



RESTful – API Text Cleansing Program

Harya Kumuda

ROADMAP

Pembuatan database untuk menyimpan data

1

Pembuatan RESTful API

3

Menjalankan API

5

Pembuatan program untuk cleansing data

2

Pembuatan Dokumentasi API

4

API dapat diakses melalui browser

6

TOOLS



Python3

Base Programming
Language



Pandas

Data Manipulation
and Analysis

RegEx

RegEx

Text Cleansing Tool



Flask

RESTful API
Framework



SQLite3

Database
Management
System



Swagger

API Documentation

DATABASE

- Database merupakan tempat menyimpan data
- Database kali ini dibuat menggunakan SQLite3
- Database dibuat dengan hanya 1 tabel berisi 3 kolom.

```
4 create table tweet
5 (tweet_id INTEGER PRIMARY KEY AUTOINCREMENT, tweet_kotor text, tweet_bersih text);
```

01

Tweet_id : nomer id untuk masing masing text yang diinput

02

Tweet_kotor : text yang dimasukkan

03

Tweet_bersih : text yang telah dicleansing

tweet_id	tweet_kotor	tweet_bersih
INTEGER	TEXT	TEXT
1	Prabowo Sudah Kalah lv	prabowo sudah kalah m
2	USER USER dan yang tal	dan yang takut dengan

DATA CLEANSING

Data Cleansing pada program ini, akan dilakukan 4 tahapan
Cleansing Text menggunakan library RegEx

Tahap ini berfungsi agar kedepannya data dapat digunakan sebagai bahan pembuatan machine learning

01. Mengubah seluruh karakter menjadi lowercase

```
11 def lowercase(text):  
12     return text.lower()
```

02. Menghapus karakter yang tidak perlu

```
14 def remove_unnecessary_char(text):  
15     text = re.sub('\n', ' ', text)  
16     text = re.sub('rt', ' ', text)  
17     text = re.sub('user', ' ', text)  
18     text = re.sub('((www\.[^\s]+)|(https?://[^\s]+)|(http://[^\s]+))', ' ', text)  
19     text = re.sub(' +', ' ', text)  
20     return text
```

03. Menghapus karakter selain alphanumeric

```
22 def remove_nonaplhanumeric(text):  
23     text = re.sub('[^0-9a-zA-Z]+', ' ', text)  
24     text = re.sub(' +', ' ', text)  
25     return text
```

04. Menormalisasi kata kata slang

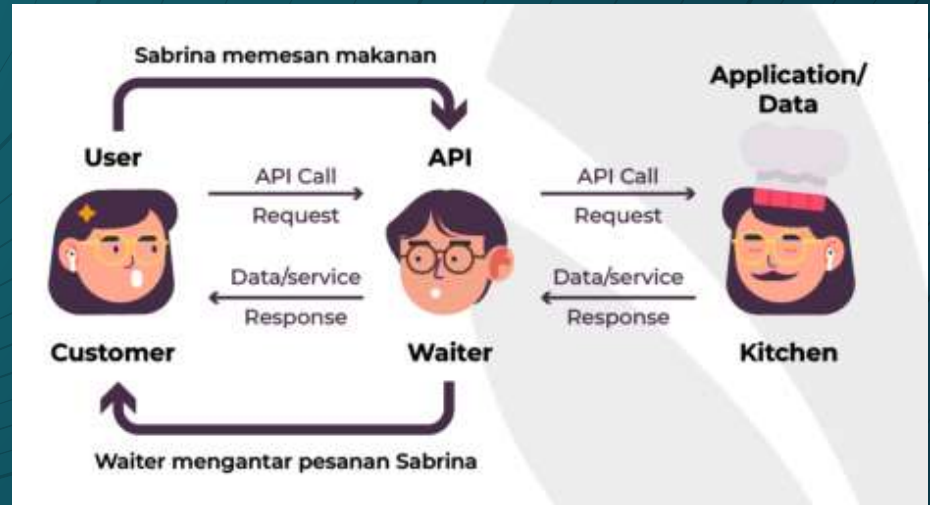
```
27 alay_dict_map = dict(zip(t_kamusalay['alay'], t_kamusalay['cleaned']))  
28 def normalize_alay(text):  
29     for word in alay_dict_map:  
30         return ' '.join([alay_dict_map[word] if word in alay_dict_map else word for word in text.split(' ')])
```

API

Ibarat sebuah restoran,

1. User sebagai Customer akan memesan makanan melalui pelayan.
2. API sebagai Pelayan akan memberitahu dapur keinginan customer
3. Lalu Server sebagai Dapur Restoran akan menyajikan apa yang diminta oleh Customer

Secara ringkas API merupakan perantara user dengan server, sehingga user dapat berkomunikasi dengan server dan tidak perlu tahu proses apa saja yang terjadi dibalik permintaannya

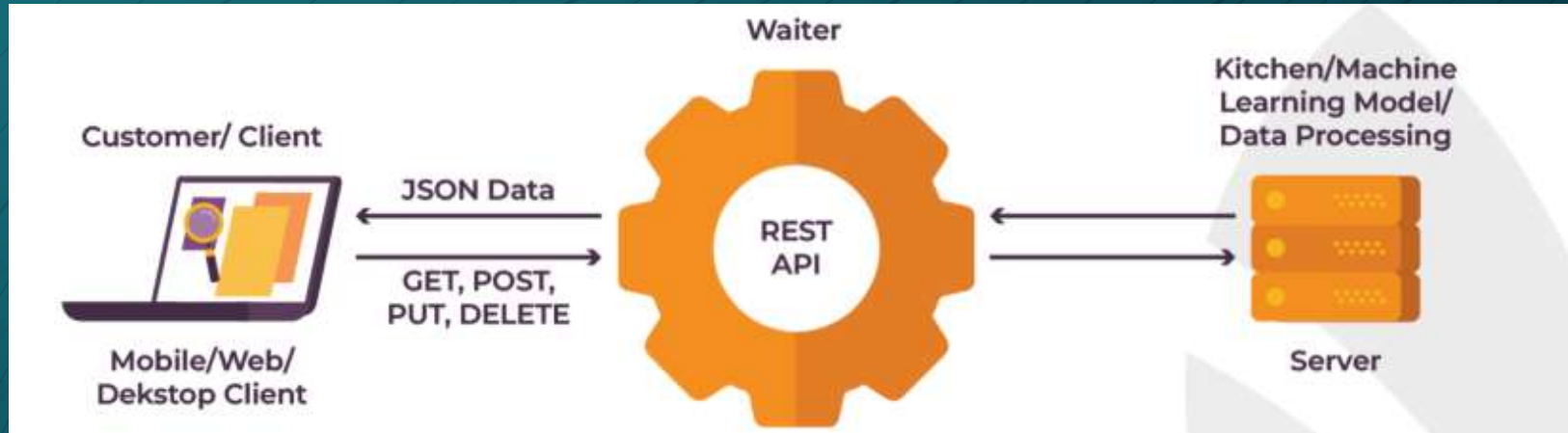


RESTful API

RESTful API adalah API yang memanfaatkan HTTP sebagai alat komunikasinya, sehingga mudah diakses melalui browser seperti Google Chrome, Mozilla Firefox, Microsoft Edge dll.

API yang akan dijalankan menggunakan server localhost / 127.0.0.1 pada port 5000,

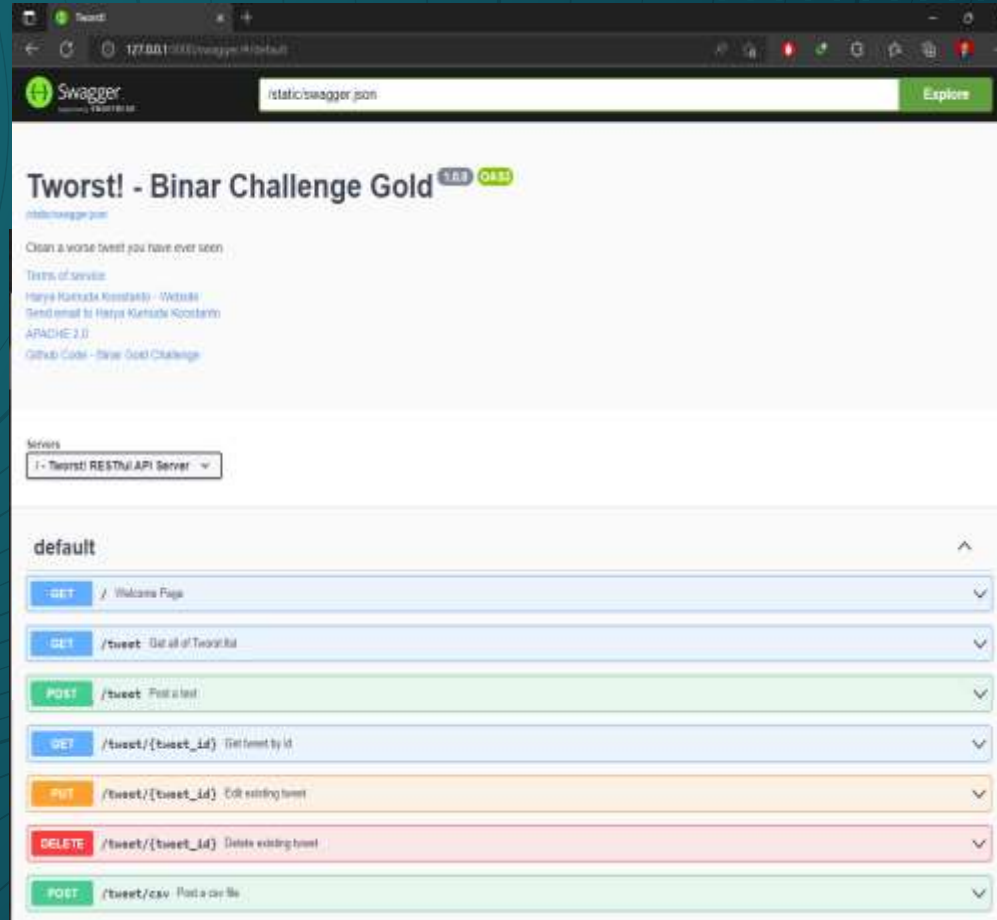
Lalu url dan metode HTTP akan ditentukan sesuai dengan kebutuhan



API DOCUMENTATION

Swagger digunakan untuk melakukan dokumentasi API.

Dokumentasi API berfungsi agar user tahu, apa saja opsi yang tersedia pada API yang telah kita buat



INPUT DATA

Saat mengirimkan data / tweet baru menggunakan metode POST, maka secara otomatis tweet tersebut masuk kedalam database dilengkapi dengan tweet ID dan hasil cleansing dari tweet tersebut.

POST /tweet Post a text

Post a text file to clean

Parameters

No parameters

Request body required

text
string

Hari Ini Begitu Cerah !!!! <https://www.haryakumuda.com>

☐ Send empty value

Execute

Response body

```
[
  {
    "tweet_id": 41,
    "tweet_kotor": "Hari Ini Begitu Cerah !!!! https://www.haryakumuda.com",
    "tweet_bersih": "hari ini begitu cerah "
  }
]
```

START SERVER

Setelah server dijalankan menggunakan Command Prompt.

API dapat diakses dengan browser yang dimiliki.

YouTube 

Video demonstrasi
dapat ditonton di
Youtube:

https://youtu.be/izxy_tXFqbs?t=389



GitHub

akses full code

GitHub:

<https://github.com/haryakumuda/RESTful-API-TWORST>



Thanks!

Let's connect!

You can find me on:

[linkedin.com/in/haryakumuda](https://www.linkedin.com/in/haryakumuda)