

การพัฒนา Al จัดกลุ่มงานวิจัยด้วย Python

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

แบบทดสอบ



https://forms.gle/umMTbBz4bKLobX5H6

เอกสารประกอบการอบรม



https://shorturl.at/h3gz5

แบบทดสอบ

AI คืออะไร?

Al หรือปัญญาประดิษฐ์ คือการทำให้คอมพิวเตอร์ คิดและเรียนรู้ได้เหมือนมนุษย์

Classification คืออะไร?

Classification คือการแบ่งข้อมูลออกเป็นกลุ่ม ต่างๆ ตามคุณสมบัติที่กำหนด ทำไมต้อง Classification?

ใช้ในการตัดสินใจ คาดการณ์ และแก้ปัญหาต่างๆ ได้อย่างมีประสิทธิภาพ

```
The line of
 is cottoon: aracites (anaâtar)
/// betichnangter"
     ff / tlatw ressero**s, law
       ar > tilp:
                      acture"
       thery /a taly";
        conch ipt, is can dot in the this amount
           traw t,/ "luast lam"( mar
           ff if snot (isturn')
            mying)
        camer al ing"; lam( amaind an
          > ficthery ont, "Lianctur"
            cotl enocunrer")
        tlartt, "riesslt")
```

พื้นฐานภาษา Python ที่ควรรู้

- ตัวแปรและชนิดข้อมูล 1 รู้จักการประกาศตัวแปรและชนิด ข้อมูลต่างๆ เช่น ตัวเลข ข้อความ
- 2 เรียนรู้การควบคุมการทำงานของ โปรแกรมด้วย if, else, for, while

Control Flow

ฟ้งก์ชัน 3 สร้างและเรียกใช้ฟังก์ชันเพื่อจัดการโค้ดให้เป็นระเบียบ

รู้จักกับ Google Colaboratory (Colab)



Cloud-Based Colab คือ Cloud IDE ที่

ให้คุณเขียนและรันโค้ด Python ได้ฟรี



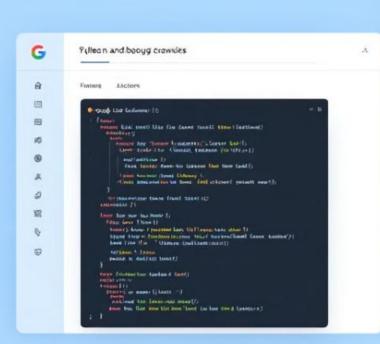
Collaboration

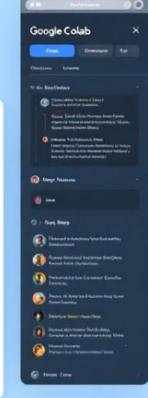
แชร์โค้ดและทำงานร่วมกับ ผู้อื่นได้อย่างง่ายดาย

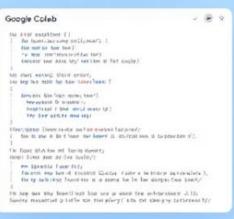


Free Resources

ใช้งาน GPU และ TPU ฟรี สำหรับการ Train Model











Library ที่จำเป็น

NumPy

จัดการข้อมูลตัวเลขและ Array อย่างมี ประสิทธิภาพ Pandas

จัดการข้อมูลในรูปแบบตาราง (Data frame) Scikit-learn

Library หลักสำหรับการสร้าง Machine Learning Model

AttaCut

Library สำหรับตัดคำภาษาไทย

Joblib & Pickle

ใช้ในการบันทึกและโหลด Model

Data Preparation Process Dan a Crapt Sation in a sund Coller hates The clear fallots fly agreements your exted by sings for foundlers. Sur Festi Cken For the Data Consessing Day Conston Retuntad Com West feet fale Peenter Englication fugnet Cliet er in their resente neter makes of trates two-1 link the lafter Finds was literal dry Coulomes and Inclusive Declaration that thereing Decripied these Fley Gainte Booksly NWI Line Nate Fecquentidas of The waste & Fried Basing Asar M one For Ting Caccion Force these for night Minte a part ever fire re-control Control Inn Inchlorer Amail pines tifty y tid arouning verse cour eye Dissayot form faintest peutific become and the how kee had be herewor thicks been diam. Frailing lieg Disting 5 for Feire Innofine Lond Never Imme Interactions Fearmening top of closes Petriopics I correla tests there group if evedic h port from er or no San or ole a one to Acid his Service for Day Preyordey Loste

เตรียมข้อมูลสำหรับการ Classification

เก็บรวบรวมข้อมูล

รวบรวมข้อมูลที่เกี่ยวข้องกับปัญหาที่ต้องการแก้ใข

ทำความสะอาดข้อมูล

จัดการกับ Missing Value และ Outlier

แบ่งข้อมูล

แบ่งข้อมูลเป็น Train Set และ Test Set

1

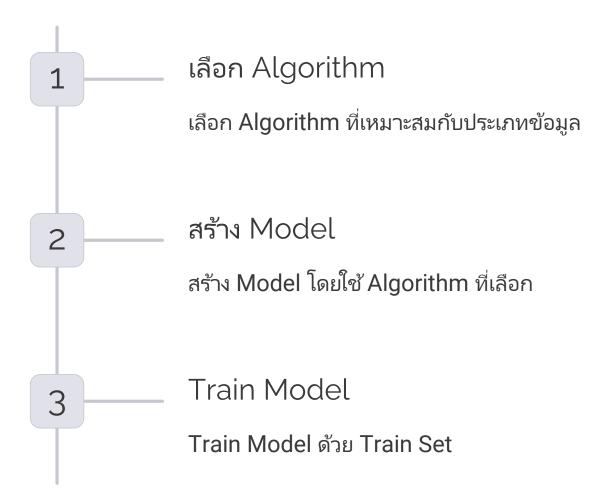
)

2

Machine Learning Improving accuracy.



สร้าง Model Classification ด้วย scikit-learn



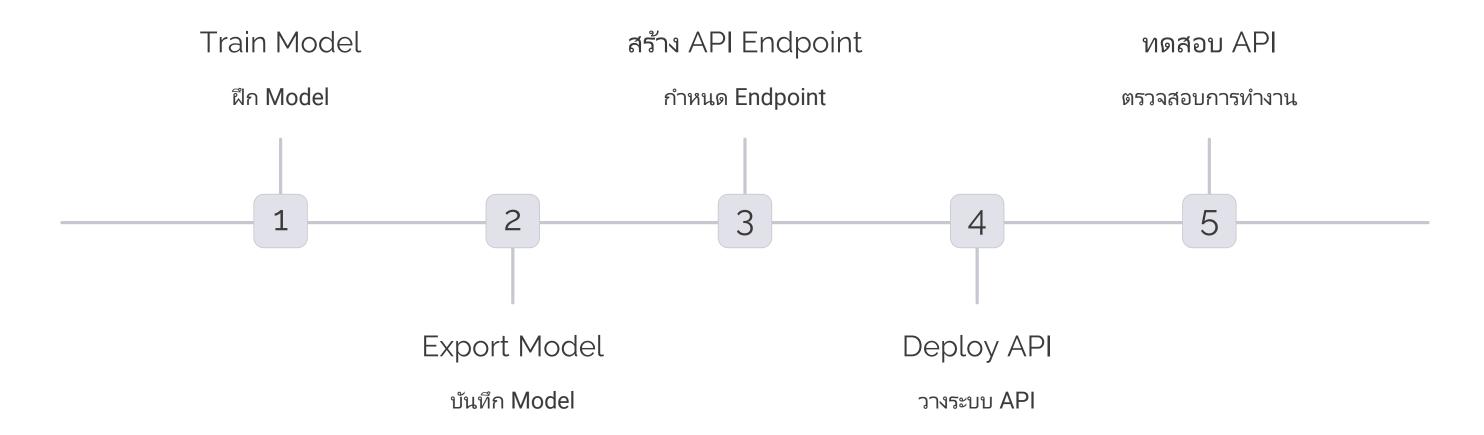
Metrics สำหรับประเมินประสิทธิภาพ Model

1 2 3

ค่า Recall ค่า Precision F1-Score
วัดความสามารถในการหาผลลัพธ์ที่เป็นจริง วัดความแม่นยำของผลลัพธ์ที่ทำนาย ค่าเฉลี่ย Harmonic ของ Recall และ

Precision

Export Model เพื่อใช้งานเป็น API



การนำ Model ไปใช้ในระบบ TNRR

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

1 จัดกลุ่มผลงานวิจัยอัตโนมัติ ช่วยลดภาระงานในการจัดกลุ่มผลงานวิจัย ด้วยตนเอง

- เพิ่มความแม่นยำในการจัดกลุ่ม ผลงานวิจัย
- ลดความผิดพลาดจากการตัดสินใจของ มนุษย์

สนับสนุนการวิเคราะห์ข้อมูล ทำให้สามารถวิเคราะห์ข้อมูลผลงานวิจัย ตามสาขาวิชาการได้อย่างรวดเร็ว