   
 Universitas Muhammadiyah Surakarta**

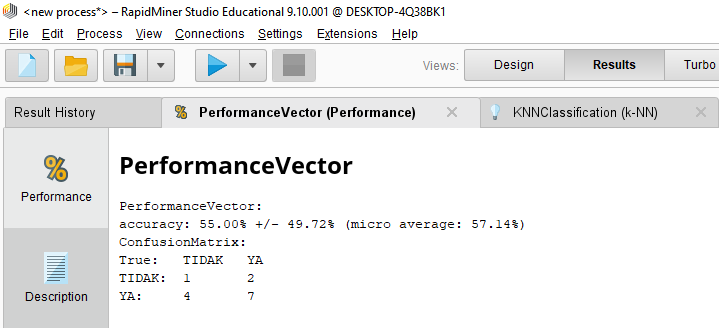
**Kegiatan 4.1 (Menentukan Nilai k dan Tingkat Akurasi Algoritma K-NN)**  
1. Menggunakan file DataCuaca\_Training dan operator Cross Validation, lalu menghubungkan port sesuai dimodul.  
  
  
2. Klik 2x pada Cross Validation, masukan ‘K-NN’, ‘Apply Model’ dan ‘Performance’, lalu sambungkan port sesuai modul.  
  


3. Ubah parameter pada k-NN menjadi 1, setelah itu klik process dan klik RUN.  


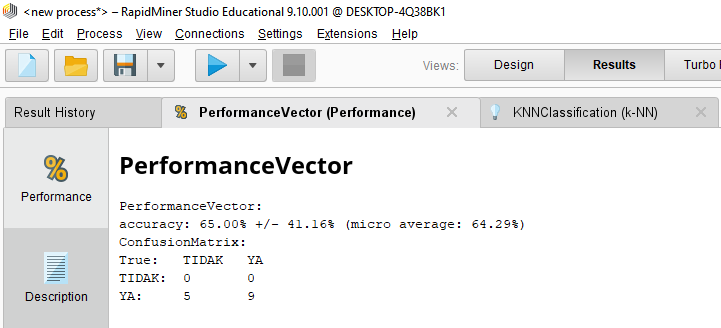
4. Hasil proses klasifikasi dengan algoritma K-NN pada bagian: PerformanceVector (Performance).  
  
 Mode Text View  


* Tingkat akrasi pada percobaan ini mencapai 35% yang menunjukkan angka

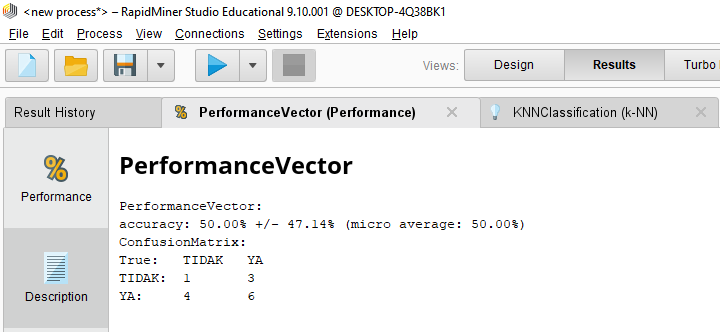
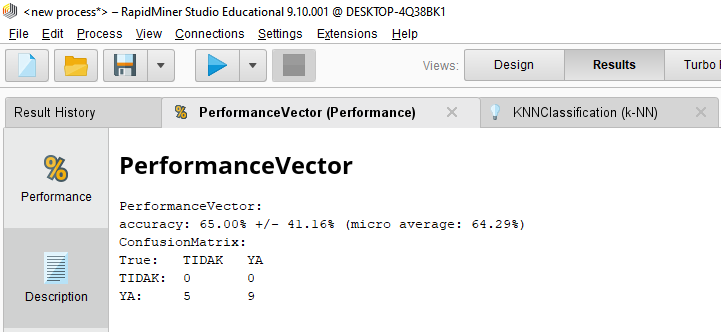
keakuratannya yang sangat rendah.

5. Dikarenakan tingkat akurasinya rendah, maka ubah parameter k pada K-NN menjadi angka 6.  


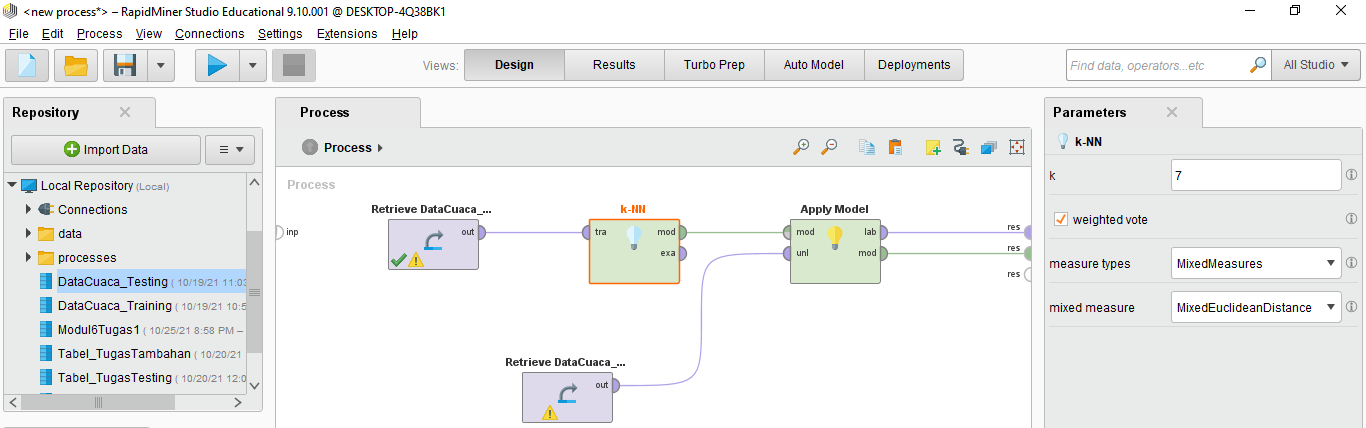
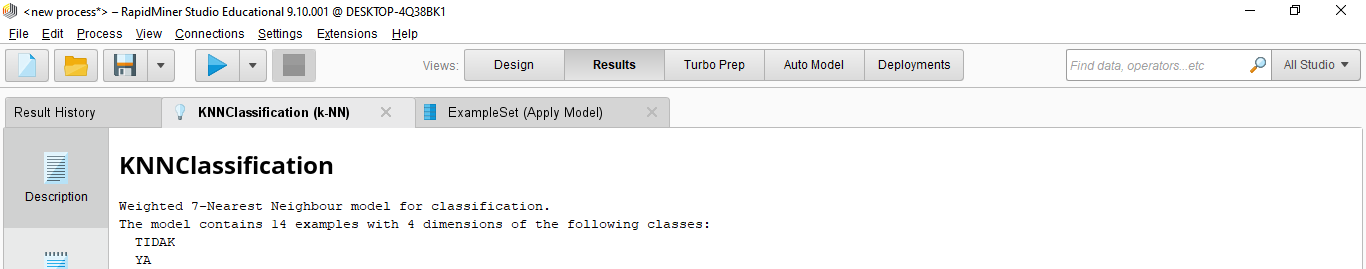
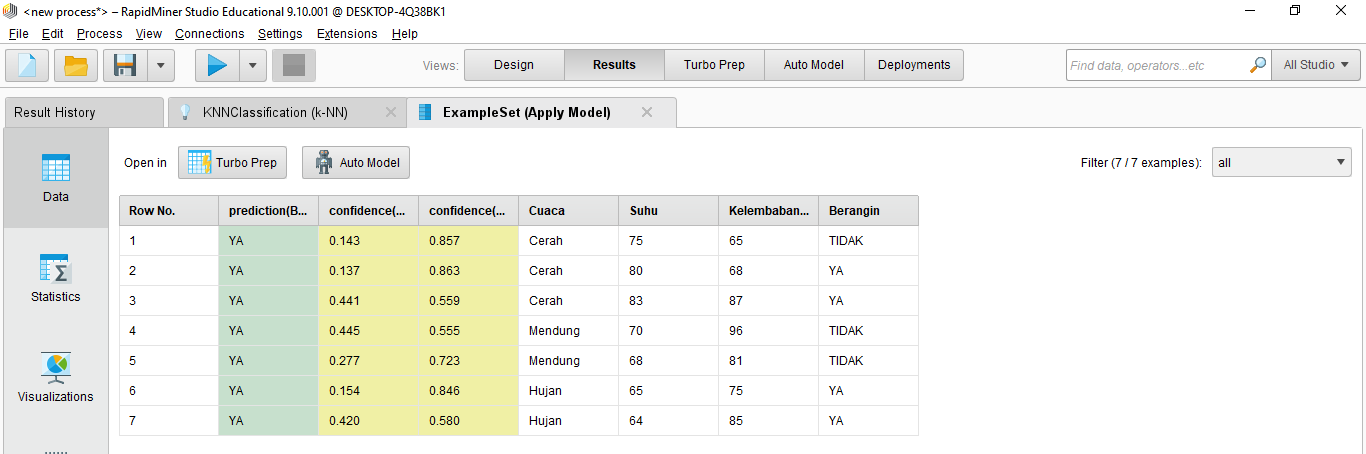
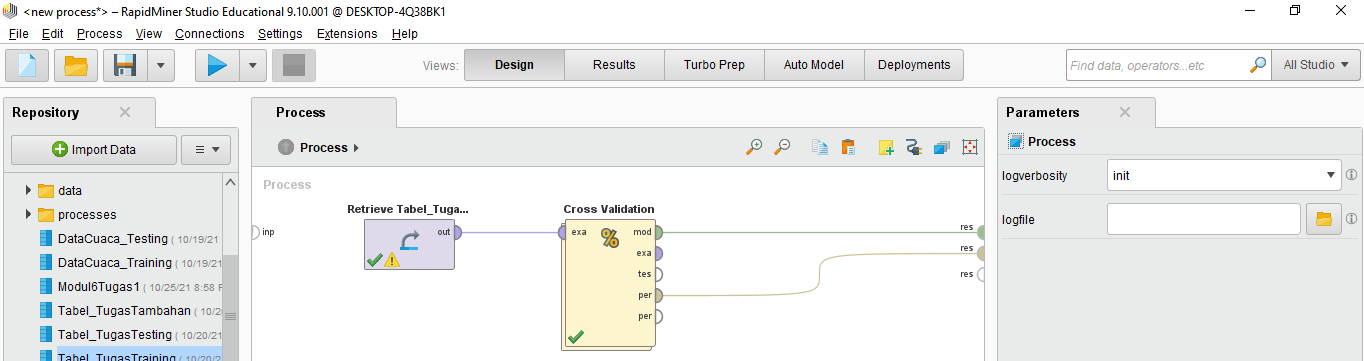
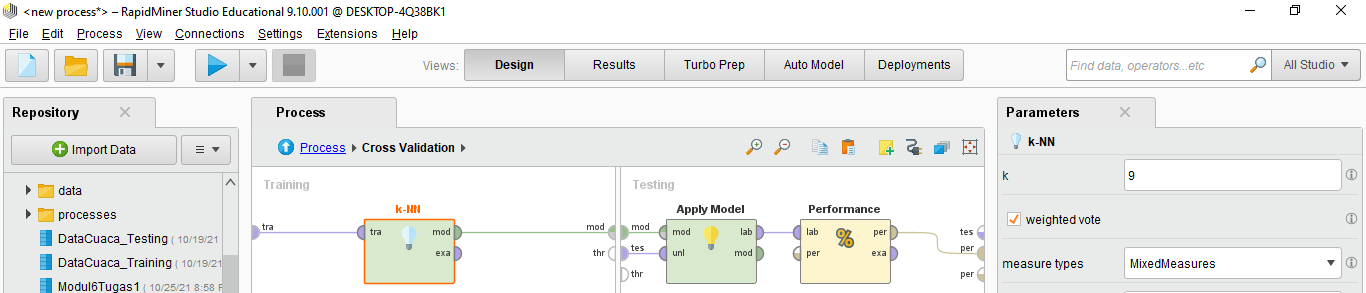
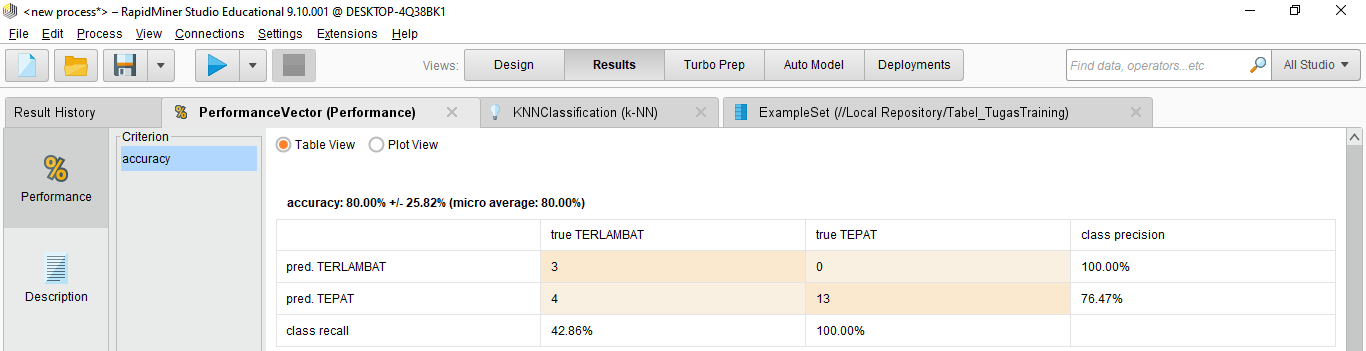
* Hasil tingkat akurasi yang diperoleh sebesar 55% dengan Mode Text View

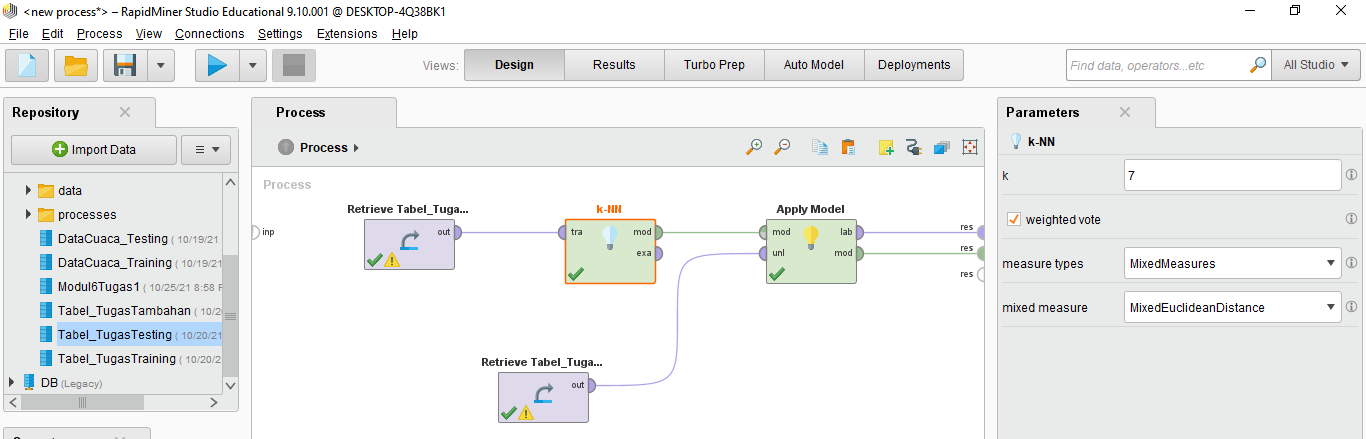
6. Ubah nilai K dengan angka sembarang dari 1 - 14 (saya menggunakan 7).   


* Ternyata, tingkat akurasi maksimal yang diperoleh sebesar 65% dan ini terjadi pada saat nilai k = 4, 7, 9, 11, 12, 13 dan 14.

7. Mengubah Kembali nilai k = 4 dan memberikan tanda cek pada pilihan weighted vote.  
  
  
  
  
  
  
  
  
  
  
8. Mengubah Kembali nilai k = 7 dan memberikan tanda cek pada pilihan weighted vote.  


* Nilai k = 4 dengan menambahkan bobot weight vote tingkat akurasi berkurang menjadi 50%, Nilai k = 7 tingkat akurasi nya tetap sama sebesar 65% meskipun diberi bobot weighted vote. Sehingga dapat disimpukan bahwa nilai k yang akan digunakan untuk melakukan prediksi terhadap data testing adalah k = 7.

**KEGIATAN 4.2 (Prediksi dengan Algoritma K-NN)**1. Masukan DataCuaca\_Testing & DataCuca\_Training, lalu masukan operator K- NN beserta Apply Model, setelah itu ubah nilai k = 7 pada K-NN. Hubungkan port sesuai modul, klik process dan tombol RUN.  
2. Berikut hasil prosesnya:  
  
  
  
**TUGAS!**1. Menggunakan DataStudi\_Training (nama di saya beda) pada respository sebagai data training, Kemudian mencari nilai k yang memiliki tingkat akurasi tertinggi dengan menggunakan algortima k-NN dari data training tersebut.  
  
  
Hasil:  


* Memasukan nilai k = 9 dan memberi weighted vote menghasilkan akurasi sebesar 80 %, angka 1-14 yang merupakan akurasi tertinggi adalah angka 9 dengan menghasilkan akurasi sebesar 80%.  
    
    
    
  2. Menggunakan file DataStudi\_Testing pada respository sebagai data testing, Lakukan prediksi data testing tesebut menggunakan algoritma k-NN.  
    
  Hasil:  
  