

```

PROBLEMS    OUTPUT    TERMINAL    DEBUG CONSOLE
PS D:\Code and Scripts\Python\Machine Learning\OCR> pip install pytesseract

```

ทำการ install pytesseract เพื่อจะทำให้เราสามารถไ้ python กับ tesseract

```
1 import pytesseract as tess
2 from PIL import Image
3 tess.pytesseract.tesseract_cmd = r'D:\ocr\tesseract.exe'
4 image = Image.open('D:\ocr\kan2.png')
5 text = tess.image_to_string(image, lang='tha+eng')
6 print(text)
```

บรรทัดที่ 1 import flies tesseract

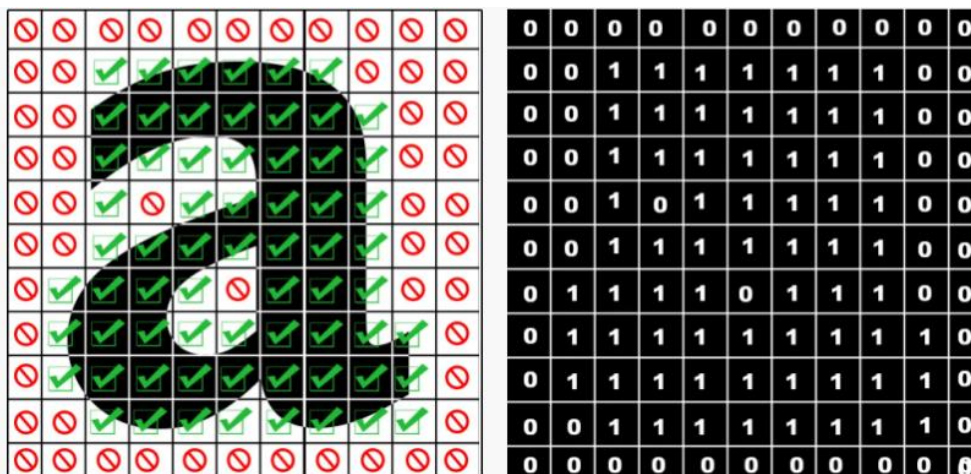
บรรทัดที่ 2 import image เพื่อใช้ในการเพิ่มรูปภาพเข้ามา

บรรทัดที่ 3 คือการอ่าน path files ของ tesseract

บรรทัดที่ 4 การอ่าน **files** รูปภาพแล้วกับไว้ในตัวแปร **image** เพื่อนำไปประมวลผล

บรรทัดที่ 5 การเอารูปภาพมาแปลงเป็น **text** จะแปลงได้แค่ ภาษาไทยและภาษาอังกฤษ

การรู้จำอักษรด้วยแสงทำงานโดยแบ่งรูปภาพของอักขระข้อความออกเป็นส่วนๆ และแยกความแตกต่างระหว่างพื้นที่ว่างและไม่ว่าง ขึ้นอยู่กับพונต์หรือสคริปต์ที่ใช้สำหรับตัวอักษร ผลรวมตรวจสอบของเมทริกซ์ผลลัพธ์จะถูกระบุในภายหลัง ตามอักขระในภาพ



รูปภาพที่ ผลรวมตรวจสอบของเมทริกซ์