

# การพัฒนา AI จัดกลุ่มงานวิจัยด้วย Python

เรียนรู้และฝึกปฏิบัติการสร้าง Model Classification ด้วย AI เพื่อจัดกลุ่มงานวิจัยอย่างมีประสิทธิภาพ เริ่มต้นจากพื้นฐาน Python ด้วยเครื่องมือ Google Colab เพื่อจัดกลุ่มผลงานวิจัยในระบบ TNRR ตามสาขาวิชาการ OECD

## แบบทดสอบ



<https://forms.gle/umMTbBz4bKLobX5H6>

## เอกสารประกอบการอบรม



<https://shorturl.at/h3gz5>

# แบบทดสอบ

AI คืออะไร?

AI หรือปัญญาประดิษฐ์ คือการทำให้คอมพิวเตอร์คิดและเรียนรู้ได้เหมือนมนุษย์

Classification คืออะไร?

Classification คือการแบ่งข้อมูลออกเป็นกลุ่มต่างๆ ตามคุณสมบัติที่กำหนด

ทำไมต้อง Classification?

ใช้ในการตัดสินใจ คาดการณ์ และแก้ปัญหาต่างๆ ได้อย่างมีประสิทธิภาพ

# พื้นฐานภาษา Python ที่ควรรู้

1

ตัวแปรและชนิดข้อมูล

รู้จักการประกาศตัวแปรและชนิด  
ข้อมูลต่างๆ เช่น ตัวเลข ข้อความ

2

Control Flow

เรียนรู้การควบคุมการทำงานของ  
โปรแกรมด้วย if, else, for, while

3

ฟังก์ชัน

สร้างและเรียกใช้ฟังก์ชันเพื่อจัดการโค้ดให้เป็นระเบียบ

# รู้จักกับ Google Colaboratory (Colab)



Cloud-Based

Colab คือ Cloud IDE ที่  
ให้คุณเขียนและรันโค้ด  
Python ได้ฟรี



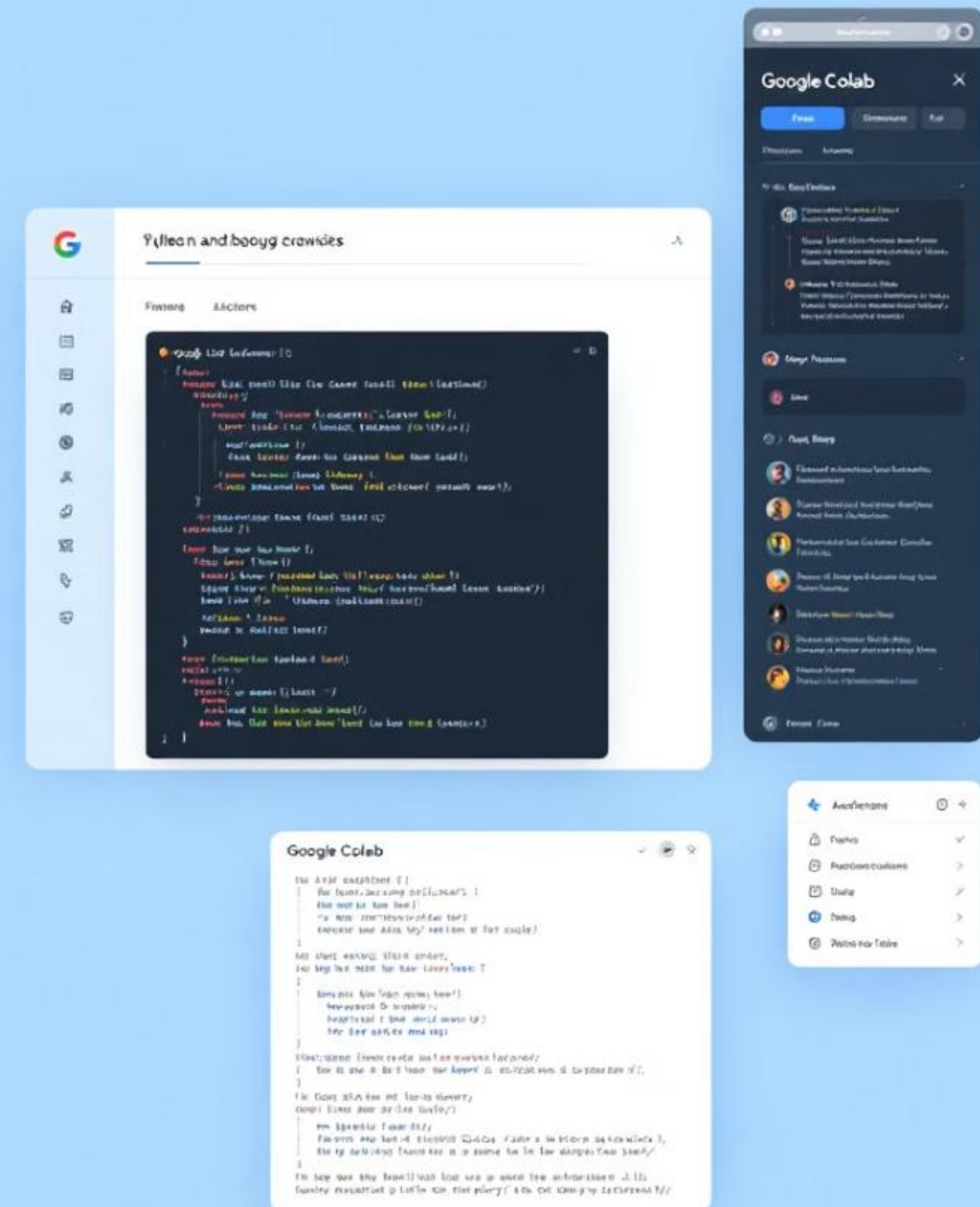
Collaboration

แชร์โค้ดและทำงานร่วมกับ  
ผู้อื่นได้อย่างง่ายดาย



Free Resources

ใช้งาน GPU และ TPU ฟรี  
สำหรับการ Train Model





# Library ที่จำเป็น

NumPy

จัดการข้อมูลตัวเลขและ Array อย่างมีประสิทธิภาพ

Pandas

จัดการข้อมูลในรูปแบบตาราง  
(Data frame)

Scikit-learn

Library หลักสำหรับการสร้าง Machine Learning Model

AttaCut

Library สำหรับตัดคำภาษาไทย

Joblib & Pickle

ใช้ในการบันทึกและโหลด Model





# เตรียมข้อมูลสำหรับการ Classification

เก็บรวบรวมข้อมูล

รวบรวมข้อมูลที่เกี่ยวข้องกับปัญหาที่ต้องการแก้ไข

ทำความสะอาดข้อมูล

จัดการกับ Missing Value และ Outlier

แบ่งข้อมูล

แบ่งข้อมูลเป็น Train Set และ Test Set



# สร้าง Model Classification ด้วย scikit-learn



1

เลือก Algorithm

เลือก Algorithm ที่เหมาะสมกับประเภทข้อมูล

2

สร้าง Model

สร้าง Model โดยใช้ Algorithm ที่เลือก

3

Train Model

Train Model ด้วย Train Set



# Metrics สำหรับประเมินประสิทธิภาพ Model

1

ค่า Recall

วัดความสามารถในการหาผลลัพธ์ที่เป็นจริง

2

ค่า Precision

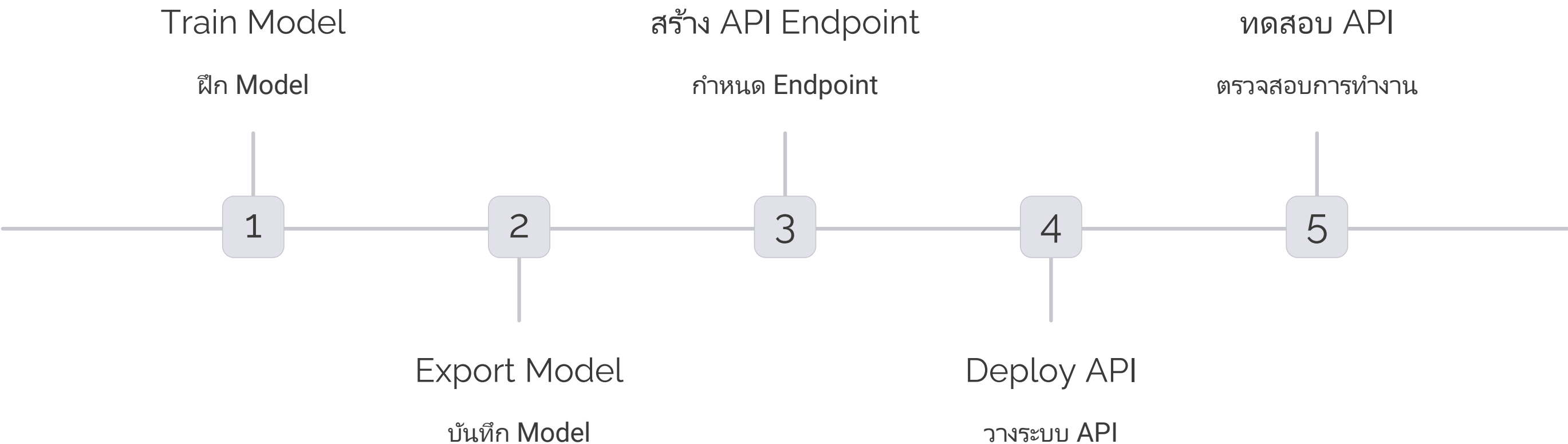
วัดความแม่นยำของผลลัพธ์ที่ทำนาย

3

F1-Score

ค่าเฉลี่ย Harmonic ของ Recall และ Precision

# Export Model เพื่อใช้งานเป็น API



# การนำ Model ไปใช้ในระบบ TNRR

Model Classification ที่สร้างขึ้น สามารถนำไปประยุกต์ใช้ในระบบ Thai National Research Repository (TNRR) เพื่อจัดกลุ่มผลงานวิจัยตามสาขาวิชาการ OECD ได้อย่างมีประสิทธิภาพ

1

จัดกลุ่มผลงานวิจัยอัตโนมัติ

ช่วยลดภาระงานในการจัดกลุ่มผลงานวิจัย  
ด้วยตนเอง

2

เพิ่มความแม่นยำในการจัดกลุ่ม  
ผลงานวิจัย

ลดความผิดพลาดจากการตัดสินใจของ  
มนุษย์

3

สนับสนุนการวิเคราะห์ข้อมูล

ทำให้สามารถวิเคราะห์ข้อมูลผลงานวิจัย  
ตามสาขาวิชาการได้อย่างรวดเร็ว