

Take-home midterm exam with Traffy Fondue dataset

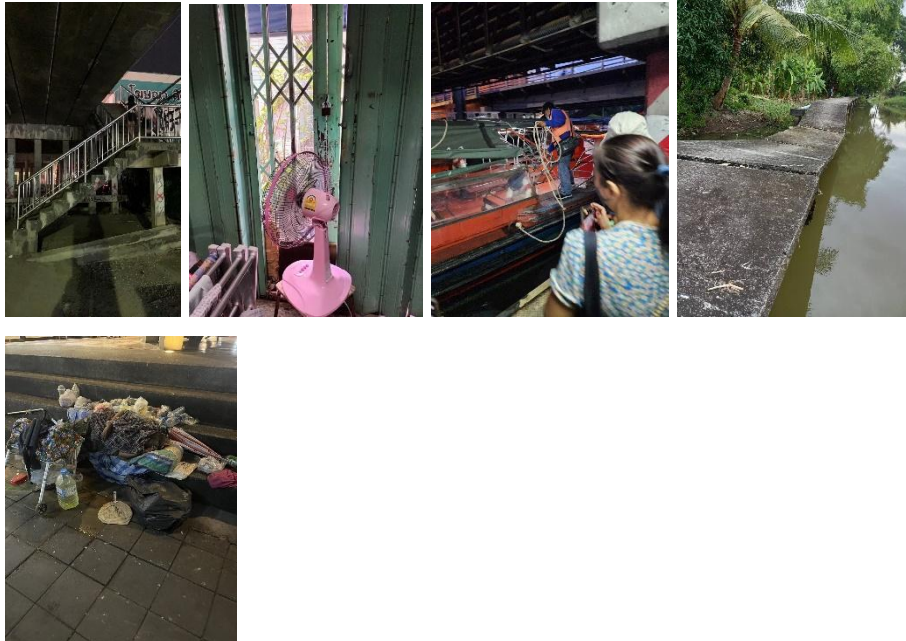
Chapter 1: Introduction

Traffy Fondue เป็น application ที่รับข้อเสนอแนะจากผู้แจ้งหรือประชาชน ทำให้การบริหารและการแก้ปัญหาในพื้นที่ได้มีประสิทธิภาพมากขึ้น โดยกระบวนการใช้งานออกแบบมาให้ง่ายต่อการใช้งาน เพียงแค่ผู้แจ้งถ่ายรูปและระบุประเภทของปัญหา ระบบจะทำการแจ้งปัญหาดังกล่าวไปยังเจ้าหน้าที่และทีมงานที่รับผิดชอบทันที โดยระบบจะทำหน้าที่คัดแยกประเภทของปัญหาและแจ้งต่อเจ้าหน้าที่ที่รับผิดชอบ

แต่เนื่องจากรูปถ่ายบางรูปมีความกำกวมรวมถึงมีจำนวนรูปเป็นจำนวนมาก ทำให้ยากต่อการคัดแยกปัญหา ทำให้เกิดโจทย์ image classification นี้ขึ้นมา เพื่อทำนายว่ารูปที่ได้รับมาเป็นปัญหาอะไร ซึ่งจะช่วยได้ทั้งเจ้าหน้าที่เพื่อคัดแยกรูปที่ส่งมาแล้ว รวมทั้งผู้ใช้งานที่ระบบจะทำการ recommend ปัญหาจากรูปที่ส่ง ก่อนกดส่งทำให้แจ้งปัญหา ทำให้ทำงานได้ถูกต้องและมีประสิทธิภาพมากยิ่งขึ้น

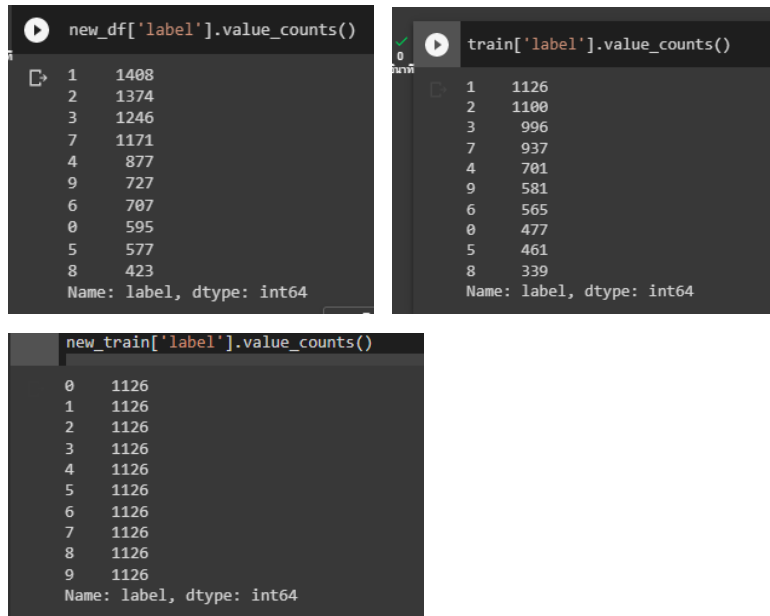
Chapter 2: Data preparation

1. เริ่มจาก manual clean data โดยการเลือกรูปในแต่ละ category และลบรูปที่มองว่าดูยากหรือไม่สามารถระบุได้ว่าอยู่ category อะไร เช่น รูปป้ายข้างทาง รูปคนไร้บ้าน(อาจทำให้ model มองว่าคนเป็นขยะได้ จึงต้องเอาออก) ทางเดินที่อยู่ริมคลอง เป็นต้น



2. หลังจากได้รูปมาจำนวนหนึ่ง ก่อนที่จะนำเข้าโมเดล ได้ตรวจสอบจำนวนรูปที่มีในแต่ละ category แล้วพบว่า imbalance ดังนั้นจึงได้ทำการ over sampling รูปใน category ที่

น้อยกว่า หลัง split data ออกมาแล้ว เช่น ต้องการ 1126 รูปต่อ category แต่มี 565 ก็จะทำ random with replacement เพิ่มมาอีก $1126 - 565 = 561$ ใน category นั้นๆ



```
new_df['label'].value_counts()
1    1408
2    1374
3    1246
7    1171
4     877
9     727
6     707
0     595
5     577
8     423
Name: label, dtype: int64
```

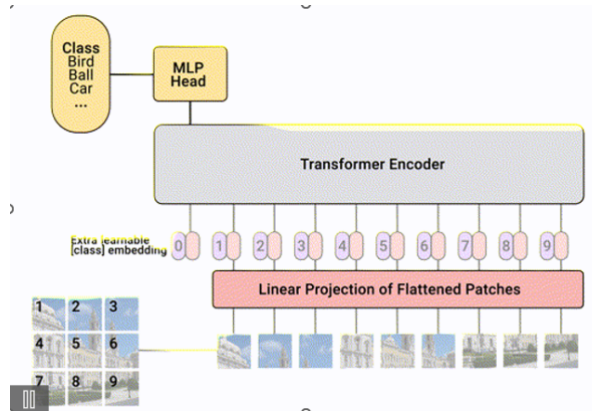
```
train['label'].value_counts()
1    1126
2    1100
3     996
7     937
4     701
9     581
6     565
0     477
5     461
8     339
Name: label, dtype: int64
```

```
new_train['label'].value_counts()
0    1126
1    1126
2    1126
3    1126
4    1126
5    1126
6    1126
7    1126
8    1126
9    1126
Name: label, dtype: int64
```

3. ทำการ Transform image ที่ได้ โดยใช้ config เหมือนที่ pre-trained ใช้ คือ [google/vit-base-patch16-224-in21k](https://www.tensorflow.org/api_guides/python/tf_image) โดยการ
 - a. Resize image เป็น 224*224
 - b. Normalize RGB ด้วย mean (0.5, 0.5, 0.5) และ standard deviation (0.5, 0.5, 0.5)

Chapter 3: Model

ใช้ Vision transformer model กับ pre-trained "google/vit-base-patch16-224-in21k" โดยนำ hidden ที่ออกมาจาก vision transformer model มา regularizing ด้วย dropout ที่ 0.1 ก่อนเข้า linear model ที่รับ input เท่ากับ hidden size ที่ออกมาจาก vision transformer และ output เท่ากับขนาด target ที่ต้องการ ซึ่งก็คือ 10 class



Vision transformer: credit <https://roboflow.com/model/vision-transformer>

Chapter 4: Results

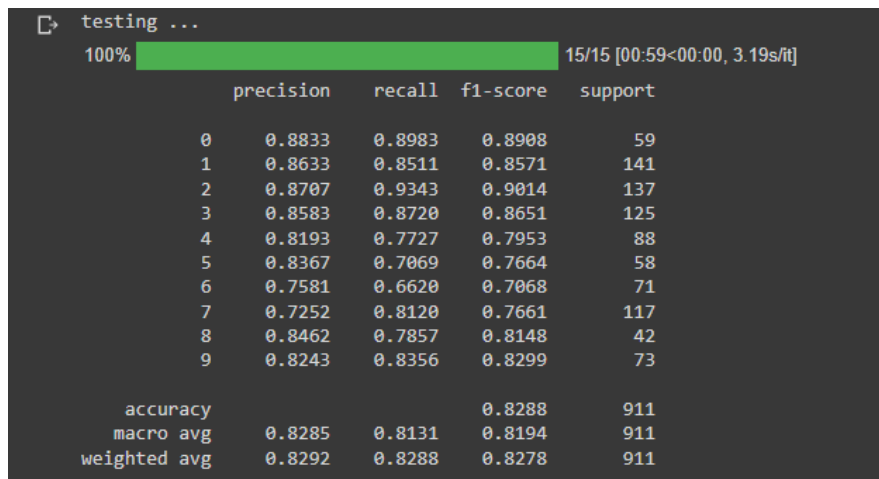
ผลที่ได้คือที่ 11 ใน private score ที่คะแนน F1 score ที่ 0.81054

#	Team	Score	Entries	Last	Solution
1	Thattithong Sci	0.84378	38	9h	
2	PinnakApichathanon633031221	0.82089	14	1d	
3	Pacharapon A. 6331325621	0.82012	17	10h	
4	Paveeran Phupempoon 633039822	0.81828	12	11h	
5	ThanabulChittamerd6330216721	0.81910	36	9h	
6	Charon Rattanasarakul	0.81762	58	9h	
7	Rada W. 6332038721	0.81392	14	11h	
8	Nonyia 6332018021	0.81301	29	2d	
9	6331321021 Paveerit Thunrat...	0.81148	2	2d	
10	Nitawat J. 6331320421	0.81128	7	14h	
11	KhunanonRattanasoe6332003721	0.81054	18	9h	

✓	submitting_ds (9).csv Complete · 9h ago	0.81054	0.81785	✓
✓	submitting_ds (8).csv Complete · 11h ago	0.80956	0.81803	✓

Chapter 5: Discussion

เช็คผล evaluate บน test set จาก cleaned dataset



	precision	recall	f1-score	support
0	0.8833	0.8983	0.8908	59
1	0.8633	0.8511	0.8571	141
2	0.8707	0.9343	0.9014	137
3	0.8583	0.8720	0.8651	125
4	0.8193	0.7727	0.7953	88
5	0.8367	0.7069	0.7664	58
6	0.7581	0.6620	0.7068	71
7	0.7252	0.8120	0.7661	117
8	0.8462	0.7857	0.8148	42
9	0.8243	0.8356	0.8299	73
accuracy			0.8288	911
macro avg	0.8285	0.8131	0.8194	911
weighted avg	0.8292	0.8288	0.8278	911

จากผลจะต้องคำนึงว่า dataset ที่ clean มาอาจไม่ได้ label ถูกต้องตามที่ควรจะเป็น เหตุมา
จากคำอธิบายของ label แต่ละ category ยังไม่ชัดเจน ทำให้ไม่สามารถเลือกได้ว่ารูปแต่ละรูปควรจะอยู่
หมวดปัญหาอะไร ทำให้จำเป็นต้องลบรูปนั้นออกทั้งที่บางทีอาจจะจะเป็นข้อมูลที่มีประโยชน์ต่อโมเดล ซึ่งส่ง
ผลต่อเนื่องไปที่ปัญหาข้อมูลน้อยจากการลบรูปไป ทำให้โมเดลไม่มีประสิทธิภาพ

สิ่งที่คิดว่าจะทำให้โมเดลดีขึ้น

1. เพิ่มคำอธิบายให้ชัดเจน มีตัวอย่างปัญหาได้ในใจยกนั้นๆ ที่หลากหลายกว่านี้
2. Image augmentation โดยการเพิ่มหรือลดแสง หรือหมุนรูปเป็นต้น เพื่อให้โมเดลเรียนรู้ได้
มากขึ้น

Chapter 5: Conclusion

จากโจทย์ Image classification ได้คะแนนจากการผลบน test set (Kaggle) F1 score ที่
0.81054 ซึ่งอยู่ในเกณฑ์ที่สามารถนำไปใช้จริงได้ เนื่องจากโจทย์เป็นลักษณะของการใช้โมเดลเป็นผู้ช่วยใน
การทำงานและลดข้อผิดพลาดในการทำงานของเจ้าหน้าที่ โดยถ้าอยากทำให้โมเดลเก่งขึ้น สามารถทำได้โดย
การเตรียมข้อมูลให้ดีขึ้น และใช้โมเดลที่มีประสิทธิภาพมากกว่านี้