**LAPORAN PRAKTIKUM**

**WORKSHOP KECERDASAN KOMPUTASIONAL**

**“Naive Bayes Classification”**



**Oleh :**

**Muhammad Rifqi Aminuddin  
NRP. 3123640039**

**PROGRAM STUDI STrLJ TEKNIK INFORMATIKA**

**DEPARTEMEN TEKNIK INFORMATIKA DAN KOMPUTER**

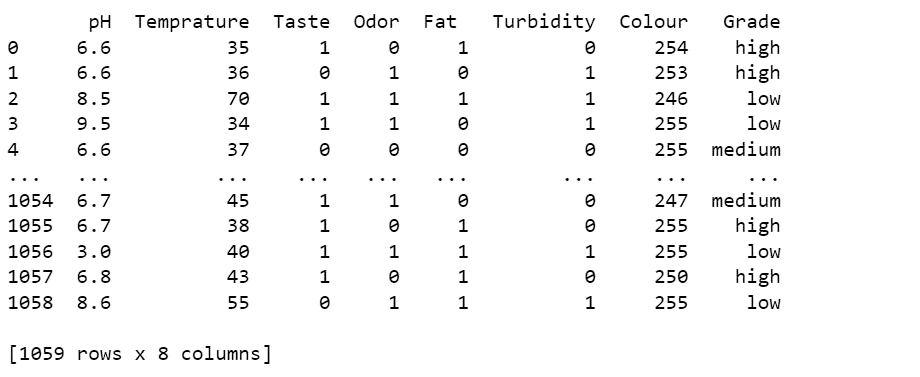
**POLITEKNIK ELEKTRONIKA NEGERI SURABAYA**

1. dataset 🡨 milk.csv

* Kode



* Keluaran

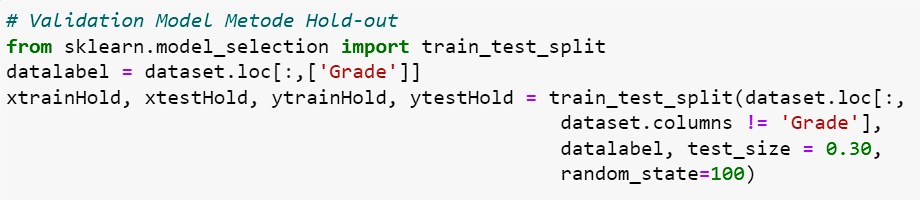


* Analisa

Kode di atas digunakan untuk menampilkan dataset dari file dataset berupa milk.csv, yang mana berisi data data susu mulai dari pH, suhu, rasa, indeks lemak, dll.

1. Lakukan validation Model dengan metode:
2. Hold-out Method (70%-30%)

* Kode



* Keluaran

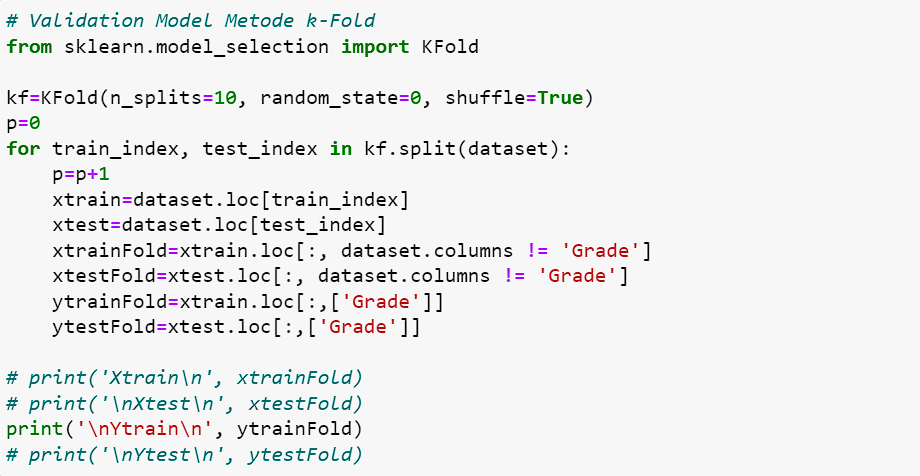


* Analisa

Kode di atas merupakan kode untuk melakukan validasi dengan metode Hold-out dengan (70%-30%). Metode ini diperlukan untuk melakukan dan memisahkan antara data\_test dengan data\_train secara otomatis berdasar algoritma masing masing metode.

1. K-Fold (k=10)

* Kode



* Keluaran

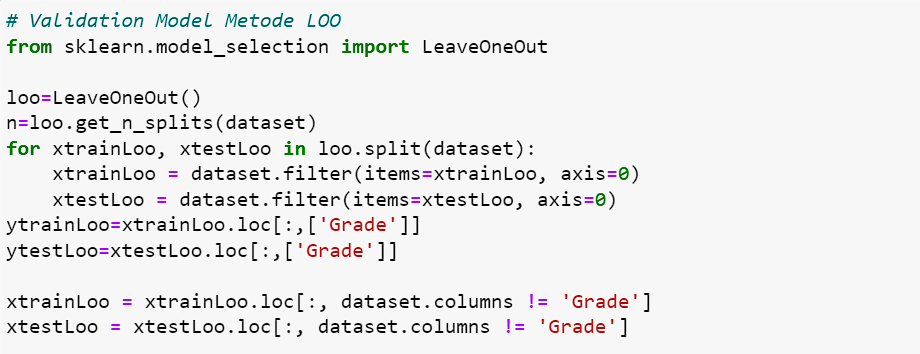


* Analisa

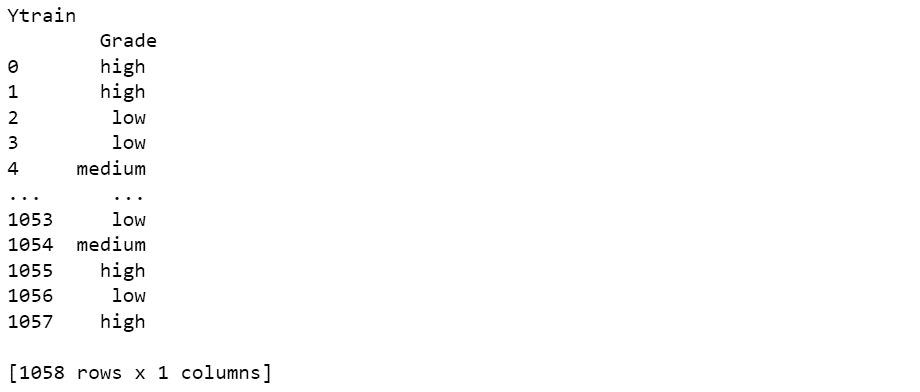
Kode di atas merupakan kode untuk melakukan validasi dengan metode k-Fold dengan k=10. Metode ini diperlukan untuk melakukan dan memisahkan antara data\_test dengan data\_train secara otomatis berdasar algoritma masing masing metode.

1. LOO

* Kode



* Keluaran

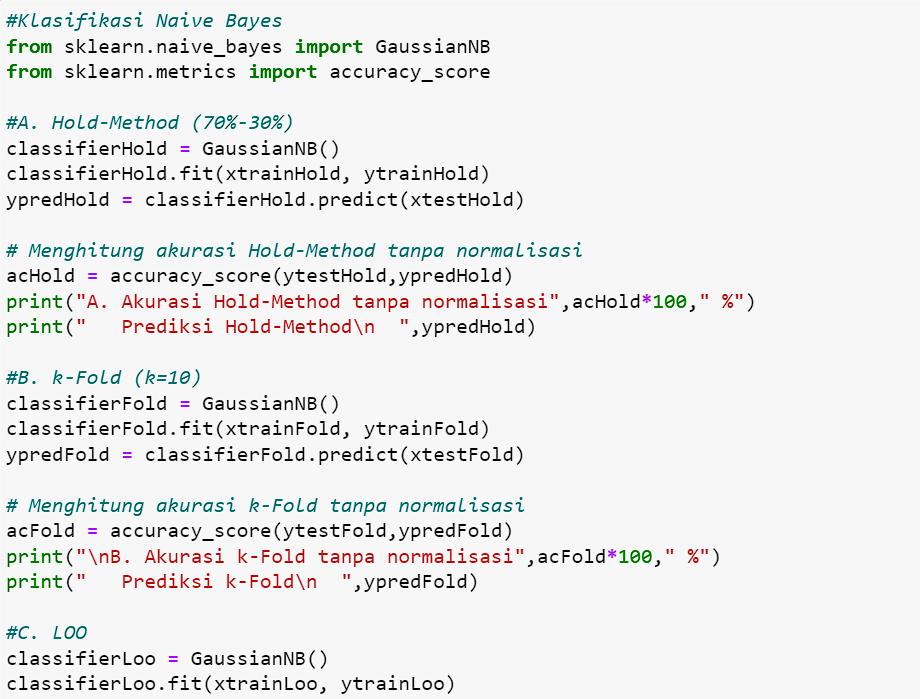


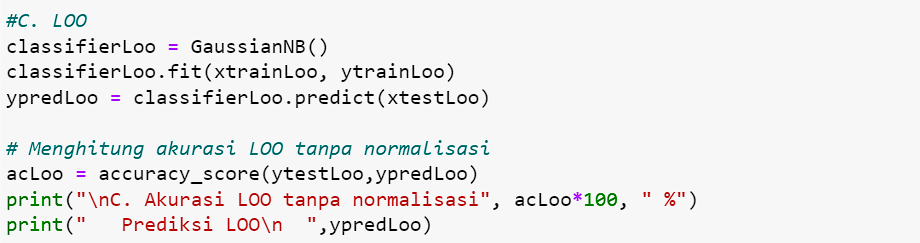
* Analisa

Kode di atas merupakan kode untuk melakukan validasi dengan metode LOO, yang mana metode mengeluarkan satu data\_test dan ini diperlukan untuk melakukan dan memisahkan antara data\_test dengan data\_train secara otomatis berdasar algoritma masing masing metode.

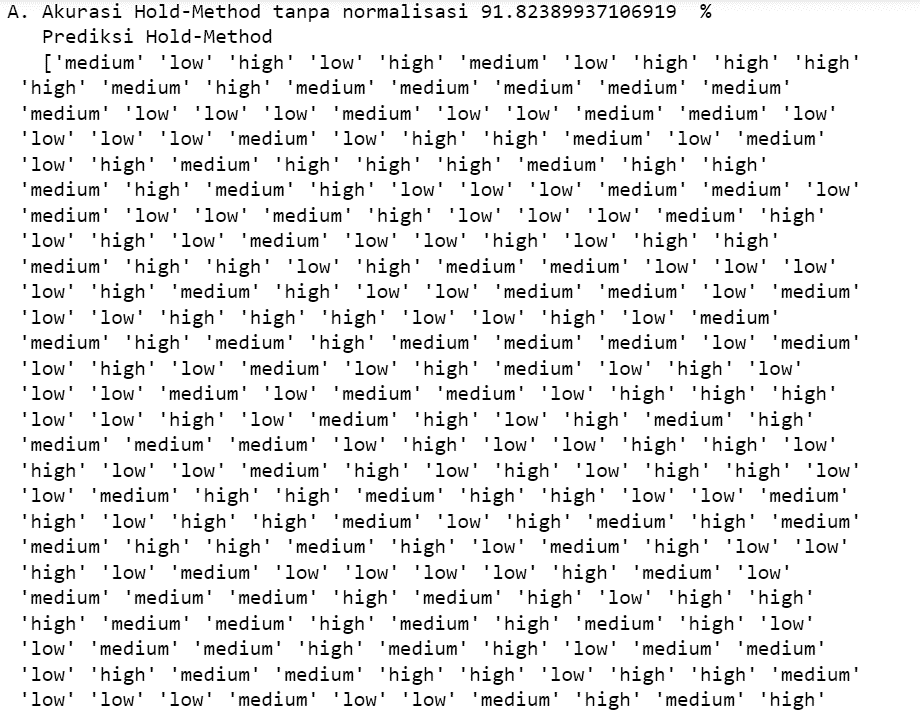
1. Lakukan klasifikasi Naïve Bayes untuk masing-masing pendekatan validasi dan hitunglah akurasi untuk masing-masing metode validation

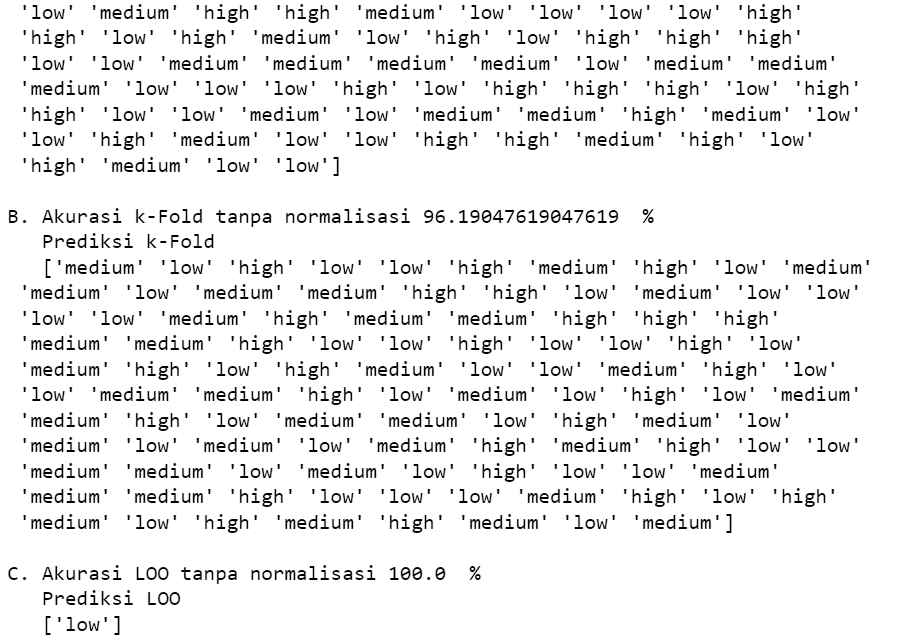
* Kode

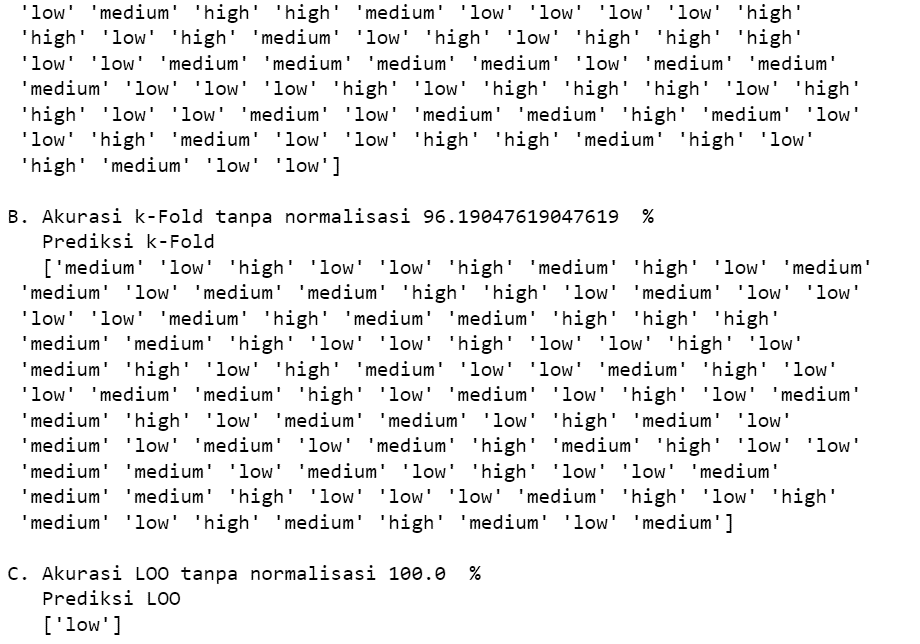




* Keluaran





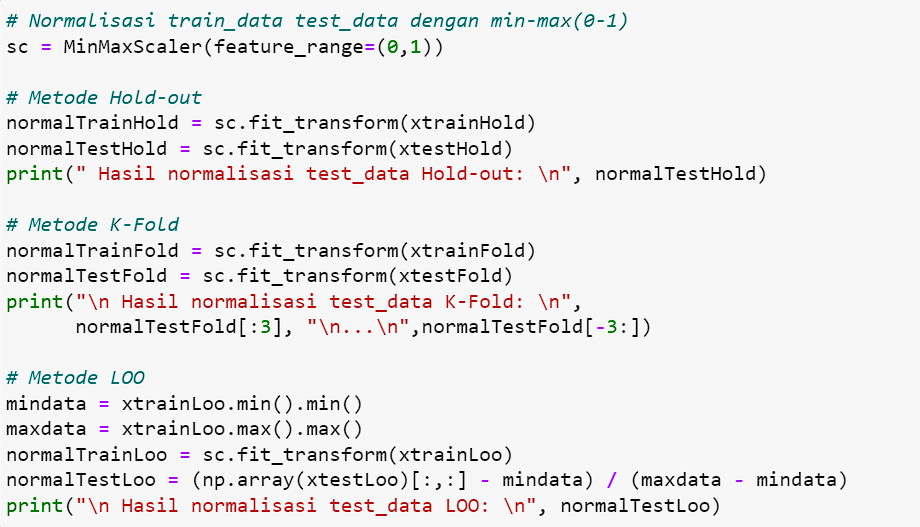


* Analisa

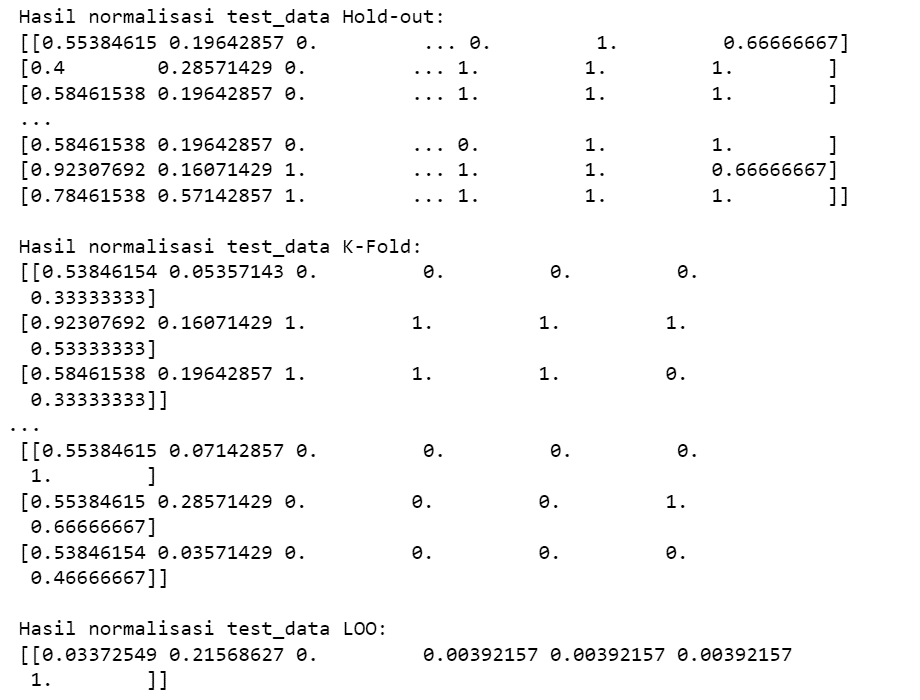
Kode di atas digunakan untuk melakukan klasifikasi naïve bayes, metode ini merupakan metode klasifikasi yang biasanya digunakan untuk memprediksi hasil keluaran dari model yang telah dilatih dengan data\_training. Pada praktiknya model akan dilatih menggunakan data dari data\_training dengan jumlah data yang beragam yaitu tergantung pada metode validasi apa yang akan digunakan. Pada studi kasus di atas digunakanlah metode validasi Hold-out, k-Fold, dan LOO, yang mana ketiganya memiliki standar dan aturan sendiri dalam memisahkan antara data\_training dengan data\_test. Apabila model telah dilatih dengan semua data yang ada, maka ia akan diujicobakan untuk diprediksi hasil keluaran seperti apa yang akan keluar apabila diberikan nilai tertentu. Sehingga dari proses di atas diperoleh beberapa hasil akurasi yang diantaranya Hold-Out dengan 91,823%, k-Fold dengan 96,1904% dan LOO dengan hasil yang sempurna yaitu 100%.

1. test\_data 🡨 lakukan normalisasi pada train\_data dengan min-max

* Kode



* Keluaran

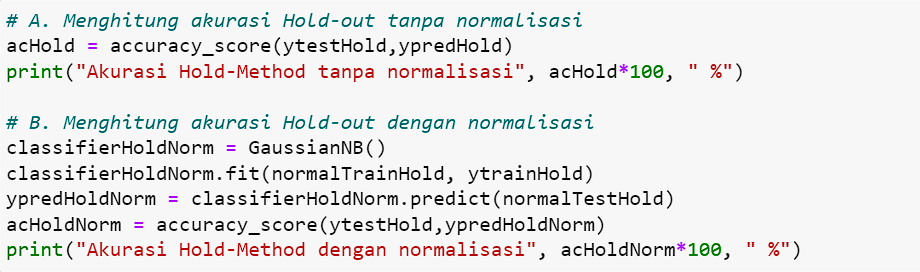


* Analisa

Langkah di atas merupakan langkah dalam melakukan normalisasi terhadap data\_training dan data\_testing untuk masing-masing metode validasi. Tahapan normalisasi ini dirasa perlu dilakukan karena untuk menghindari bias pada data yang akan diolah, sehingga hasil keluaran yang diperoleh dapat maksimal dan tidak terganggu dengan adanya data bias tersebut. Selain itu normalisasi juga digunakan untuk menghilangkan jarak antara data maksimal dengan data minimal yang mana terpaut cukup jauh. Oleh karenanya pada tahapan normalisasi ini data yang tersedia diolah menjadi dalam rentang 0 hingga 1.

1. Bandingkan nilai akurasi klasifikasi dengan Naïve Bayes pada salah satu metode validasi jika data training & data test dilakukan normalisasi & tidak dinormalisasi

* Kode



* Keluaran



* Analisa

Langkah di atas digunakan untuk mendeteksi hasil akurasi antara klasifikasi Naïve Bayes dengan normalisasi dan tanpa normalisasi. Dari percobaan di atas, diperoleh hasil yang mana tingkat akurasi klasifikasi Naïve Bayes dan validasi Hold-out dengan normalisasi maupun normalisasi ternyata ialah sama, yaitu di angka 91,823%. Hal tersebut biasa terjadi apabila data yang digunakan sudah stabil ataupun optimal yaitu terhindar dari bias maupun residu residu data lainnya yang membuat perhitungan data dari model kurang optimal dan berakibat penurunan tingkat akurasi dalam mendeteksi data\_test.