LAPORAN MACHINE LEARNING

PRAKTIKUM 2



OLEH:

NUGROHO ARIFIANTO 24060117140076

DEPARTEMEN ILMU KOMPUTER

FAKULTAS SAINS DAN MATEMATIKA

UNIVERSITAS DIPONEGORO

2018/2019

Bab 1

Pendahuluan

* 1. Latar Belakang

Pada zaman ini data bisa diperoleh dengan mudah dan cepat. Data-data tersebut dikumpulkan dan digunakan untuk pengambilan kesimpulan atau untuk melakukan klasifikasi dari data tersebut. Pengolahan data menjadi hal yang penting saat ini, adanya kemudahan memperoleh data membuat pengolahan data yang menjadi lebih mudah dan cepat.

Selain data itu diolah data tersebut juga harus dievaluasi untuk memperkirakan akurasi dari dataset yang kita dapat.

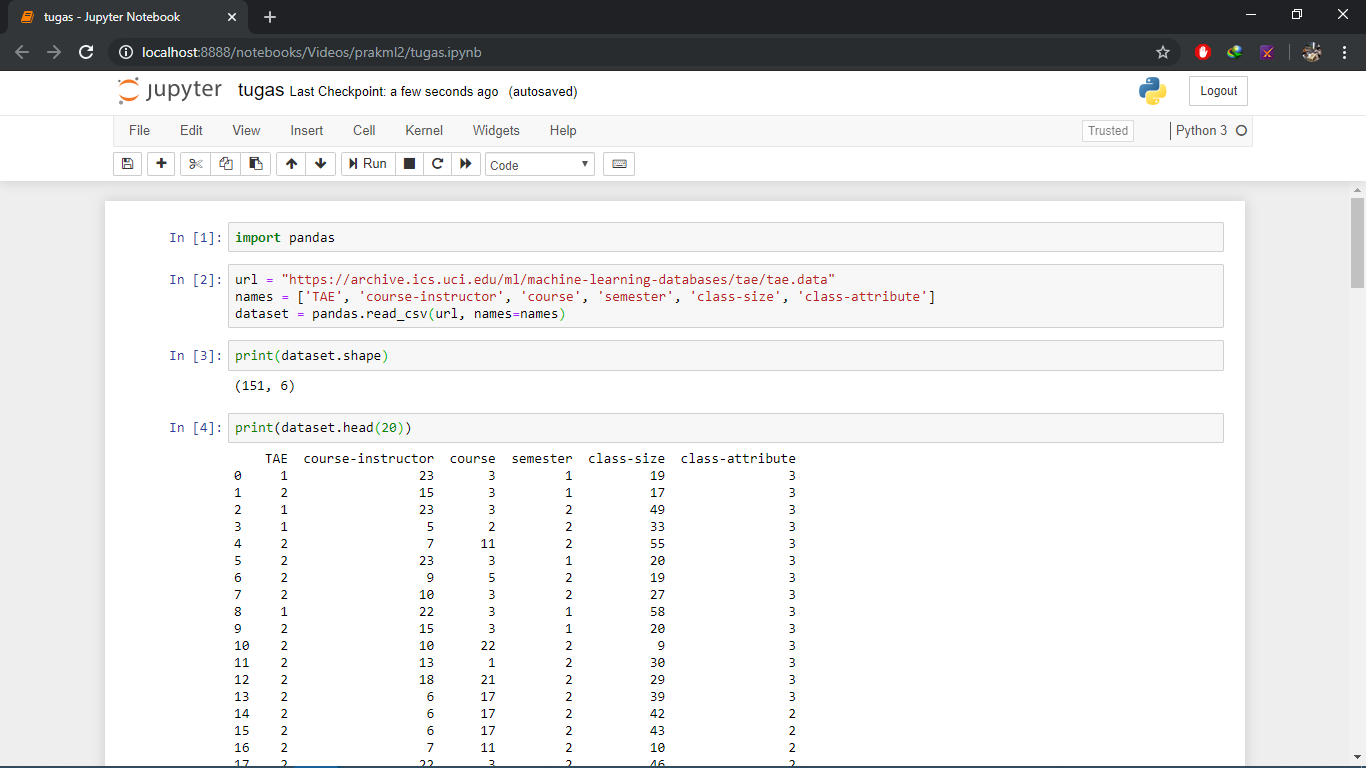
* 1. Rumusan Masalah
     1. Bagaimana hasil evaluasi dari dataset yang dipunya?
  2. Tujuan
     1. Mahasiswa dapat melakukan evaluasi terhadap data yang dipunya menggunakan python.

Bab 2

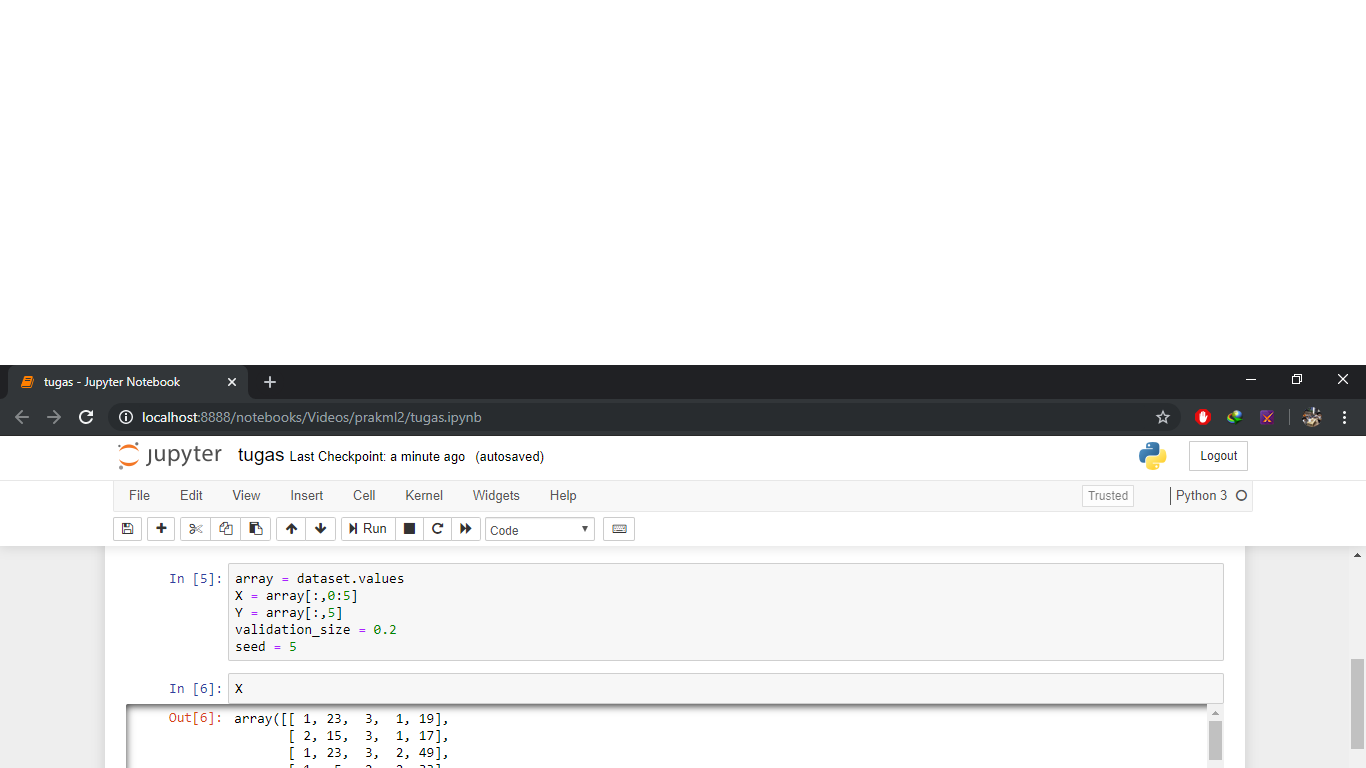
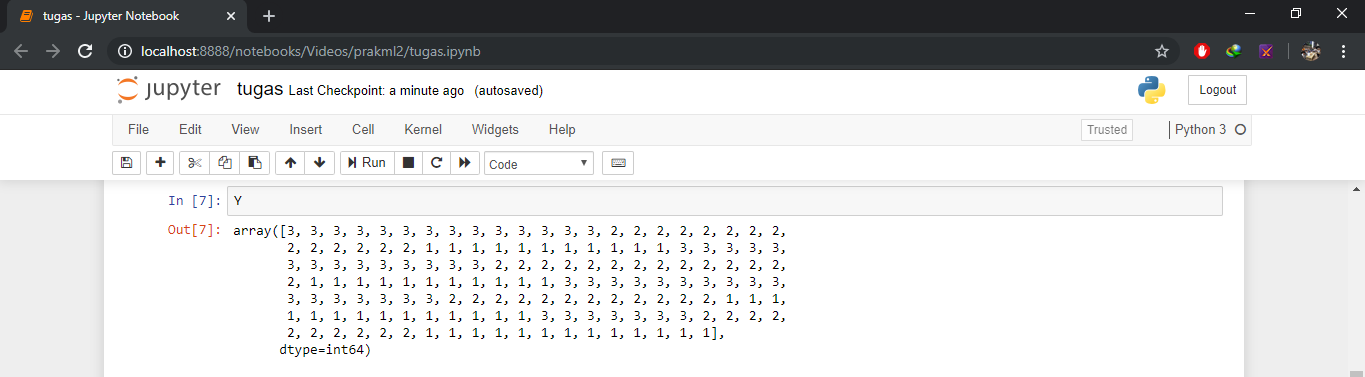
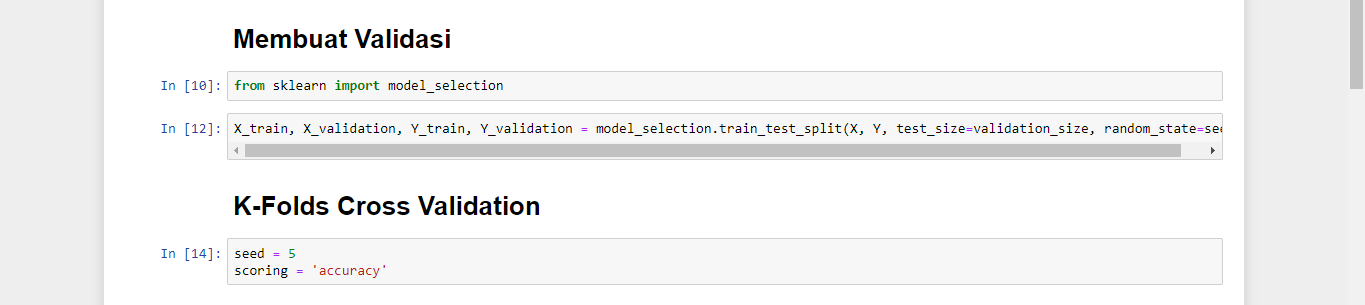
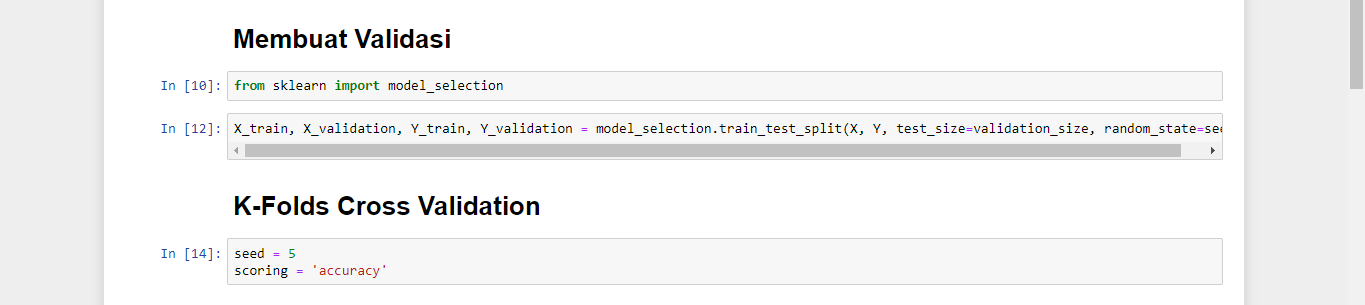
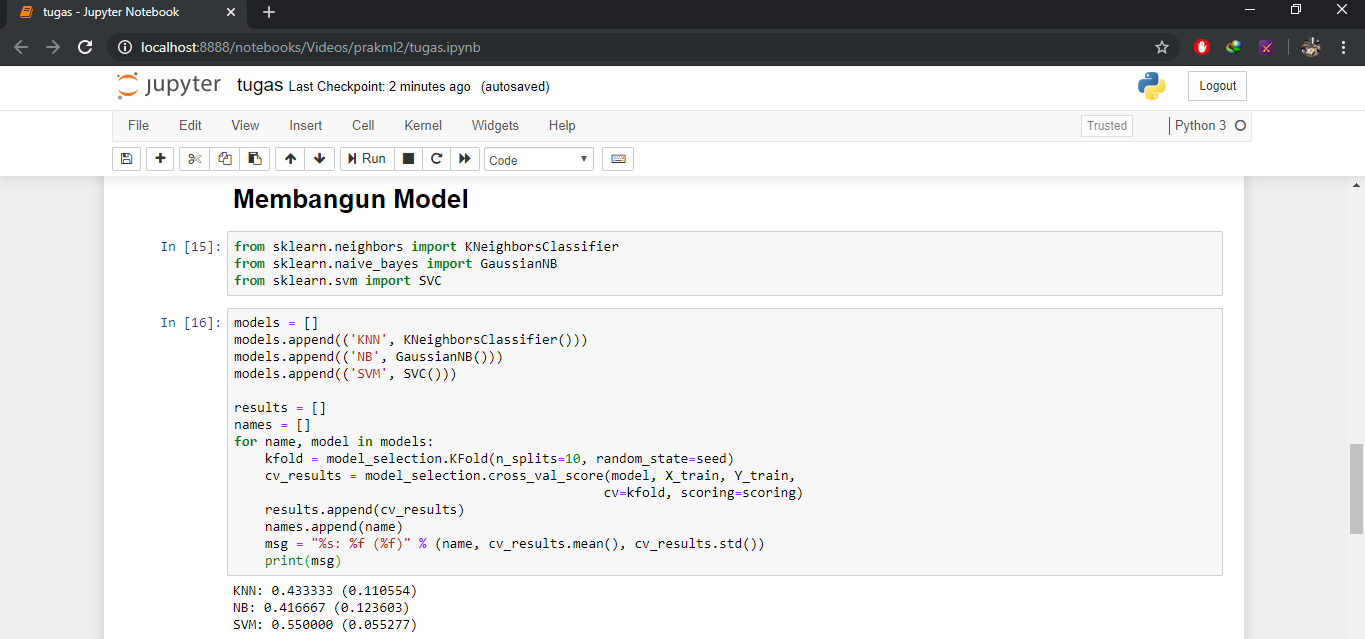
Pembahasan

2.1 Data

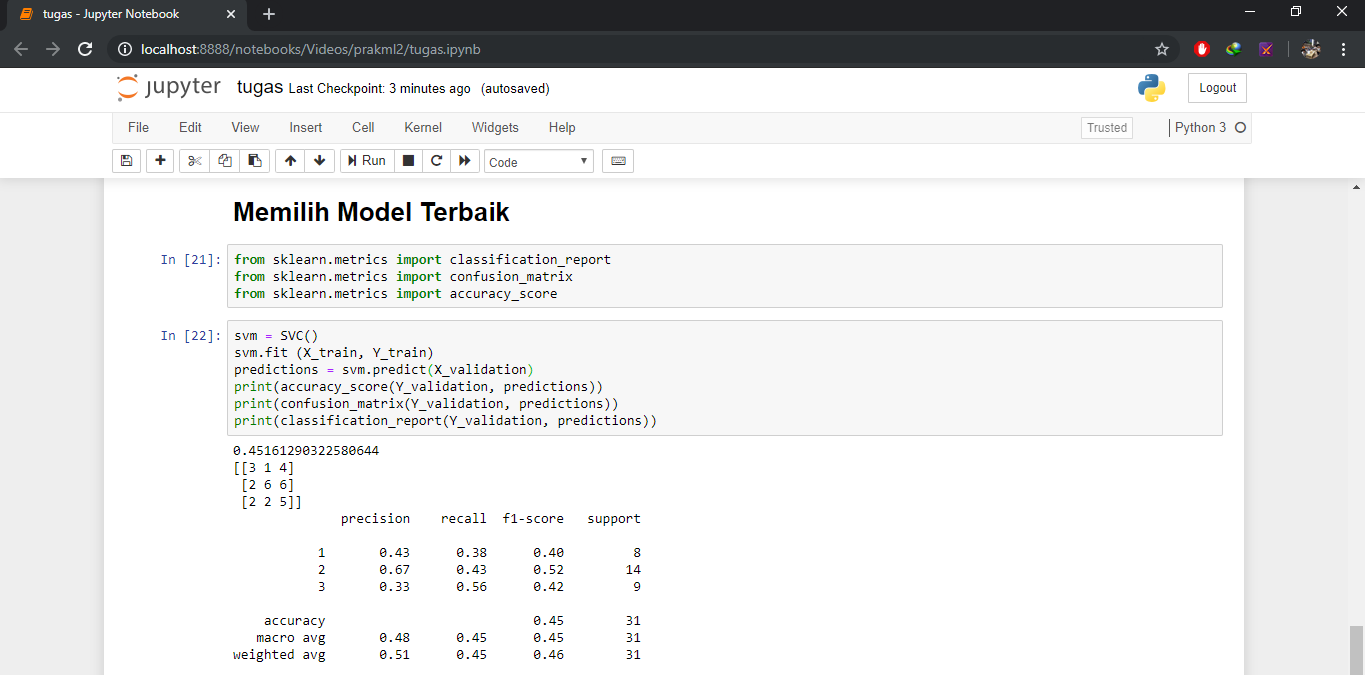
Dataset yang saya gunakan adalah dataset Teaching Assistant Evaluation. Dataset ini berisi evaluasi tentang performa mengajar selama tiga regular semester dan dua summer semester dari 151 Teaching Assistant.

2.2 Pengolahan Dataset

Dataset yang saya ambil ada pada file tae.data dataset tersebut memiliki 6 atribut, yaitu TAE untuk mengetahui apakah Teaching Assistant tersebut adalah English Speaker atau Non-English Speaker dengan nilai 1 untuk English Speaker dan 2 untuk Non-English Speaker, lalu ada atribut course-instructor, course, semester dengan nilai 1 untuk Summer dan nilai 2 untuk Regular Semester, class-size, dan class-attribute.

Setelah itu membuat validasi terhadap dataset dengan X menampilkan data yang memliki nilai dari kolom pertama hingga kolom kelima, dan Y menampilkan data yang memiliki nilai dari kolom keenam, lalu membagi dataset menjadi dua bagian, yaitu data training sebesar 0,8 dan data validasi yang diinisialisasi dengan validation\_size sebesar 0,2.

Setelah itu kita membangun model dengan menggunakan tiga algoritma, yaitu KNN, Naives Bayes, dan SVM untuk mengetahui algoritma apa yang cocok digunakan dan diketahui nilai dari masing masing algoritma yaitu KNN sebesar 0,433333(0,110554), Naives Bayes sebesar 0,416667(0,123603), dan SVM sebesar 0,550000.

Berdasarkan hasil di atas algoritma SVM yang memiliki nilai akurasi terbesar maka dari itu dilakukan pengujian tentang keakuratan model SVM terhadap data yang diuji.

Bab 3

Kesimpulan

Kesimpulan dari pembahasan adalah data yang kita punya setelah diolah dilakukan evaluasi untuk memperkirakan akurasi dari data yang diketahui. Pengecekan akurasi dapat dilakukan menggunakan beberapa algortima, yaitu KNN, Naives Bayes, SVM, dan lain-lain