**PENGENALAN PYTHON**

**PRAKTIKUM PEMBELAJARAN MESIN**



**Disusun Oleh:**

Nashirudin Baqiy 24060119130045

**Lab A2**

**DEPARTEMEN ILMU KOMPUTER / INFORMATIKA**

**FAKULTAS SAINS DAN MATEMATIKA**

**UNIVERSITAS DIPONEGORO**

**2021**

**BAB I PENDAHULUAN**

* 1. **Rumusan Masalah**
     1. Cari dataset dataset lain dari repository UCI Machine Learning!
     2. Lakukan keenam langkah untuk menampilkan summary dataset!
  2. **Tujuan** 
     1. Mencari dan menemukan dataset di repository UCI machine learning
     2. Melakukan enam langkah preprocessing data hingga menampilkan summary dataset.
  3. **Dasar Teori**

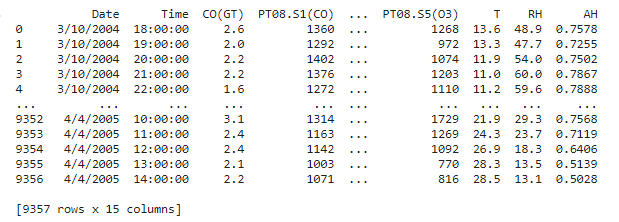
Pembelajaran mesin (ML) adalah studi tentang algoritma komputer yang dapat ditingkatkan secara otomatis melalui pengalaman dan penggunaan data. Itu dilihat sebagai bagian dari kecerdasan buatan. Algoritma pembelajaran mesin membangun model berdasarkan data sampel, yang dikenal sebagai "data pelatihan", untuk membuat prediksi atau keputusan tanpa diprogram secara eksplisit untuk melakukannya. Algoritma pembelajaran mesin digunakan dalam berbagai macam aplikasi, seperti dalam kedokteran, penyaringan email, pengenalan suara, dan visi komputer, di mana sulit atau tidak mungkin untuk mengembangkan algoritma konvensional untuk melakukan tugas-tugas yang diperlukan.

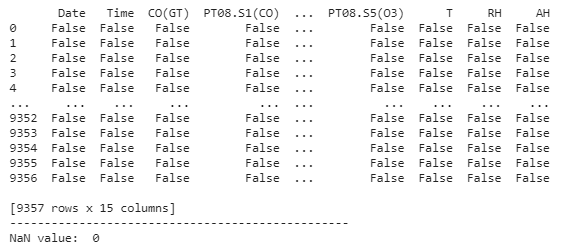
Bagian dari pembelajaran mesin terkait erat dengan statistik komputasi, yang berfokus pada pembuatan prediksi menggunakan komputer; tetapi tidak semua pembelajaran mesin adalah pembelajaran statistik. Studi optimasi matematika memberikan metode, teori, dan domain aplikasi ke bidang pembelajaran mesin. Data mining adalah bidang studi terkait, dengan fokus pada analisis data eksplorasi melalui pembelajaran tanpa pengawasan. Beberapa implementasi pembelajaran mesin menggunakan data dan jaringan saraf dengan cara yang meniru kerja otak biologis. Dalam penerapannya di seluruh masalah bisnis, machine learning juga disebut sebagai analitik prediktif.

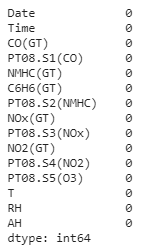
**BAB II PEMBAHASAN**

* 1. **Deskripsi Dataset**

Pada praktikum kali ini, dataset yang akan digunakan adalah dataset dari UCI yang berisi data mengenai prediksi pasar saham dengan 9357 baris dan 15 kolom. File dataset yang digunakan bernama AirQualityUCI.csv

****

****

****

Jika dilihat lebih teliti lagi, maka dataset ini mengandung banyak sekali missing value atau NaN value 0 nilai data jadi tidak menghilangkan beberapa fiturnya.

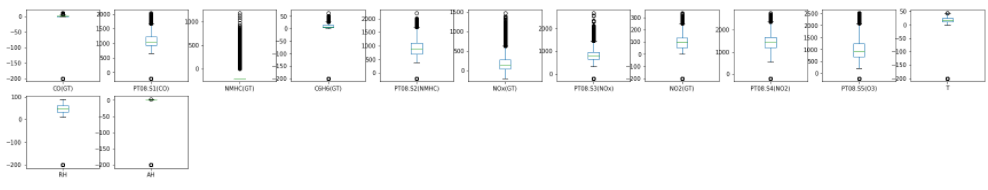
* 1. **Statistik**

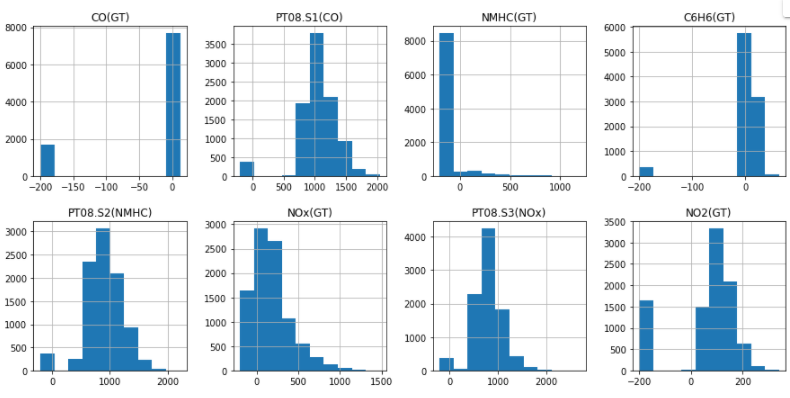
Berikut karakteristik secara statistik dari dataset ini

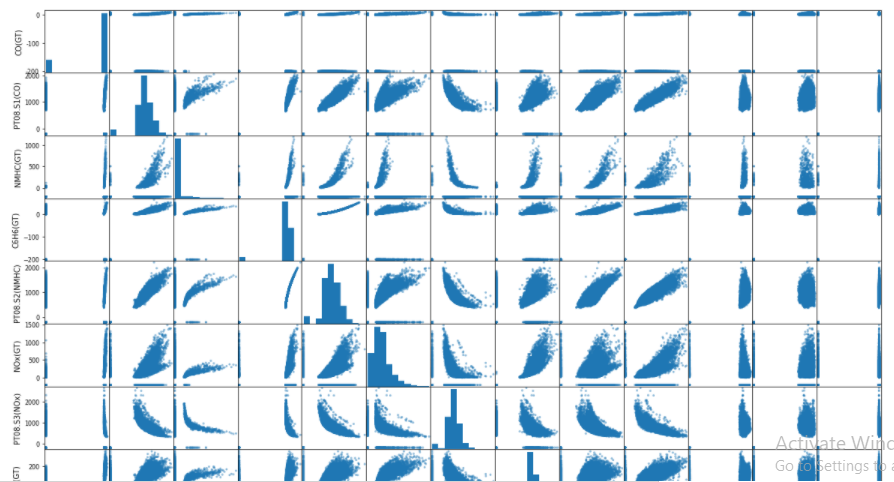
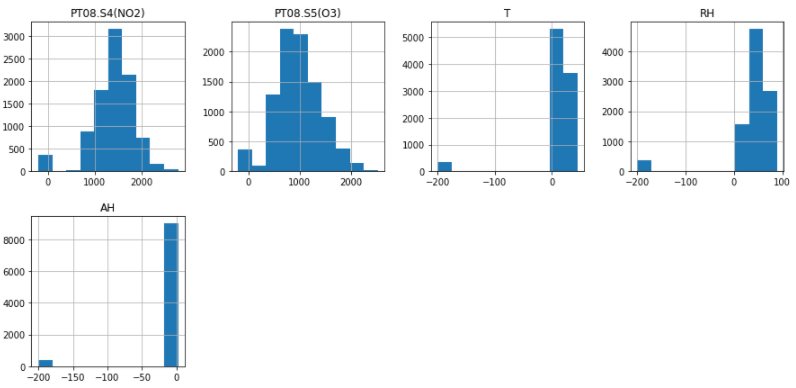
****

* 1. **Visualisasi**

Tahap terakhir dari preprocessing kali ini adalah melakukan plotting grafik dengan univarian dan multivarian







**BAB III KESIMPULAN**

* 1. **Kesimpulan**

Pada praktikum pertemuan pertama di mata kuliah pembelajaran mesin kali ini, mempelajari mengenai dasar Python untuk machine learning khususnya untuk preprocessing dataset. Dalam belajar mengenai machine learning di praktikum kali ini, dipergunakan beberapa library Python guna memudahkan dalam melakukan preprocessing, seperti Pandas dan Numpy untuk mengolah array dan matriks dan Matplotlib sebagai library untuk membuat grafik dari data yang ada. Dataset yang digunakan adalah dataset mengenai kualitas udara prediction (*AirQuality*) yang didapat dari situs web UCI

**DAFTAR PUSTAKA**

S. De Vito, E. Massera, M. Piga, L. Martinotto, G. Di Francia, On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario, Sensors and Actuators B: Chemical, Volume 129, Issue 2, 22 February 2008, Pages 750-757, ISSN 0925-4005

Saverio De Vito, Marco Piga, Luca Martinotto, Girolamo Di Francia, CO, NO2 and NOx urban pollution monitoring with on-field calibrated electronic nose by automatic bayesian regularization, Sensors and Actuators B: Chemical, Volume 143, Issue 1, 4 December 2009, Pages 182-191, ISSN 0925-4005,

S. De Vito, G. Fattoruso, M. Pardo, F. Tortorella and G. Di Francia, 'Semi-Supervised Learning Techniques in Artificial Olfaction: A Novel Approach to Classification Problems and Drift Counteraction,' in IEEE Sensors Journal, vol. 12, no. 11, pp. 3215-3224, Nov. 2012.

doi: 10.1109/JSEN.2012.2192425