

# โครงการวิศวกรรมคอมพิวเตอร์ บทที่ 3

เรื่อง

พัฒนาโปรแกรมประยุกต์สำหรับทำนายโรคจากฟิล์มเอกซเรย์  
โดยใช้การเรียนรู้ของเครื่อง

Develop application to predict diseases from X-ray film  
using Machine Learning

โดย

นาย ธนกานต์ ปราโมทย์  
รหัสประจำตัว 5810502334

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์  
มหาวิทยาลัยเกษตรศาสตร์

อาจารย์ที่ปรึกษาโครงการ ลงนาม .....  
(ผศ.ดร.ธนาวิทย์ รักธรรมานนท์)

## บทที่ 3

### วิธีการดำเนินการศึกษา

#### 3.1 เครื่องมือที่ใช้ดำเนินงาน

เครื่องมือ อุปกรณ์ฮาร์ดแวร์ ซอฟต์แวร์ที่ใช้ในการดำเนินโครงการมีดังนี้

##### 3.1.1 ฮาร์ดแวร์ (Hardware)

##### 3.1.1.1. เครื่องคอมพิวเตอร์พกพาสำหรับการพัฒนา



Notebook HP Omen 15

หน่วยความจำ 8 GB

หน่วยความจำสำรอง 1 TB

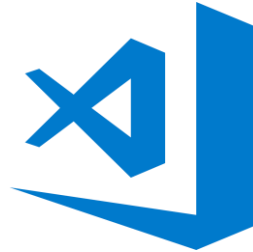
หน่วยประมวลผล 64-bit ความเร็ว 2.6 GHz Intel Core i7-6700HQ

หน่วยประมวลผลกราฟิก Nvidia GeForce GTX 960M

ระบบปฏิบัติการ Windows 10 64-bit

### 3.1.2 ซอฟต์แวร์ (Software)

3.1.2.1. Visual Studio Code ใช้ในการพัฒนาและทดสอบตัวโมเดล และโปรแกรมประยุกต์

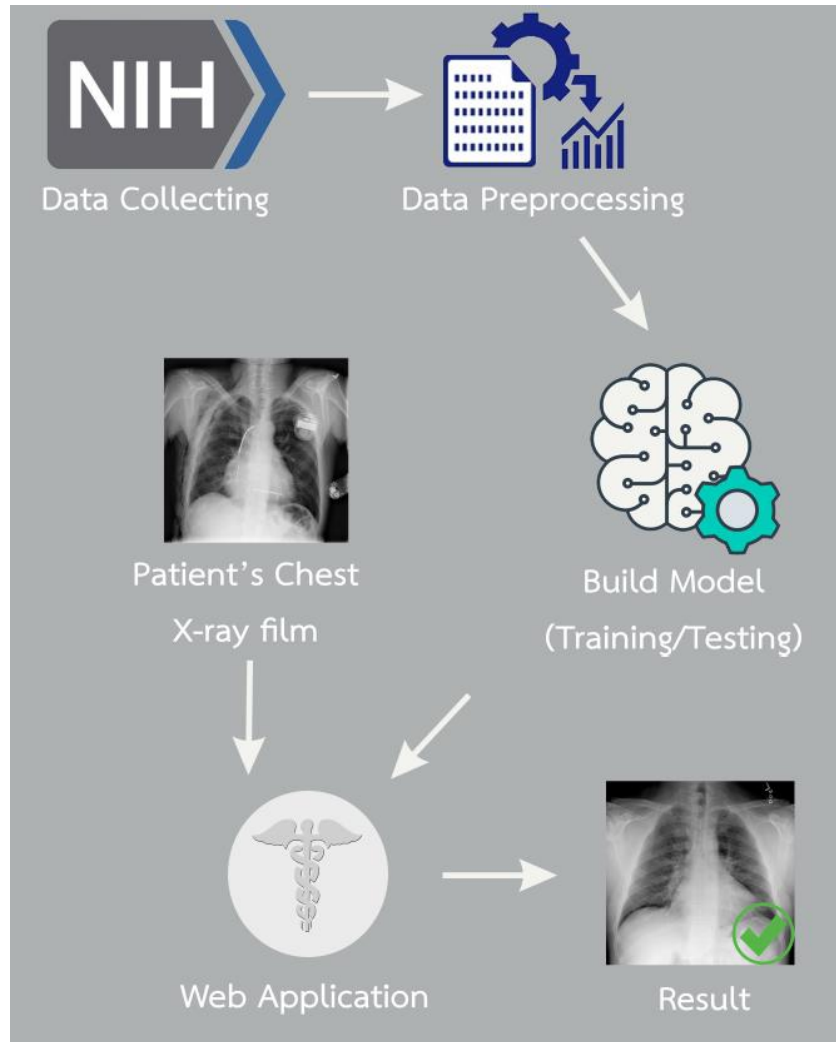


3.1.2.2. Flask Framework ใช้เป็น Framework ในการเขียน Website Application



3.1.2.3. Python เป็นภาษาที่ใช้สำหรับเขียนโปรแกรม ในส่วนของการทำ Machine Learning และคำนวณบนหน้าเว็บไซต์

### 3.2 โครงสร้างของระบบ



กระบวนการทำงานของระบบมีส่วนประกอบดังนี้

3.2.1. การเก็บข้อมูลฟิล์มเอกซเรย์จาก NIH

3.2.2. ทำการจัดเตรียมข้อมูลให้เหมาะสมและพร้อมต่อการสร้างโมเดล

3.2.3.

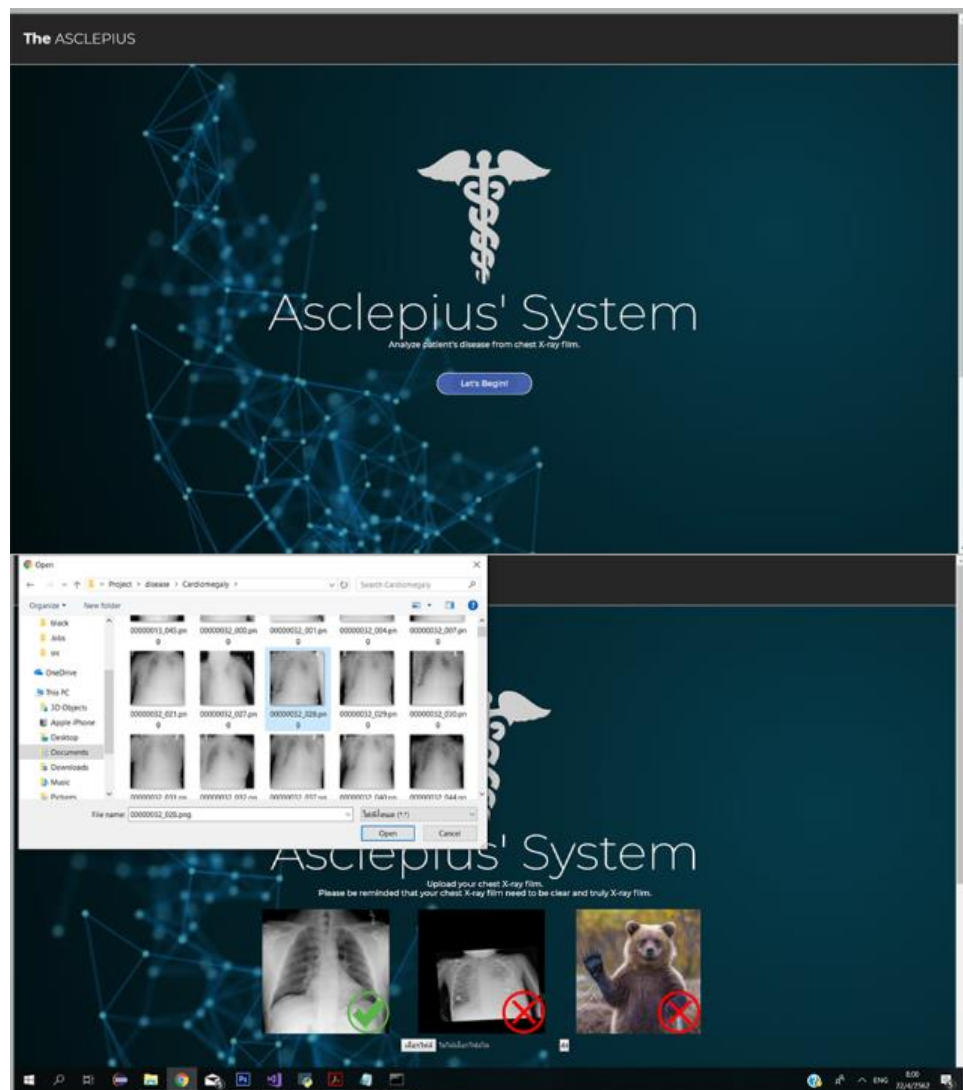
การสร้างโมเดลด้วยการใช้เรียนรู้ของเครื่องเพื่อพัฒนาโมเดลสำหรับทำนายโรคจากฟิล์มเอกซเรย์

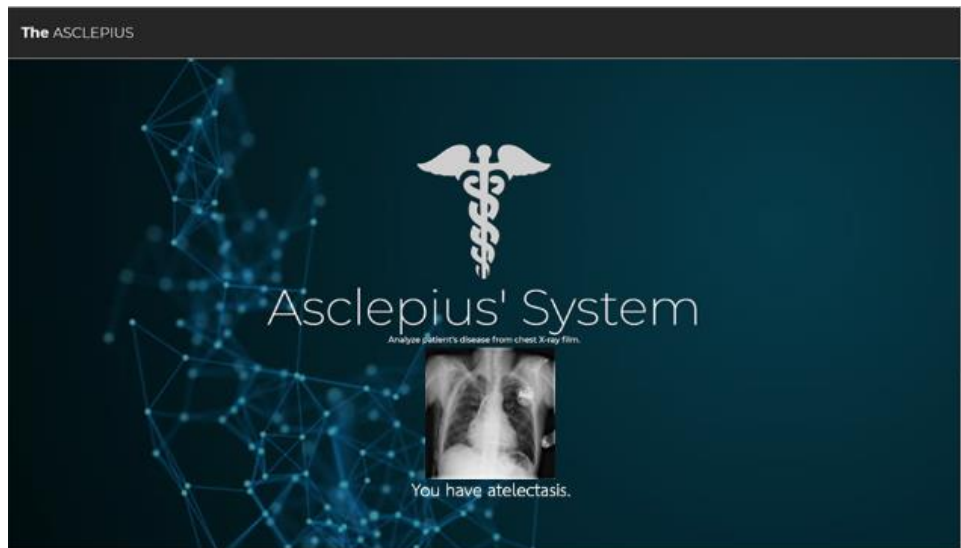
3.2.4. สร้าง Web Application

เพื่อให้สามารถทำนายฟิล์มเอกซเรย์ที่ต้องการผ่านหน้าเว็บได้

### 3.3 รายละเอียดของระบบที่พัฒนา

#### 3.3.1. การออกแบบ User Interface





### 3.4 ขั้นตอนการพัฒนา

3.4.1. รวบรวมข้อมูลฟิล์มเอกซเรย์ช่วงอกจาก National Institute of Health (NIH)

3.4.2. นำข้อมูลเข้าสู่กระบวนการจัดเตรียมข้อมูล (Preprocessing) ดังนี้

3.4.2.1. จำแนกฟิล์มเอกซเรย์ไปตามประเภทของโรค

3.4.2.2. จำแนกข้อมูลสอนและข้อมูลทดสอบตามนี้

3.4.2.2.1. สำหรับโมเดลทำนายแยกแต่ละโรค มีข้อมูลสอน 90% และข้อมูลทดสอบ 10% จากข้อมูลทั้งหมดของแต่ละประเภท

3.4.2.2.2. สำหรับโมเดลทำนายรวมโรค โรคที่มีรูป  $\geq 5000$

จะมีข้อมูลสอน 4500 รูป ส่วนโรคที่มีรูป  $< 5000$

จะมีข้อมูลสอนเท่ากับ

จำนวนรูปทั้งหมดของโรคนั้นๆ หารด้วย 500 โดยทั้งสองจะมี

ข้อมูลสอนเท่ากันคือ 500 รูป

3.4.2.3.

ทำการตัดและย่อรูปข้อมูลเฉพาะช่วงช่องอกเท่านั้นพร้อมกำจัดข้อมูลNoise

3.4.3. ทำการเรียนรู้จากข้อมูลสอนผ่าน Tensorflow โดยใช้ โมเดล Convolutional-Neural Network (CNN) สำหรับโมเดลแต่ละโรคเป็น CNN

พื้นฐานที่สร้างขึ้นมาเอง

และสำหรับโมเดลรวมโรคเป็นการประยุกต์จาก

ResNet50 ของ Keras

#### 3.4.4.

ทำการสร้างเดโมเว็บแอปพลิเคชันสำหรับใช้ทำนายโรคจากฟิล์มเอกซเรย์ช่วงอกที่ผู้ใช้  
อัปโหลดขึ้นไป โดยคำนวณผลลัพธ์จากทั้งสองโมเดลเพื่อให้ได้ขอบเขตโรคที่ดีที่สุด