

บทที่ 1

บทนำ

1.1 ความเป็นมาของปัญหา

ปัญญาประดิษฐ์มีบทบาทสำคัญต่อการบริหารและการจัดการ โดยการจัดเก็บชุดข้อมูลถือเป็นปัจจัยสำคัญในการที่จะนำไปสู่การพยากรณ์ หรือการสร้างโมเดลที่ช่วยซัพพอร์ตงานในองค์กรได้ ปัจจุบันมีองค์กรหลายองค์กรที่พยายามเก็บรวบรวมข้อมูลในหลายรูปแบบ เช่น รูปภาพ เสียง สัญญาณ และข้อความ โดยข้อมูลรูปภาพถือว่าเป็นชุดข้อมูลที่สำคัญที่ช่วยมาแก้ปัญหาในงานด้านธุรกิจ เช่น การรู้จำใบหน้า เพื่อยืนยันตัวตน การทำนายวัตถุ เพื่อใช้ในการตรวจสอบจำแนกข้อผิดพลาดของผลผลิต หรือผลิตภัณฑ์ต่างๆ เป็นต้น แต่ปัจจุบันคนส่วนใหญ่ยังเก็บข้อมูลในรูปแบบไฟล์ ทำให้เกิดปัญหาเกิดความล่าช้า หรือใช้เวลานาน ในกรณีการศึกษานี้ ผู้จัดทำจึงได้จัดเก็บเป็นรูปภาพจำพวก Classification รวบรวมข้อมูลให้กับคอมพิวเตอร์ และส่งไปให้ Machine Learning เพื่อจดจำการเรียนรู้

จากปัญหาดังกล่าว ผู้จัดทำจึงมีแนวคิดที่จะพัฒนา ระบบจัดเก็บชุดข้อมูลรูปภาพ โดยผู้พัฒนาจะจัดทำเว็บไซต์จัดเก็บชุดข้อมูล ที่สามารถค้นหา วิเคราะห์ สืบค้น และเรียนรู้เพิ่มเติมเกี่ยวกับข้อมูลประเภทรูปภาพ เป็นต้น เพื่อช่วยให้ง่ายต่อการทำนายผลและคาดการณ์ง่ายต่อการวางแผน การประเมินผลในหลายๆด้าน และอำนวยความสะดวกในกระบวนการวิเคราะห์และคัดกรองข้อมูล เช่น การวิเคราะห์โครงสร้างข้อมูล หรือการคัดกรองข้อมูลที่ไม่เหมาะสมหรือไม่เหมือนจริง

1.2 วัตถุประสงค์

1.2.1 เพื่อพัฒนาเว็บไซต์สำหรับจัดเก็บชุดข้อมูลที่สามารถนำข้อมูลประเภทรูปภาพไปใช้งานในการทดลองเพื่อทำนายผลด้วยโมเดลการจำแนกประเภทข้อมูล Classification

1.2.2 เพื่อช่วยอำนวยความสะดวก และลดโอกาสในการเกิดข้อผิดพลาดของชุดข้อมูลรูปภาพ

1.2.3 เพื่อให้ผู้ใช้งานสามารถจัดเก็บข้อมูล และเผยแพร่ข้อมูลต่อสาธารณชนที่สามารถนำไปใช้ประโยชน์

1.3 ขอบเขตโครงการ

ระบบจัดเก็บชุดข้อมูล มีการแบ่งงานออกเป็น 3 ส่วนหลักๆ คือ ส่วนการจัดการข้อมูลพื้นฐาน ส่วนการจัดเก็บชุดข้อมูล และส่วนสุดท้ายการออกรายงาน โดยแบ่งกลุ่มผู้ใช้งานออกเป็น 2 กลุ่ม ได้แก่ ผู้ดูแลระบบ และผู้ใช้งานทั่วไป ซึ่งแต่ละส่วนมีรายละเอียด ดังนี้

1.3.1 ส่วนการจัดการข้อมูลพื้นฐาน

ส่วนของผู้ดูแลระบบ

- 1.) สามารถเพิ่ม แก้ไข ค้นหา ข้อมูลผู้ใช้งานระบบ
- 2.) สามารถเพิ่ม แก้ไข ค้นหา ข้อมูลประเภทชุดข้อมูล

ส่วนของผู้ใช้งานทั่วไป

- 1.) สามารถ เพิ่ม แก้ไข ข้อมูลของผู้ใช้งานระบบ

1.3.2 ส่วนการจัดเก็บชุดข้อมูล

ส่วนของผู้ดูแลระบบ

- 1.) สามารถ ค้นหา เพิ่ม แก้ไข ข้อมูลการจัดเก็บชุดข้อมูลรูปภาพทีละไฟล์
- 2.) สามารถ ค้นหา เพิ่ม แก้ไข ข้อมูลการจัดเก็บชุดข้อมูลรูปภาพหลาย
- 3.) สามารถกำหนดสิทธิ์การเผยแพร่ชุดข้อมูล
- 4.) สามารถกำหนดขนาดTraining และ Testing Datasets เพื่อดาวนโหลดชุดข้อมูล พร้อมไฟล์เฉลยในรูปแบบ CSV

ส่วนของผู้ใช้งานทั่วไป

- 1.) สามารถ ค้นหา เพิ่ม แก้ไข ข้อมูลการจัดเก็บชุดข้อมูลรูปภาพทีละไฟล์
- 2.) สามารถ ค้นหา เพิ่ม แก้ไข ข้อมูลการจัดเก็บชุดข้อมูลรูปภาพหลายไฟล์
- 3.) สามารถกำหนดขนาดTraining และ Testing Datasets เพื่อดาวนโหลดชุดข้อมูลพร้อมไฟล์เฉลยในรูปแบบ CSV

1.3.3 ออกรายงาน

ส่วนของผู้ดูแลระบบ

- 1.) รายงานสรุปผลดาวนโหลดชุดข้อมูล
- 2.) รายงานข้อมูลผู้ใช้งาน
- 3.) รายงานยอดผู้เข้าชมชุดข้อมูล

ส่วนของผู้ใช้งานทั่วไป

- 1.) รายงานชุดข้อมูลที่ดาวนโหลด

1.3.4 ขอบเขตด้านฮาร์ดแวร์

| | |
|---|---------|
| 1.) Laptop ASUS TUF Gaming A15 | 1 ชุด |
| 2.) หน่วยประมวลผลกลาง (CPU) AMD Ryzen 7 | 1 หน่วย |
| 3.) หน่วยความจำหลัก (SDRAM) | 16 GB |
| 4.) หน่วยความจำสำรอง (Hard disk) | 512 GB |
| 5.) จอภาพ 15.6 นิ้ว | 1 หน่วย |
| 6.) เมาส์ | 1 ชุด |

1.3.5 ขอบเขตด้านซอฟต์แวร์

- 1.) ระบบปฏิบัติการ Window 11
- 2.) ฐานข้อมูล MySQL
- 3.) เครื่องมือสำหรับเขียนโปรแกรม Visual Studio Code
- 4.) ภาษา PHP
- 5.) โปรแกรม xampp v3.3.0
- 6.) โปรแกรมสำหรับการแสดงผล Microsoft Edge

1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1.4.1 ช่วยให้สามารถนำข้อมูลประเภทรูปภาพไปใช้งานเพื่อทำนายผลด้วยโมเดลการจำแนกประเภทข้อมูล Classification
- 1.4.2 ช่วยอำนวยความสะดวก และลดโอกาสในการเกิดข้อผิดพลาดของชุดข้อมูลรูปภาพ
- 1.4.3 ช่วยให้การจัดการเก็บชุดข้อมูล และสามารถดาวน์โหลดเผยแพร่ต่อสาธารณชน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการจัดทำโครงงานครั้งนี้ ผู้เขียนมีจุดมุ่งหมายที่จะศึกษาถึงกระบวนการจัดเก็บข้อมูลของการทำ Machine learning โดยมุ่งโฟกัสของการรวบรวม Classification โดยทฤษฎีและแนวทางที่จะนำมาศึกษาในการทำโครงงานครั้งนี้ ประกอบด้วย

2.1 กระบวนการเก็บข้อมูล

2.2 ภาพดิจิทัล

2.3 กระบวนการ Classification

2.4 พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล (PDPA)

2.1 กระบวนการเก็บข้อมูล

ระบบเก็บข้อมูลรูปภาพ (Image Storage System) คือระบบที่ออกแบบมาเพื่อจัดการและเก็บรูปภาพต่าง ๆ ในรูปแบบอิเล็กทรอนิกส์หรือดิจิทัล เพื่อให้สามารถเข้าถึงและค้นหาข้อมูลรูปภาพได้อย่างรวดเร็วและมีประสิทธิภาพ เราสามารถใช้ระบบเก็บข้อมูลรูปภาพในหลายวัตถุประสงค์ เช่น เก็บภาพสำหรับเว็บไซต์, แชรรูปภาพในสื่อสังคมออนไลน์, การจัดการภาพในองค์กรหรือธุรกิจ มีกระบวนการทำงาน ดังนี้



- 1.) การจัดเก็บ: ระบบนี้ช่วยในการจัดเก็บรูปภาพในรูปแบบดิจิทัล โดยปกติจะมีการจัดเก็บในเซิร์ฟเวอร์หรือคลาวด์เก็บข้อมูล (cloud storage) เพื่อความคงทนและความปลอดภัย
- 2.) การค้นหา: ระบบเก็บข้อมูลรูปภาพจะมีการจัดการข้อมูลเช่นป้ายกำกับ (tags) หรือคีย์เวิร์ดเพื่อช่วยในการค้นหารูปภาพที่ต้องการ

- 3.) การแบ่งปัน: ระบบนี้มีความสามารถในการแบ่งปันรูปภาพกับผู้อื่น ๑ ผ่านลิงก์หรือการแชร์ในโซเชียลมีเดีย
- 4.) การจัดการ: ระบบนี้ช่วยในการจัดการและแก้ไขรูปภาพ เช่นการปรับขนาด, การตัดต่อ, การปรับแสงและสี, และการทำภาพถ่าย
- 5.) ความปลอดภัย: การควบคุมการเข้าถึงข้อมูลรูปภาพเพื่อให้มีความปลอดภัยและความเป็นส่วนตัว
- 6.) การสำรองข้อมูล: การสำรองข้อมูลรูปภาพเพื่อป้องกันข้อมูลสูญหายหรือเสียหาย

2.2 ภาพดิจิทัล



2.2.1 ภาพดิจิทัลสี (Color Digital Image)

- 1.) สี: ภาพดิจิทัลสีมีความสามารถในการแสดงสี โดยประกอบไปด้วยสีหลายสีที่สอดคล้องกันในแต่ละพิกเซล
- 2.) ความสว่างและเข้มสี: สีและความสว่างของแต่ละพิกเซลสามารถระบุความสว่างและความเข้มของสี ทำให้สามารถสร้างรูปภาพที่สีสดใสและสวยงาม
- 3.) รูปแบบไฟล์: ภาพดิจิทัลสีสามารถบันทึกในรูปแบบไฟล์ที่รองรับการจัดเก็บสี เช่น JPEG, PNG, TIFF
- 4.) ความสามารถในการแก้ไข: ภาพดิจิทัลสีมีความสามารถในการแก้ไขสี, ปรับสี, และปรับแต่งความสว่าง-เข้มของภาพอย่างอิสระ

2.2.2 ภาพดิจิทัลขาว-ดำ (Black and White Digital Image)

- 1.) ขาว-ดำ: ภาพดิจิทัลขาว-ดำมีเพียงสองสีเท่านั้น: ขาวและดำ ซึ่งอาจจะดับเทา (grayscale) ได้หากมีการใช้ระดับสีเทาในภาพ
- 2.) ความสว่างและเข้ม: ในภาพขาว-ดำ, ความสว่างและความเข้มจะถูกสะท้อนผ่านการใช้ระดับเทา (grayscale) แทนการใช้สี
- 3.) แบบไฟล์: ภาพดิจิทัลขาว-ดำสามารถบันทึกในรูปแบบไฟล์ที่ไม่จำเป็นต้องรองรับสีเช่น BMP, TIFF, หรือ PNG
- 4.) ความสามารถในการแก้ไข: ภาพขาว-ดำสามารถใช้เทคนิคการปรับความสว่างและความเข้มเพื่อเปลี่ยนรูปแบบและรูปลักษณ์ของภาพ

2.3 กระบวนการ Classification

Image classification (การจำแนกภาพ) คือ กระบวนการที่ใช้เทคนิค และ ศาสตร์ด้าน Machine learning ในการแยกแยะและจำแนกภาพเป็น กลุ่มหรือหมวดหมู่ต่าง ๆ เช่น การจำแนกว่าภาพนั้นเป็น สุนัข หรือ แมว

2.3.1 ขั้นตอนในการทำ Image Classification

- 1.) การเตรียมข้อมูลที่ต้องการที่จะจำแนกประกอบด้วย รูปภาพต่างๆ และข้อมูลที่ระบุประเภทเพื่อทำการจำแนก(label)
- 2.) แบ่งข้อมูลเพื่อเป็นชุดฝึกฝน(training set) และ ชุดทดสอบ(test set)
- 3.) ทำการกำหนดและฝึกฝน โมเดล(model) โดยการฝึกฝน โมเดลจะใช้ข้อมูลฝึกฝน (training) ในการฝึก
- 4.) ประเมินและปรับปรุง โดยจะใช้ข้อมูลทดสอบ(test set) เพื่อทดสอบประสิทธิภาพของ โมเดล โดยการประเมินความแม่นยำ เราจะใช้ accuracy , precision , recall และ F1-score และปรับปรุงโมเดลให้ได้ประสิทธิภาพ
- 5.) ใช้ model ในการจำแนกภาพ



2.4 พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล (PDPA)



PDPA (Personal Data Protection Act) คือ ข้อบังคับสำหรับองค์กรที่จัดเก็บประมวลผล และเผยแพร่ข้อมูลส่วนบุคคล โดยจุดมุ่งหมายที่สำคัญในกฎหมายฉบับนี้ จะเป็นการจัดการแนวทางการเก็บข้อมูลส่วนบุคคล พร้อมทั้งยกระดับความปลอดภัยในการเก็บข้อมูล

2.4.1 รายละเอียดข้อบังคับ PDPA

● การขออนุญาตเก็บข้อมูล

การขออนุญาตข้อมูลส่วนบุคคลจากเจ้าของข้อมูลนั้นจะต้องมีเงื่อนไขการยินยอมที่ชัดเจน ซึ่งหมายความว่าจุดประสงค์ของการขอเก็บข้อมูลต้องเข้าใจได้ง่าย มีการกำหนดระยะเวลาการเก็บข้อมูล มีช่องที่เหมาะสมสำหรับการติดต่อ DPO ให้ชัดเจน เพื่อไม่ให้เจ้าของข้อมูลเกิดความสับสน เพื่อให้เจ้าของข้อมูลส่วนบุคคลสามารถถอนหรือลบเมื่อไหร่ก็ได้

- **ระยะเวลาในการแจ้งเตือนหากว่ามีข้อมูลส่วนบุคคลรั่วไหล**

หาก Data Subject ขอสำเนาข้อมูลของตนในองค์กรจะต้องออกแบบฟอร์ม หรือระบุช่องทางการขอข้อมูลที่มีรายละเอียดครบถ้วน ในส่วนนี้องค์กรต้องส่งข้อมูลที่มีอยู่ทั้งหมดเกี่ยวกับเจ้าของข้อมูล รวมถึงวิธีต่าง ๆ ที่คุณใช้ข้อมูลของเจ้าของข้อมูลด้วย

- **สิทธิในการขอลบข้อมูลของเจ้าของข้อมูล**

เมื่อข้อมูลส่วนบุคคลมีการใช้งานจนบรรลุวัตถุประสงค์ หรือครบระยะเวลาในการเก็บแล้ว ผู้ควบคุมข้อมูล (Data Controller) จะต้องทำการลบข้อมูลส่วนบุคคลทั้งหมดตามที่ระบุไว้ใน Consent ขออนุญาตเก็บข้อมูล อีกทั้งเจ้าของข้อมูลมีสิทธิที่จะขอลบข้อมูลเมื่อไหร่ก็ได้

- **สิทธิในถ่ายโอน หรือเลือกใช้ข้อมูลส่วนบุคคล**

เจ้าของข้อมูลส่วนบุคคล (Data Owner) สามารถให้สิทธิ์แก่องค์กรนำข้อมูลไปใช้กับอีกองค์กรในเครือได้โดยการใช้สิทธิ์นั้นจะต้องไม่ขัดต่อกฎหมาย สัญญา หรือละเมิดสิทธิของผู้อื่น

- **ออกแบบระบบ และรูปแบบการเก็บข้อมูลที่มีความปลอดภัย**

ในตัว พ.ร.บ. คุ้มครองข้อมูลส่วนบุคคลมีข้อกำหนดที่ว่า องค์กรที่มีการเก็บประมวลผล และใช้ข้อมูลส่วนบุคคลต้องมีระบบการเก็บข้อมูลที่มีประสิทธิภาพ และชัดเจน สามารถ Monitor และแจ้งเตือนหากมีการลวง หรือละเมิดข้อมูลส่วนบุคคลได้

- **แต่งตั้งเจ้าหน้าที่คุ้มครองข้อมูลที่มีศักยภาพ**

หากองค์กรมีการจัดเก็บข้อมูลส่วนบุคคลเป็นปริมาณมาก ๆ จำเป็นจะต้องมีตำแหน่ง DPO (Data Protection Officer) ไว้กำกับดูแลข้อมูลส่วนบุคคล ทั้งนี้ก็ใช้ว่าทุกองค์กรจะต้องมี DPO เสมอไป ขึ้นอยู่กับขนาดของข้อมูลที่เก็บ และกระบวนการภายใน



บทที่ 3

วิธีการดำเนินงาน

จากทฤษฎีที่เกี่ยวข้องนำมาซึ่งการออกแบบของระบบ โดยผู้เขียนได้ทำการวิเคราะห์และออกแบบระบบโดยใช้คอนเท็กซ์ไดอะแกรม (Context Diagram) และออกแบบฐานข้อมูลด้วยแผนภาพ Entity Relationship Model