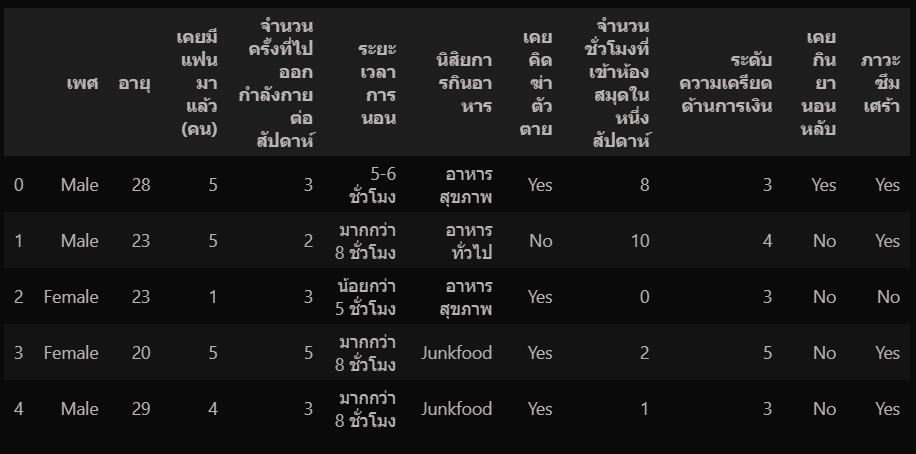
**การนำเข้าข้อมูลและการเตรียมข้อมูล**

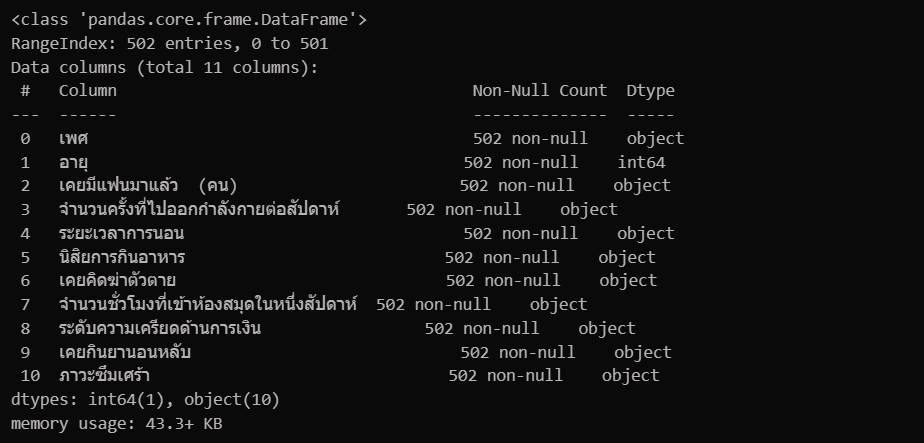
* โดยในชุดข้อมูลประกอบไปด้วย feature ดังนี้

|  |  |
| --- | --- |
| **features** | **Description** |
| เพศ | ระบุเพศ |
| อายุ | ระบุอายุ |
| เคยมีแฟนมาแล้ว (คน) | จำนวนแฟนที่เคยมีมาแล้ว |
| จำนวนครั้งที่ไปออกกำลังกายต่อสัปดาห์ | จำนวนครั้งความถี่ในการออกกาลังกายในหนึ่งสัปดาห์ |
| ระยะเวลาการนอน | ช่วงเวลาการนอนหลับ |
| นิสัยการกินอาหาร | ประเภทอาหารที่ชอบรับประทาน |
| เคยคิดฆ่าตัวตาย | ระบุว่าเคยมีความคิดฆ่าตัวตายหรือไม่ |
| จำนวนชั่วโมงที่เข้าห้องสมุดในหนึ่งสัปดาห์ | ระบุเวลาที่ใช้ในห้องสมุดต่อสัปดาห์ |
| ระดับความเครียดด้านการเงิน | ระดับความเครียดเกี่ยวกับการเงิน |
| เคยกินยานอนหลับ | ระบุว่าเคยใช้ยานอนหลับหรือไม่ |
| ภาวะซึมเศร้า | สถานะภาวะซึมเศร้า |

* แสดงชุดข้อมูลที่นำเข้ามา



* การแสดงรายละเอียดของชุดข้อมูล



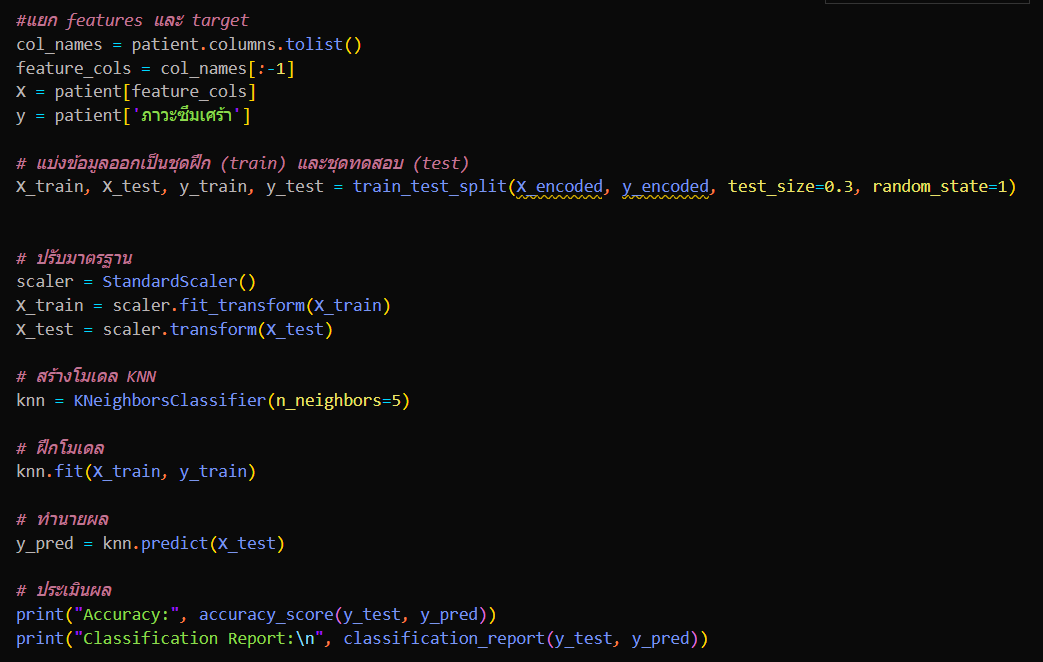
* แสดงรายละเอียดของ features target

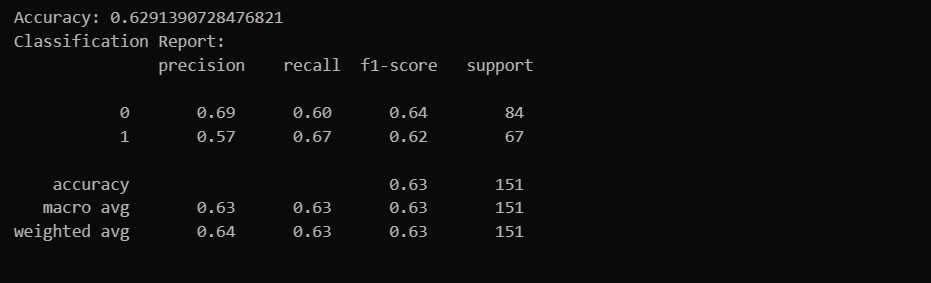


* แสดงชุดข้อมูลอีกครั้งหลังจากทำการแปลงข้อมูล

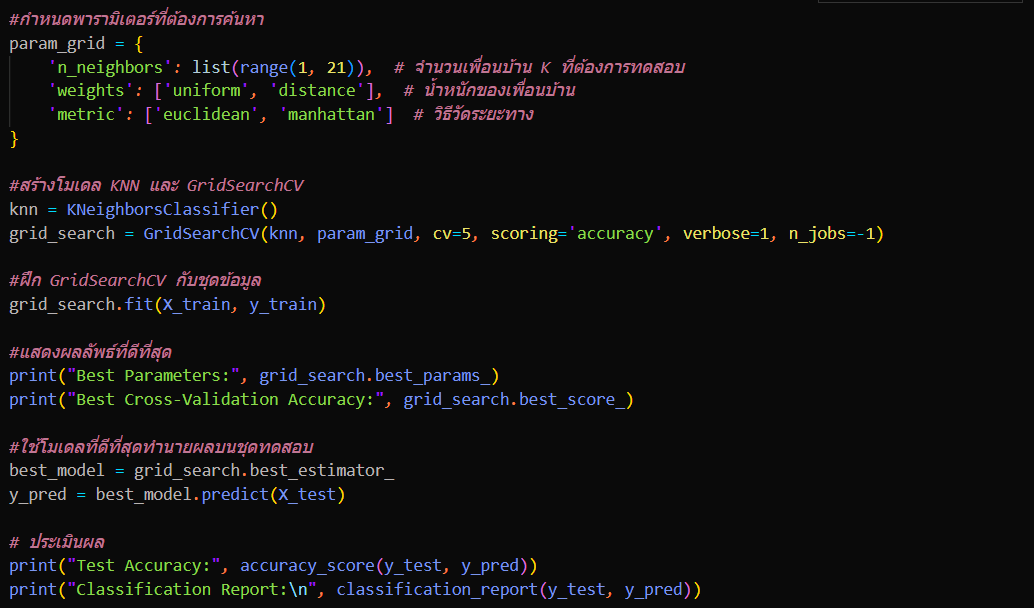


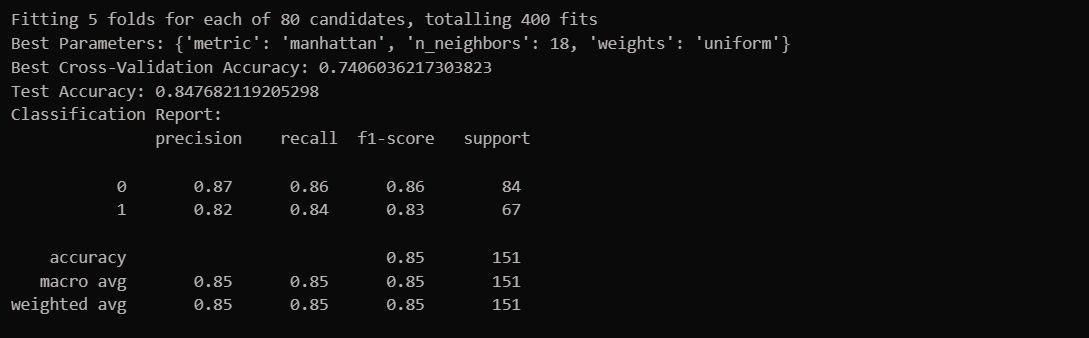
**การทำนายผล**

* การทำนายครั้งแรก

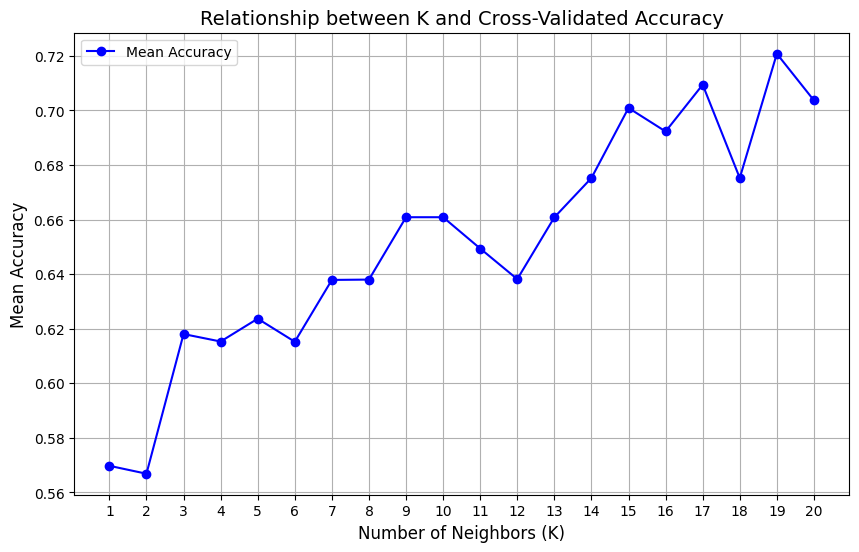
****

* ทำการหา GridSearchCV เพื่อหาค่า K ที่เหมาะสมที่สุดในการทำนายผล





* กราฟแสดงความสัมพันธ์ระหว่างค่า K กับค่าเฉลี่ยของ accuracy scores



**แกน X** แสดงจำนวนเพื่อนบ้าน (Number of Neighbors, K) ที่ใช้ในโมเดล KNN โดยเริ่มจาก 1-20

**แกน Y** แสดงค่าเฉลี่ยความแม่นยำ (Mean Accuracy) ของโมเดลในแต่ละ K โดยจะมีค่าอยู่ในช่วง 0-1

**จุดข้อมูล** แต่ละจุดแทนค่าความแม่นยำเฉลี่ยที่ได้จากการทำ Cross-Validation สำหรับ K แต่ละค่า

**การวิเคราะห์**

**1. ช่วงเริ่มต้น (K=1K = 1K=1 ถึง K=4K = 4K=4):**

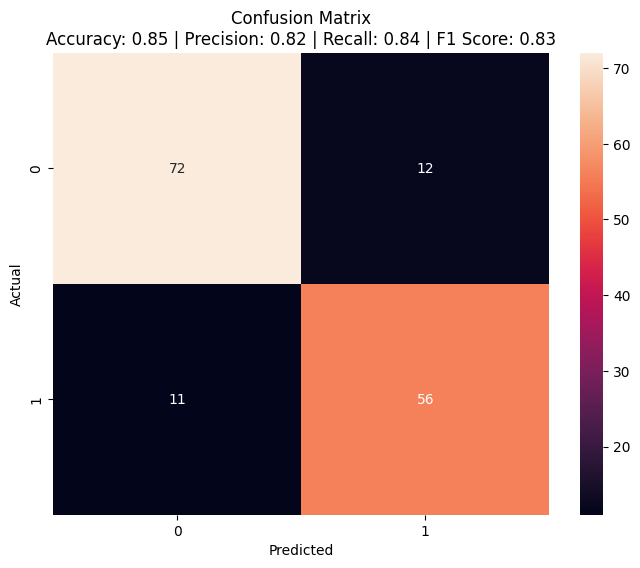
* + ค่า Mean Accuracy มีค่าต่ำ (ประมาณ 0.56) ที่ K=1K = 1K=1.
  + ความแม่นยำเพิ่มขึ้นอย่างรวดเร็วเมื่อเพิ่ม KKK เป็น 3 หรือ 4.

**2. ช่วงค่ากลาง (K=5K = 5K=5 ถึง K=14K = 14K=14):**

* มีการขึ้นและลงของความแม่นยำเล็กน้อย.
* ความแม่นยำค่อนข้างนิ่งในช่วง K=9K = 9K=9 ถึง K=13K = 13K=13.

**3. ช่วงปลาย (K=15K = 15K=15 ถึง K=20K = 20K=20):**

* ค่า Mean Accuracy มีแนวโน้มเพิ่มสูงขึ้นอย่างชัดเจน.
* จุดสูงสุดของความแม่นยำอยู่ที่ประมาณ K=19K = 19K=19, มีค่าเฉลี่ยประมาณ 0.72.
* แสดงผลการทำนายในรูปแบบ Confusion Matrix



**แกน X (Predicted)** แสดงค่าผลลัพธ์ที่โมเดลทำนาย (0 หรือ 1)

**แกน Y (Actual)** แสดงค่าจริงที่เกิดขึ้นในชุดข้อมูล (0 หรือ 1)

โดยค่าตัวเลขในแต่ละช่องแสดงจำนวนตัวอย่างที่โมเดลจัดให้อยู่ในกลุ่มนั้น

**ช่อง (0, 0):** 72

ตัวอย่างที่โมเดลทำนายว่าเป็นคลาส 0 และค่าจริงคือคลาส 0 (True Negatives - TN)

**ช่อง (0, 1):** 12

ตัวอย่างที่โมเดลทำนายว่าเป็นคลาส 1 แต่ค่าจริงคือคลาส 0 (False Positives - FP)

**ช่อง (1, 0):** 11

ตัวอย่างที่โมเดลทำนายว่าเป็นคลาส 0 แต่ค่าจริงคือคลาส 1 (False Negatives - FN)

**ช่อง (1, 1):** 56

ตัวอย่างที่โมเดลทำนายว่าเป็นคลาส 1 และค่าจริงคือคลาส 1 (True Positives - TP)

**โดยที่มีค่า Accuracy** (ความแม่นยำ): 0.85 , **Precision** (ความถูกต้องของการทำนายคลาส 1): 0.82, **Recall (ความครอบคลุมของการทำนายคลาส 1):** 0.84, **F1 Score** (คะแนนเฉลี่ยถ่วงน้ำหนักของ Precision และ Recall): 0.83