

1.ทำการโหลดไลบรารีที่จำเป็นทั้งหมดในภาพมาก่อนการทดลองใช้

```
import pandas as pd
import json
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, OneHotEncoder, LabelEncoder
from sklearn.compose import ColumnTransformer
from sklearn.utils.class_weight import compute_sample_weight
from xgboost import XGBClassifier
from sklearn.pipeline import Pipeline
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics import classification_report, f1_score, precision_score, recall_score

from sklearn.metrics import confusion_matrix
import matplotlib.pyplot as plt
import seaborn as sns
import matplotlib.font_manager as fm
import random
```

## 2.กำหนด path ของ data set ให้ถูกต้อง

```
path = r"C:\Users\PC\Desktop\data\[CONFIDENTIAL] AI symptom picker data (Agnos candidate assignment) - ai_symptom_picker.csv"
df = pd.read_csv(path)
```

โครงสร้างข้อมูลที่จะนำมาใช้งานได้

gender	age	summary	search_term
male	28	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "เส้มนะ", "answers": ["ลักษณะ เสี้นะเป็นสีเหลืองเขียว"], "text": "โ", "answ": "สีน้มนะ โ"}]	
male	27	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "โ", "answers": ["ระยะเวลา 1-3 สัปดาห์", "ลักษณะ โดไม่มีเม้น เสี้นะ โนแห้ง หรือ โ", "น้มนะ โ"}]	
female	26	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "ป้ะลวงโ", "answers": ["ป้ะริเวณ รอยจุดสี", "ระยะเวลา สัปดาห์ 1 วัน ถึง 1 สัปดาห์", "ป้ะลวงโ"}]	
male	42	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "น้มนะโ", "answers": ["ระยะเวลา น้มนะกว่า 10 วัน", "ป้ะริเวณ ATK สัปดาห์โ", "น้มนะโ"}]	
female	40	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "าานะโ", "answers": [{"text": "การรักษานานกว่า", "answers": ["การรักษานานกว่า านะโ"}]	
female	49	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "ป้ะลวงโ", "answers": ["ระยะป้ะลวง", "answers": ["ระยะป้ะลวงนานกว่า 1 สัปดาห์", "ป้ะริเวณ ATK สัปดาห์โ", "ป้ะลวงโ"}]	
female	44	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "น้มนะโ", "answers": ["ระยะเวลา น้มนะกว่า 10 วัน", "ป้ะริเวณ ATK สัปดาห์โ", "น้มนะโ", "น้มนะจากนานโ, โ"}]	
female	18	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "ป้ะลวงโ", "answers": ["ป้ะริเวณ รอยจุดสี", "ระยะเวลา สัปดาห์ 1 วัน ถึง 1 สัปดาห์", "ป้ะลวงโ"}]	
female	40	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "โ", "answers": ["ป้ะริเวณ โ", "ป้ะริเวณ ATK สัปดาห์โ", "โ", "โ", "โ", "โ"}]	
male	48	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "โ", "answers": ["ระยะเวลา โ", "ลักษณะ โ", "น้มนะ โ"}]	
female	40	"diseases": [], "procedures": [], "no_symptoms": [{"text": "ป้ะลวงโ", "answers": [{"text": "ป้ะริเวณ น้มนะ", "answers": [{"text": "tdk_symptoms": [], "yes_symptom": "าานะโ"}]}]}]	
female	49	"diseases": [], "procedures": [], "no_symptoms": [{"text": "ป้ะลวงโ", "answers": [{"text": "ป้ะริเวณ น้มนะ", "answers": [{"text": "tdk_symptoms": [], "yes_symptom": "าานะโ"}]}]}]	
female	34	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "เส้มนะโ", "answers": [{"text": "ป้ะริเวณ น้มนะ", "answers": [{"text": "tdk_symptoms": [], "yes_symptom": "าานะโ"}]}]}]	
female	37	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "Feve", "answers": [{"text": "Duration 1-3 days", "Grade Low grade between 3-5.5 Fever"}]	
female	37	"diseases": [], "procedures": [], "no_symptoms": [], "tdk_symptoms": [], "yes_symptoms": [{"text": "น้มนะโ", "answers": ["ระยะเวลา น้มนะกว่า 10 วัน", "ป้ะริเวณ ATK สัปดาห์โ", "น้มนะโ", "น้มนะโ"}]	

3.เมื่อรันโค้ดจนไปถึงฟังก์ชันโมเดลเสร็จสิ้นแล้วให้ลองทดสอบความถูกต้องของโมเดลโดยการรันโค้ดนี้

```
|  
sympom_options = [  
    ['ummm', 'la'],  
    ['la', 'squnlna'],  
    ['dsmlas'],  
    ['squnlna'],  
    ['auuuu'],  
    ['dusnnegn'],  
    ['squnln', 'la'],  
    ['duwslal'],  
    ['tuumm', 'la'],  
    ['la'],  
    ['awtkm'],  
    ['dwadudnuwadua'],  
  
]  
  
n_samples = 5  
  
new_data_list = []  
  
for _ in range(n_samples):  
    age = random.randint(1, 100)  
    gender = random.choice(['male', 'female'])  
    sympoms = random.choice(sympom_options)  
    sympoms_str = ", ".join(sympoms)  
  
    new_data_list.append({  
        'age': age,  
        'gender': gender,  
        'yes_syptoms_textis': sympoms_str  
    })  
  
new_data = pd.DataFrame(new_data_list)  
  
new_data['yes_syptoms_textis'] = new_data['yes_syptoms_textis'].str.replace(',', ' ').str.lower().str.strip()  
  
X_new_vec = preprocessor.transform(new_data[['age', 'gender', 'yes_syptoms_textis']])  
y_new_prob = clf.predict_proba(X_new_vec)  
classes = le.classes_  
  
default_threshold = 0.1  
  
y_new_pred_main = []  
y_new_pred_top5 = []  
  
for row in y_new_prob:  
    filtered_idx = [i for i, prob in enumerate(row) if prob > best_thresholds.get(classes[i], default_threshold)]
```

	age	gender	yes_symptoms_texts	predicted_main	\
0	71	male	ปวดท้อง	แสนท้อง	
1	39	male	น้ำมูกไหล	น้ำมูกไหล	
2	58	male	ไอ	ไอ	
3	40	female	ปวดกระดูก	ปวดกระดูก	
4	91	male	ปวดเมื่อยกล้ามเนื้อ	ปวดเมื่อยกล้ามเนื้อทั่วๆ	

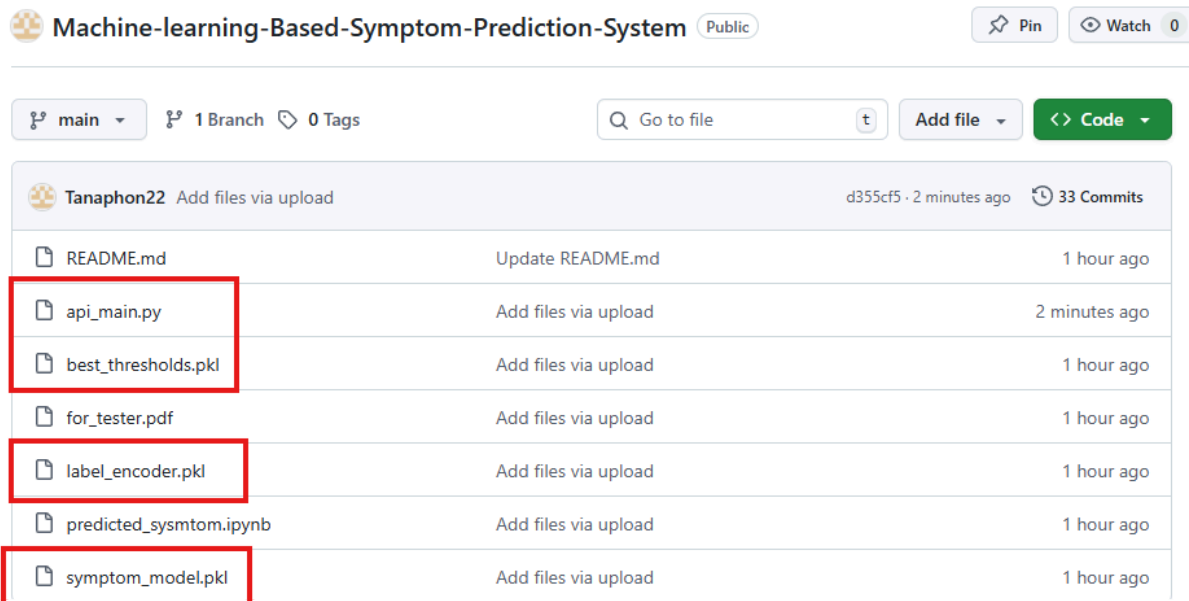
	predicted_top5
0	เรอเปรี้ยว, จุกแน่นท้องปวดท้อง, คลื่นไส้, ก้อน...
1	ไอไอลางคิน, ปวดหัว, คันจมูกจากน้อย, มีน้ตึระะ,...
2	ผิวแห้งผืน, มีเสมหะน้ำมูกไหล, หูอื้อ, คัน, ตั...
3	มีน้ตึระะ, ปวดหัวปวดท้ายทอย, ปวดกระบอกตา, ตาแห้ง...
4	คัน, หูอื้อ, ฝ่าขาที่ล้น, ผื่น, กลืนติด

4.ถ้าจะทดสอบ api ต้องให้เครื่องหนึ่ง host และทำตามขั้นตอนนี้

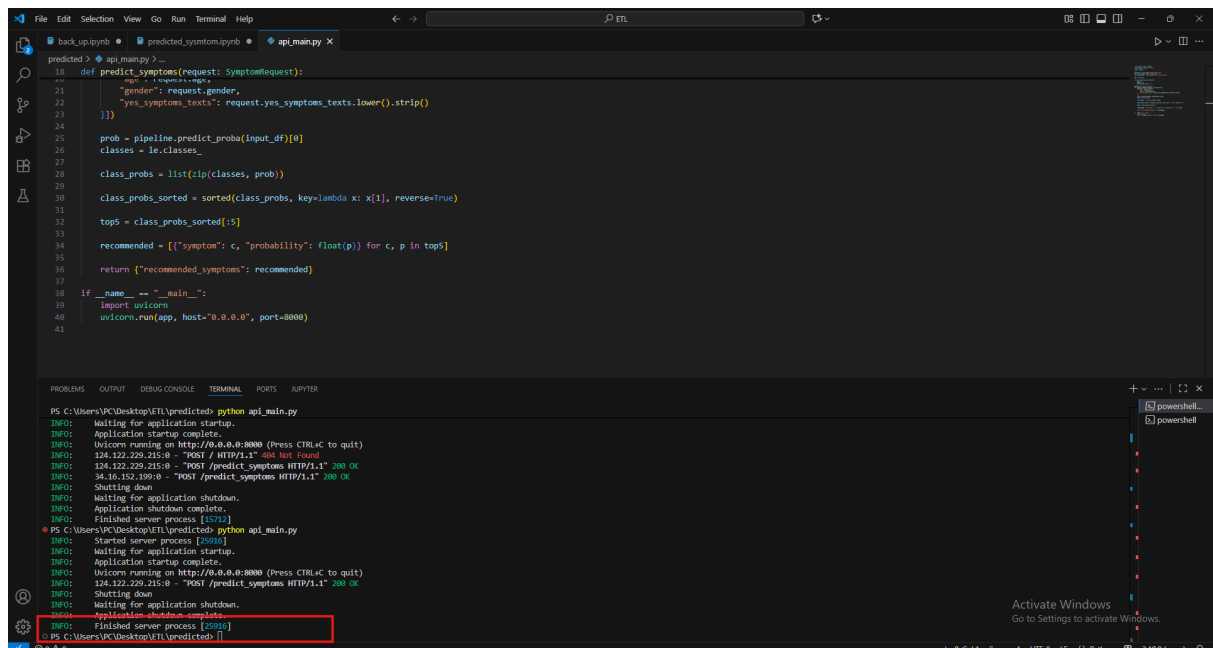
4.1 ดาวนโหลดและติดตั้ง [Node.js](#) เข้ามาในเครื่อง host ก่อน

4.2. โหลดไฟล์ที่มีชื่อ

- api\_main.py
- symptom\_model.pkl
- label\_encoder.pkl
- best\_thresholds.pkl



4.3 ต่อมาให้นำไฟล์จากข้อ 4.2 มารวมไว้ใน folder เดียวกัน และจากภาพ vscode ให้ทำการเปิด powershell ออกมาและลงไลบรารีให้ครบตามข้อที่ 1



4.4 ต่อไปให้ทำการรันคำสั่ง 'python api\_main.py' ใน powershell เพื่อรัน fastapi

```
PS C:\Users\PC\Desktop\ETL\predicted> python api_main.py
INFO: Started server process [11884]
INFO: Waiting for application startup.
INFO: Application startup complete.
INFO: Uvicorn running on http://0.0.0.0:8000 (Press CTRL+C to quit)
```

หน้าตาจะประมาณนี้

4.4 เมื่อรัน fastapi ได้แล้วให้ทำการเปิด powershell อีกหน้าต่างหนึ่งและรัน 'lt --port 8000' ตามรูปข้างล่าง

```
(.venv) PS C:\Users\PC\Desktop\ETL> lt --port 8000
your url is: https://violet-apes-heal.loca.lt
```

เราจะได้ url มาให้ทำการ copy เอาไว้และใช้คำสั่ง python ข้างล่างนี้ในการใช้งาน api

```
import requests

url = "https://lucky-readers-ask.loca.lt/predict_symptoms"
data = {
    "age": 30,
    "gender": "ชาย",
    "yes_symptoms_texts": "ไอ มีเสมหะ"
}

response = requests.post(url, json=data)
print(response.json())

✓ 24s Python
{'recommended_symptoms': [{'symptom': 'ไอล่างคั่นมีเสมหะ', 'probability': 0.6738181710243225}, {'symptom': 'มีเสมหะน้ำูกไหล', 'probability': 0.07257305085659027}, {'symptom': 'มีเสมหะ', 'probability': 0.07257305085659027}]}
```

Note : url + /predict\_symptoms ถ้าเอา url ที่ได้จาก [Node.js](#) อย่างเดียวมันจะใช้งานไม่ได้