

UJIAN AKHIR SEMESTER GENAP 2023/2024

Nama Mata Kuliah	: Bengkel Koding	Sifat	: Take Home
Hari / Tanggal	: Jumat, 12 Juli 2024	Waktu	: -
Pukul	: -	Dosen	: Tim Data Science

Masa Pengerjaan : **Tanggal 3 Juni s.d 12 Juli 2024 23.59 WIB**

KETENTUAN CAPSTONE PROJECT BENGKEL KODING DATA SCIENCE

Bagi mahasiswa Bengkel Koding Data Science diwajibkan untuk mengerjakan capstone berikut untuk tugas akhir dari mata kuliah Bengkel Koding ini. Ketentuan project sebagai berikut :

Dataset

Dataset diberikan oleh asisten Data Science Bengkel Koding sebagai berikut :

Dataset : Hungarian Dataset

[processed.hungarian.data](https://archive.ics.uci.edu/dataset/45/hungarian)

Sumber : Klasifikasi Data Penyakit Jantung

<http://archive.ics.uci.edu/dataset/45/heart+diseas>

Tahapan Pengerjaan Capstone**1. Pengumpulan Data**

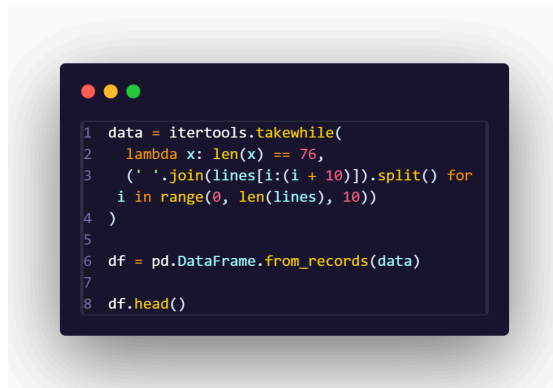
Dikarenakan bentuk dataset yang digunakan tidak terstruktur maka gunakanlah code dibawah ini untuk membaca file dataset

```
1 dir = 'hungarian.data'
2 with open(dir, encoding='Latin1') as file:
3     lines = [line.strip() for line in file]
4
5 lines[0:10]
```

output:

```
['1 0 63 1 -9 -9 -9',  
 '-9 1 145 1 233 -9 50 20',  
 '1 -9 1 2 2 3 81 0',  
 '0 0 0 0 1 10.5 6 13',  
 '150 60 190 90 145 85 0 0',  
 '2.3 3 -9 172 0 -9 -9 -9',  
 '-9 -9 -9 6 -9 -9 -9 2',  
 '16 81 0 1 1 1 -9 1',  
 '-9 1 -9 1 1 1 1 1',  
 '1 1 -9 -9 name']
```

2. Menelaah Data



code di atas akan menghasilkan tabel berisi 76 kolom fitur yang nantinya akan diseleksi kembali manakah fitur yang akan digunakan.

output:

	0	1	2	3	4	5	6	7	8	9	...	66	67	68	69	70	71	72	73	74	75
0	1	0	63	1	-9	-9	-9	-9	1	145	...	1	1	1	1	1	1	1	-9	-9	name
1	2	0	67	1	-9	-9	-9	-9	4	160	...	1	1	1	1	1	1	1	-9	-9	name
2	3	0	67	1	-9	-9	-9	-9	4	120	...	2	2	1	1	1	7	3	-9	-9	name
3	4	0	37	1	-9	-9	-9	-9	3	130	...	1	1	1	1	1	1	1	-9	-9	name
4	6	0	41	0	-9	-9	-9	-9	2	130	...	1	1	1	1	1	1	1	-9	-9	name

5 rows × 76 columns

3. Validasi Data

Pada tahap ini bertujuan untuk mengetahui dan memahami isi dari dataset agar dapat dilakukan penanganan sesuai dengan kondisinya. Didalam dataset ini **nilai -9.0 merupakan nilai NULL** oleh karena itu ubahlah agar dataset ini sesuai dengan kondisi tersebut.

4. Menentukan Objek Data

Dalam dataset Hungarian fitur yang hanya digunakan hanya 14 fitur, pilih dan kemudian berikan pada masing-masing fitur dengan:

- (1) Kolom 3 (age)
- (2) Kolom 4 (sex)
- (3) Kolom 9 (cp)
- (4) Kolom 10 (trestbps)
- (5) Kolom 12 (chol)
- (6) Kolom 16 (fbs)
- (7) Kolom 19 (restecg)
- (8) Kolom 32 (thalach)
- (9) Kolom 38 (exang)
- (10) Kolom 40 (oldpeak)
- (11) Kolom 41 (slope)
- (12) Kolom 44 (ca)
- (13) Kolom 51 (thal)
- (14) Kolom 58 (num)

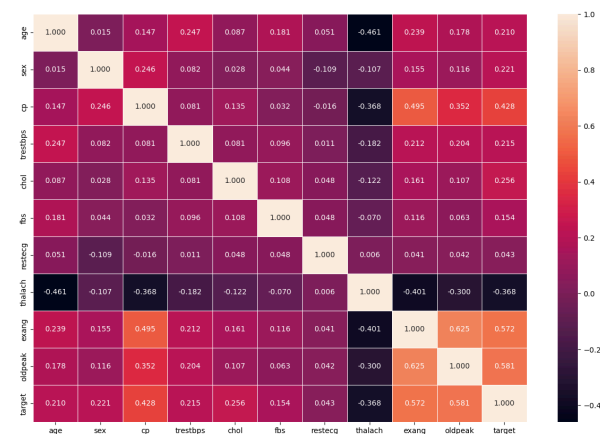
Output:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	target
0	40.0	1.0	2.0	140.0	289.0	0.0	0.0	172.0	0.0	0.0	0.0
1	49.0	0.0	3.0	160.0	180.0	0.0	0.0	156.0	0.0	1.0	1.0
2	37.0	1.0	2.0	130.0	283.0	0.0	1.0	98.0	0.0	0.0	0.0

5. Membersihkan Data

Sebelum melakukan pemodelan dilakukan pembersihan data agar model yang dihasilkan lebih akurat.

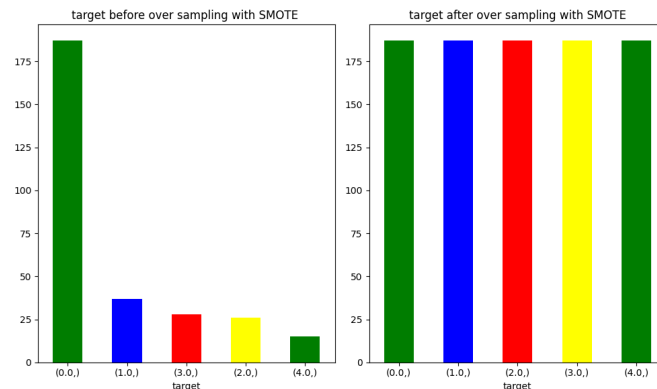
output:



6. Konstruksi Data

Dalam tahap ini Konstruksi data salah satu tujuannya yaitu untuk menyesuaikan semua tipe data yang ada di dalam dataset. Namun pada tahap ini dataset sudah memiliki tipe data yang sesuai sehingga tidak perlu dilakukan penyesuaian kembali.

output:

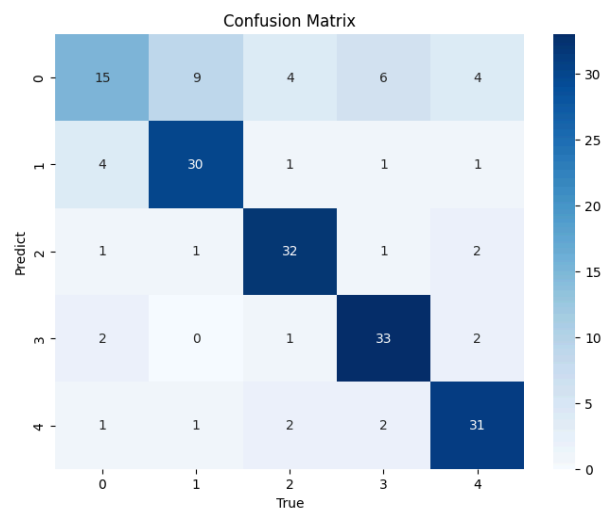


7. Pemodelan

Project **wajib** menggunakan **minimal 3 algoritma** yang kemudian dikomparasi akurasi satu sama lain.

output:

berikan confusion matrix untuk setiap model contoh:

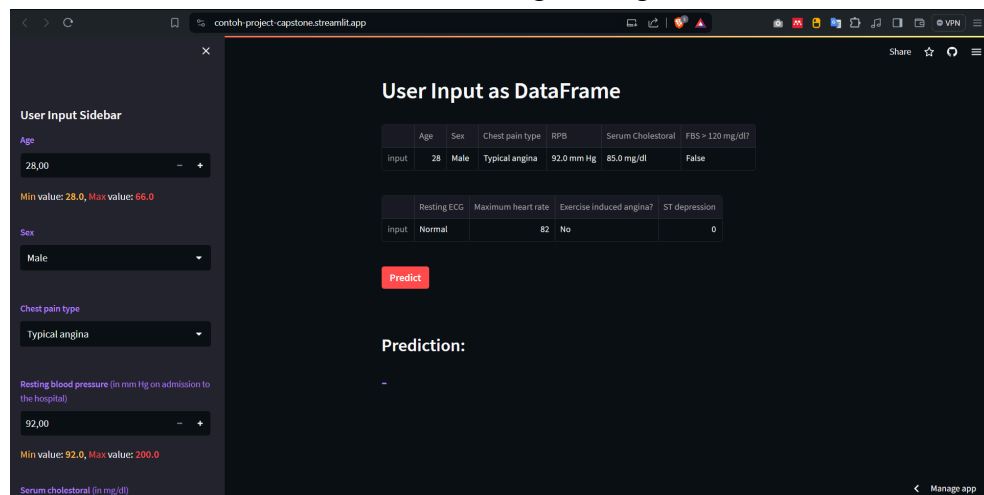


8. Evaluasi

- Setelah melakukan pemodelan harus dilakukan evaluasi model yang telah dibuat
- Setiap model / algoritma harus menghasilkan **minimal akurasi 70%**.

9. Deployment

- Deployment dapat menggunakan **satu model terbaik** atau **semua model sekaligus**.
- Deployment dilakukan menggunakan Streamlit, ketentuan template deployment berdasarkan kreativitas mahasiswa masing-masing.



Tambahan:

- Akan ada nilai tambahan jika project telah deploy secara **online** (contoh: Streamlit Cloud). Tutorial deployment ke streamlit cloud:



10. Kesimpulan

Presentasi Capstone Project

Jenis Presentasi dapat dipilih sesuai dengan kesepakatan kelas

Pilihan 1 : Mahasiswa dapat datang ke Bengkel Koding dan bertemu secara langsung dengan Mentor untuk melakukan presentasi (**Offline**)

Pilihan 2 : Mahasiswa dapat melakukan presentasi secara **online** dengan syarat sebagai berikut :

- (1) Mahasiswa **Wajib menghidupkan kamera (OnCam)**
- (2) Durasi presentasi minimal 10 menit, dengan ketentuan:
 - a. 2 menit untuk memperkenalkan Nama, NIM.
 - b. 3 menit untuk demo/menguji hasil program.
 - c. 5 menit untuk menjelaskan program yang dibuat

Ketentuan Pengumpulan Capstone Project

Pengumpulan project ini dikumpulkan di website Bengkel Koding menggunakan format file pdf, berupa :

- Link Github source code project
- Link Streamlit yang sudah di deploy (jika ada)
- Link Video Presentasi

- Bahan untuk presentasi (Sesuai kesepakatan masing-masing kelas)

Capstone Project ini BERSIFAT WAJIB,

Diharapkan mahasiswa mengumpulkan tepat waktu project capstone ini, bila ada keterlambatan dimohon untuk konfirmasi terlebih dahulu ke asisten Bengkel Koding masing-masing sebelum waktu tenggat pengumpulan. Jika tidak atau melanggar ketentuan ada sanksi pengurangan nilai.

Sekian, terimakasih.

Mengetahui,

**Asisten
Bengkel Koding Data Science**

**Dosen Pengampu
Bengkel Koding Data Science**