**LAPORAN TUGAS KECIL 2**

**Eksplorasi WEKA**

**IF3170 Inteligensi Buatan**

****

**Dibuat oleh:**

**Harry Octavianus Purba 13514050**

**Albertus Kelvin 13514100**

**SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA**

**TEKNIK INFORMATIKA**

**INSTITUT TEKNOLOGI BANDUNG**

**2016**

**Hasil Pengujian Implementasi Kelas**

**Dataset: iris.arff**

Sebuah *class* Java yang bernama **TaskTucil02** mengimplementasikan beberapa API WEKA untuk menjalankan fungsi-fungsi yang dimiliki oleh *WEKA Explorer*. Salah satu fungsi yang akan diimplementasikan adalah pengklasifikasian kelas dari dataset iris.arff. Selanjutnya, proses klasifikasi ini bersifat *supervised learning*, karena data train memiliki label kelas yang sudah terisi sebelumnya.

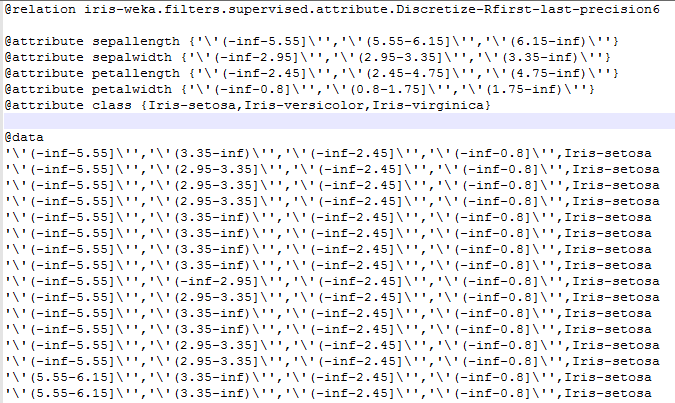
Berikut merupakan hasil implementasi kelas:

1. **Membaca dataset yang diberikan**

Pembacaan dataset merupakan langkah awal yang dilakukan program untuk menyimpan data dari file ARFF ke sebuah objek *instances*. Selain itu, pada langkah ini juga dilakukan penetapan atribut yang akan menjadi atribut kelas (yang akan diprediksi).

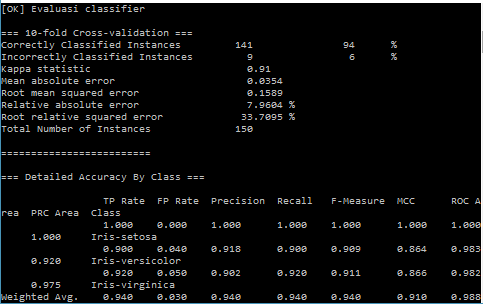
1. **Mengaplikasikan filter Discretize**

Pada tahap ini dilakukan pemfilteran data train dengan menggunakan filter Discretize untuk *supervised learning*. Filter ini digunakan agar memungkinkan digunakannya *classifier* yang menuntut tipe data nominal. Berikut adalah contohnya:



1. **Melakukan pembelajaran dataset dengan skema 10-fold cross validation**

Proses ini akan melatih dan membentuk suatu model dengan menggunakan *classifier* Naïve Bayes. Proses pembelajaran akan dilakukan sebanyak 10 kali dan akan diambil model hasil rata-rata dari setiap model yang terbentuk saat pembelajaran.



Selain itu, juga akan dilakukan evaluasi model yang terbentuk dengan data test yang diambil secara parsial dari dataset untuk diambil rata-rata persentase keberhasilannya.

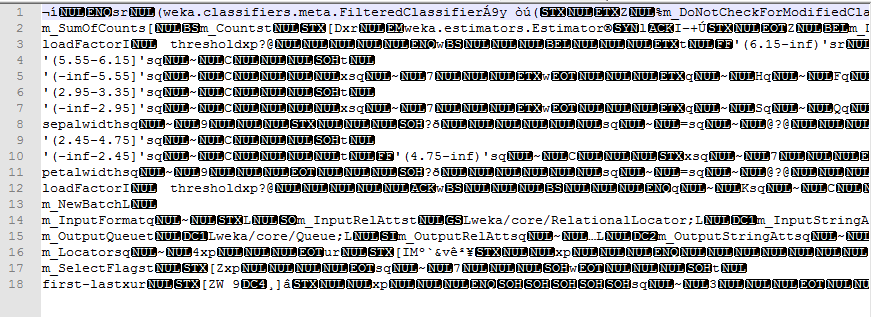
1. **Melakukan pembelajaran dataset dengan skema full training**

Proses ini akan melatih dan membentuk suatu model dengan menggunakan *classifier* Naïve Bayes. Proses pembelajaran akan dilakukan terhadap satu file dataset secara utuh, sehingga tidak ada pembagian dataset menjadi data train dan data test.

1. **Menyimpan model ke dalam file eksternal**

Proses pembelajaran akan menghasilkan suatu objek model. Objek model tersebut kemudian disimpan ke dalam sebuah file eksternal untuk digunakan kemudian saat melakukan proses klasifikasi *instance* baru.

Berikut hasil dari model yang disimpan dalam file eksternal:

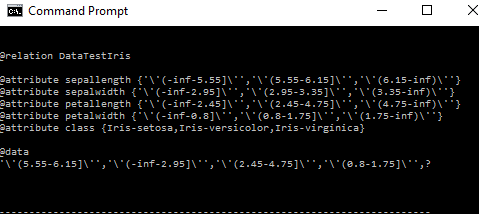


1. **Membaca model dari file eksternal**

Proses ini dilakukan pada tahap klasifikasi *instance* baru. Model yang sudah terbentuk oleh proses *learning* sebelumnya akan digunakan sebagai acuan dalam mencari hasil klasifikasi terbaik dari *instance* baru tersebut.

1. **Membuat *instance* baru dari pengguna untuk setiap atribut**

Proses ini dilakukan untuk membuat sebuah dataset baru yang bersifat sebagai data test, dimana isi data ini hanya berupa *instance* baru yang dimasukkan pengguna. Format dari dataset baru ini harus sama dengan data train sebelumnya, mulai dari nama relasi, nama dan jumlah atribut, tipe data atribut, dan susunan data. Tahap pertama yang dilakukan adalah membentuk nilai atribut yang akan diprediksi (*class*) dan akan disimpan sementara dalam sebuah *ArrayList*. Tahap kedua adalah membentuk objek atribut yang terdiri dari lima buah, yaitu *class* (yang akan diprediksi) dan atribut lainnya (sepallength, sepalwidth, dll). Tahap ketiga adalah menyimpan objek atribut yang sudah terbentuk ke dalam *ArrayList* untuk sementara. Tahap keempat adalah membentuk objek yang berupa dataset baru dengan nama relasi "DataTestIris". Masih di tahap yang sama, kemudian program menginisiasi atribut mana yang akan diprediksi (atribut terakhir) dan menambahkan input pengguna dari file input ke dalam atribut lainnya (menggunakan *DenseInstance*). Hasil dari objek *DenseInstance* akan dimasukkan ke dalam objek yang berupa *instances*. Berikut hasilnya:



1. **Melakukan klasifikasi dengan memanfaatkan model yang telah didapat**

Pada proses ini akan dilakukan klasifikasi kelas bagi *instance* baru yang dimasukkan pengguna. Proses klasifikasi akan memanfaatkan model yang telah dibaca dari file eksternal sebelumnya. Data test yang akan digunakan adalah *instances* baru yang dibentuk pada langkah sebelumnya. Selain itu, pada tahap ini juga akan ditentukan nilai probabilitas dari hasil klasifikasi yang dipilih.

Berikut gambaran hasilnya:

