

# โครงการวิทยาศาสตร์

การทำนายและเปรียบเทียบความสนใจในการจดบันทึก  
ของผู้เรียนจากมัลติมีเดียโดยใช้แบบจำลองการเรียนรู้ของ  
เครื่องเพื่อเพิ่มประสิทธิภาพในการเรียนรู้

Predicting and comparing learners' interest in  
note-taking from multimedia using a machine  
learning model to enhance learning efficiency

นายโชติพิสิฐ อุดลสีหวัตร

รายงานนี้เป็นส่วนหนึ่งของวิชา ว30303 โครงการวิทยาศาสตร์เพิ่มเติม 3

สาขาวิชา คณิตศาสตร์และวิทยาการคำนวณ

โรงเรียนมหิดลวิทยานุสรณ์

ปีการศึกษา 2564

## ชื่อโครงงานภาษาไทย

การทำนายและเปรียบเทียบความสนใจในการจดบันทึกของผู้เรียนจากมัลติมีเดียโดยใช้แบบจำลองการเรียนรู้ของเครื่องเพื่อเพิ่มประสิทธิภาพในการเรียนรู้

## ชื่อโครงงานภาษาอังกฤษ

Predicting and comparing learners' interest in note-taking from multimedia using a machine learning model to enhance learning efficiency

จากรายวิชาโครงงานวิทยาศาสตร์ 1 ที่ได้ทำมา คณะผู้วิจัยศึกษาเกี่ยวกับประสิทธิภาพในการเรียนรู้จากการจดบันทึกบนอุปกรณ์สองประเภท คือ กระดาษ และอุปกรณ์อิเล็กทรอนิกส์ โดยทำการทดลองกับนักเรียนโรงเรียนมหิดลวิทยานุสรณ์ 61 คน แบ่งนักเรียนเป็นสองกลุ่ม แล้วให้ดูวิดีโอสองวิดีโอที่แตกต่างกันและจดด้วยอุปกรณ์ที่แตกต่างกันขณะดู ช่วงระยะเวลาที่ทดลองผู้วิจัยจะเก็บทั้งเชิงคุณภาพและปริมาณจากการทำแบบทดสอบเกี่ยวกับเนื้อหาได้จากวิดีโอและแบบสอบถาม หลังจากนั้นนำข้อมูลจากทั้งสองกลุ่มมาเปรียบเทียบกันได้ว่า ข้อมูลคะแนนของทั้งสองกลุ่มไม่มีความแตกต่างอย่างมีนัยสำคัญที่ระดับนัยสำคัญเท่ากับ 0.05 และไม่มีความสอดคล้องกันของข้อมูลเชิงคุณภาพโดยการทำ Pearson Correlation ทำให้สามารถสรุปได้ว่าการจดจากทั้งสองอุปกรณ์ไม่ส่งผลต่อประสิทธิภาพในการเรียนรู้

ในรายวิชาโครงงานวิทยาศาสตร์เพิ่มเติม 3 ผู้วิจัยได้ศึกษาข้อมูลองค์ประกอบรูปภาพในวิดีโอที่ใช้ โดยแบ่งเป็นหัวข้อย่อยและข้อมูลสถิติการจดบันทึกของแต่ละหัวข้อย่อยจากรายวิชาโครงงานวิทยาศาสตร์ 1 โดยนำข้อมูลภาพมาทำกระบวนการ Image Segmentation โดยใช้โมเดล Mask R-CNN ที่ฝึกโดยชุดข้อมูล COCO เพื่อจำแนกวัตถุแล้วนำข้อมูลไปใช้สร้างโมเดลปัญญาประดิษฐ์ที่ใช้ Features จากวัตถุที่อยู่ในภาพ เช่น ตำแหน่ง ขนาด พื้นที่ จำนวน เพื่อนำข้อมูลเหล่านั้นมาทำนายเป็นความน่าจะเป็นที่ผู้เรียนจะจดในหัวข้อนั้นโดยใช้โมเดล Random Forest ที่มีต้นไม้ 1000 ต้น ได้ว่าผลลัพธ์จากชุดข้อมูลที่ใช้ทดสอบไม่มีความแม่นยำพอ และ Features ที่ส่งผลต่อผลลัพธ์มากที่สุดคือ Features ที่เกี่ยวข้องกับพิกัดตำแหน่ง และระยะทางในแนวตั้ง ซึ่งไม่สามารถหาทางวิจัยเพื่ออธิบายผลลัพธ์ดังกล่าวได้ จึงสรุปได้ว่า ไม่สามารถสร้างโมเดลที่สามารถทำนายจากเพียงองค์ประกอบรูปภาพได้ มีความคลาดเคลื่อนของข้อมูลและจำนวนข้อมูลไม่มากพอที่จะใช้ทดสอบโมเดลได้อย่างแม่นยำได้

คำรับรอง “ข้าพเจ้าขอรับรองว่าโครงงานนี้เป็นโครงงานที่มีส่วนของการพัฒนาต่อยอดเพิ่มเติมจากวิชาโครงงานวิทยาศาสตร์ 1”

ลงชื่อ

ครูที่ปรึกษาโครงงาน.....วันที่ 14 เดือน ..ค.ค.....พ.ศ. 2564

(นางขวัญสกุล อุปพันธ์)



ใบรับรองโครงงานวิทยาศาสตร์

โรงเรียนมหิตลวิทยานุสรณ์

มัธยมศึกษาตอนปลาย

คณิตศาสตร์และวิทยาการคำนวณ

หลักสูตร

สาขาวิชา

ชื่อโครงงานภาษาไทย

การทำนายและเปรียบเทียบความสนใจในการจดบันทึกของผู้เรียนจากมัลติมีเดียโดยใช้แบบจำลองการเรียนรู้ของเครื่องเพื่อเพิ่มประสิทธิภาพในการเรียนรู้

ชื่อโครงงานภาษาอังกฤษ

Predicting and comparing learners' interest in note-taking from multimedia using a machine learning model to enhance learning efficiency

นามผู้ทำโครงงาน นายโชติพิสิฐ อตุลสีหวัตต์ ม.6/5 เลขประจำตัวนักเรียน 07967

ได้พิจารณาเห็นชอบโดย

อาจารย์ที่ปรึกษาโครงงาน ..... วันที่ 14 เดือน ตุลาคม พ.ศ.2564

(นายทศพร แสงจำ)

อาจารย์ที่ปรึกษาโครงงาน ..... วันที่ 14 เดือน ตุลาคม พ.ศ.2564

(นางขวัญสกุล อุปพันธ์)

หัวข้อโครงการ	การทำนายและเปรียบเทียบความสนใจในการจดบันทึกของผู้เรียนจากมัลติมีเดียโดยใช้แบบจำลองการเรียนรู้ของเครื่องเพื่อเพิ่มประสิทธิภาพในการเรียนรู้	
ผู้ทำโครงการ	นายโชติพิสิฐ อุดลสีหัตต์	
อาจารย์ที่ปรึกษา	นายทศพร แสงจำ, นางขวัญสกุล อุปพันธ์	
สาขาวิชา	คณิตศาสตร์และวิทยาการคำนวณ	
โรงเรียน	มหิดลวิทยานุสรณ์	ปีการศึกษา 2564

### บทคัดย่อ

ความก้าวหน้าทางด้านเทคโนโลยีถูกนำมาประยุกต์ใช้เพื่อช่วยเหลือในด้านการศึกษา โดยมีการรวมข้อความรูปภาพ ภาพเคลื่อนไหว วิดีโอและเสียงเข้าไว้ด้วยกันเป็นมัลติมีเดียเพื่อใช้ในการสอน หากแต่ผู้เรียนบางส่วนไม่สามารถเรียนจากมัลติมีเดียได้อย่างมีประสิทธิภาพเต็มที่ ได้ การกระตุ้นความอยากรเรียนรู้ของผู้เรียนสามารถทำได้โดยการกระตุ้นให้ผู้เรียนจดจ่อกับสื่อการสอนมากขึ้น ซึ่งการจดบันทึกสามารถบ่งบอกความจดจ่อกับสื่อการสอนได้ ผู้วิจัยจึงต้องการที่จะตรวจวัดและเพิ่มประสิทธิภาพของมัลติมีเดียที่ใช้สอน โดยการสร้างโมเดลปัญญาประดิษฐ์ที่สามารถทำนายความสนใจในการจดของผู้เรียนจากองค์ประกอบในมัลติมีเดียได้ ผู้จัดทำใช้ข้อมูลจากการทดสอบนักเรียนโรงเรียนมหิดลวิทยานุสรณ์จำนวน 61 คน โดยให้ผู้เรียนดูวิดีโอสองเรื่อง ได้แก่ วิดีโออ่าวเปอร์เซีย (Persian) วิดีโอโรคทางจิตเวช (Schizophrenia) แล้วจดบันทึกตาม หลังจากนั้นแบ่งวิดีโอเป็นหัวข้อย่อยจำนวน 45 หัวข้อ (วิดีโอที่ 1 แบ่งเป็น 21 หัวข้อ และวิดีโอที่ 2 แบ่งเป็น 24 หัวข้อ) แล้วนับจำนวนคนที่จดบันทึกแต่ละหัวข้อ เพื่อหาความน่าจะเป็นที่ผู้เรียนจะจดในหัวข้อนั้น หลังจากนั้นเก็บข้อมูลภาพที่ใช้อธิบายหัวข้อนั้น แล้วทำ Image Segmentation ด้วย Mask R-CNN ที่ถูกฝึกโดยชุดข้อมูล COCO ได้เป็นข้อมูลของวัตถุในภาพ ผู้จัดทำสร้าง Features จากข้อมูลดังกล่าวจำนวน 10 Features แล้วแบ่งชุดข้อมูลเป็น 3 ชุด คือ ชุดข้อมูลวิเคราะห์วิดีโออ่าวเปอร์เซีย (Persian) ชุดข้อมูลวิเคราะห์โรคทางจิตเวช (Schizophrenia) และรูปแบบผสมผสาน (Mixed) โดยใช้โมเดล Random Forest เพื่อทำนายค่าความน่าจะเป็นโดยใช้ Features เหล่านั้นเป็นข้อมูลนำเข้า หลังจากที่ได้ผลจากโมเดลแล้ว วิเคราะห์ความแม่นยำของโมเดลด้วย Mean Absolute Percentage Error ได้ว่า ความแม่นยำของผลจากทั้งสามชุดข้อมูลเป็น 0.5113 0.9084 และ 0.3998 ตามลำดับ และมี Feature y1 y2 และ avg\_height มี Feature Importance ที่มากที่สุด สรุปผลได้ว่าความแม่นยำของโมเดลต่ำเกินกว่าจะนำไปใช้จริงได้ เพราะ ความคลาดเคลื่อนจาก Mask R-CNN ปัจจัยหลักของผลลัพธ์ที่ไม่ใช่ภาพ ข้อมูลที่ใช้มีไม่หลากหลายทำให้เกิดการ overfit ของโมเดล ผู้วิจัยจะศึกษาองค์ประกอบอื่นของมัลติมีเดียและปัญหาอื่นที่พบเจอ เพื่อให้สามารถสร้างปัญญาประดิษฐ์นี้ได้จริง

<b>Research Title</b>	Predicting and comparing learners' interest in note-taking from multimedia using a machine learning model to enhance learning efficiency	
<b>Researchers</b>	Mr. Chotpisit Adunsehawatt	
<b>Advisor</b>	Mr. Tossaporn Saengja, Mrs. Kwansakul Ouppaphan	
<b>Department</b>	Mathematics and Computing Science	
<b>School</b>	Mahidol Wittayanusorn	Year 2021

---

### ABSTRACT

Technological advancements have been applied to improve education. Text, images, animations, video, and audio are combined into educational multimedia. The problem is that some learners are unable to learn from multimedia effectively. Stimulating learners' curiosity can be achieved by encouraging them to concentrate on the material, which can be indicated by their note-taking. Therefore, the author wants to measure and increase the efficiency of multimedia teaching. By making an artificial intelligence model that can predict learners' interest in note-taking from elements in multimedia. The authors used data from a test of 61 Mahidol Wittayanusorn School students. The students were asked to watch Persian and Schizophrenia videos and take notes. After that, the video was divided into 45 sub-topics (21 for Persian and 24 for Schizophrenia) and counted the number of students' notes for each topic to find the probability that the learner will take note of the topic. Image data was collected and used to describe the topic and then does Image Segmentation with Mask R-CNN trained by COCO data set as the data of the objects in the image and 10 features were created. The data sets were then divided into three, Persian, Schizophrenia, and Mixed, using the Random Forest model to predict probabilities using those features as input. The model's accuracy is analyzed by the Mean Absolute Percentage Error of the results. It was found that the accuracy of the results from the three datasets were 0.5113, 0.9084, and 0.3998, respectively, and Feature y1, y2, and avg\_height had the greatest Feature Importance. In conclusion, the model's accuracy was too low to be practical because there are inaccuracies from Mask R-CNN, image is not the main determinant of results, the lack of data used causes an overfit of the model. To create the artificial intelligence, the study about other multimedia's elements and the other encountered problems could be further study.

### กิตติกรรมประกาศ

โครงการวิทยาศาสตร์นี้สำเร็จได้ด้วยความช่วยเหลือของ อ.ทศพร แสงจำ และ อ.ขวัญสกุล อุปพันธ์ ผู้เป็นอาจารย์ที่ปรึกษาโครงการ ผู้ซึ่งคอยให้คำปรึกษา คำแนะนำ เสนอแนวคิด ช่วยแก้ไขปัญหาที่ผู้จัดประสบ ตลอดช่วงเวลาที่ทำโครงการ และสอนกระบวนการจัดทำโครงการและงานวิจัยที่เป็นประโยชน์ต่อผู้จัดทำในอนาคต ทางผู้จัดทำจึงขอขอบพระคุณเป็นอย่างสูงมา ณ ที่นี้

ขอขอบพระคุณ อาจารย์อัญญรัตน์ บุญวัฒน์ อาจารย์ณอมศักดิ์ เหล่ากุล และอาจารย์เดี่ยว ใจบุญ คณะกรรมการตัดสินโครงการนี้ในรายวิชาโครงการวิทยาศาสตร์เพิ่มเติม 3 โรงเรียนมหิตลวิทยานุสรณ์ที่ได้ให้การประเมินและให้ข้อเสนอแนะในการปรับปรุงแก้ไขให้โครงการนี้ดีขึ้นตลอดระยะเวลาจัดทำโครงการ

ขอขอบพระคุณข้อมูลงานวิจัยเป็นแหล่งความรู้ สถาบันต่างวิจัยหลายแห่งที่ได้รวบรวมแหล่งข้อมูลและแหล่งอ้างอิงในการจัดทำโครงการนี้จนสำเร็จ

สุดท้ายขอขอบคุณกำลังใจและคำแนะนำจากเพื่อน รุ่นพี่ รุ่นน้องโรงเรียนมหิตลวิทยานุสรณ์คอยสนับสนุนเสมอตลอดมา จึงทำให้โครงการนี้สำเร็จตามจุดประสงค์ของรายวิชาโครงการวิทยาศาสตร์

ผู้จัดทำ

13 ตุลาคม 2564

## สารบัญ

	หน้า
บทคัดย่อ	ก
Abstract	ข
กิตติกรรมประกาศ	ค
สารบัญ	ง
สารบัญตาราง	จ
สารบัญรูป	ฉ
บทที่	หน้า
<b>1    บทนำ</b>	
1.1 ที่มาและความสำคัญของโครงการ	1
1.2 วัตถุประสงค์	1
1.3 ขอบเขตการศึกษา	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ	2
1.5 ระยะเวลาในการดำเนินการ	2
1.6 สถานที่ทำโครงการ	3
1.7 นิยามเชิงปฏิบัติการ	3
<b>2    งานวิจัย/เอกสารที่เกี่ยวข้อง</b>	
2.1 ทฤษฎีการเรียนรู้ของมนุษย์	4
2.2 การใช้ CNN เพื่อวิเคราะห์ภาพ	4
<b>3    วิธีการดำเนินการทดลอง</b>	
3.1 วัสดุอุปกรณ์	5
3.2 วิธีการทดลอง	5
<b>4    ผลการทดลองและอภิปรายผลการทดลอง</b>	
4.1 การคำนวณความแม่นยำ	8
4.2 ผลลัพธ์	8
4.3 อภิปรายผล	9
<b>5    สรุปผลการทดลอง</b>	11
บรรณานุกรม	13
ภาคผนวก (ถ้ามี)	15
ประวัติผู้ทำโครงการ	17

## สารบัญตาราง

	หน้า
ตารางที่ 1.5 ตารางแสดงแผนการทำงานของผู้จัดทำโครงการเรียงตามลำดับเวลา	2
ตารางที่ 4.2.2 ค่าความแม่นยำจากการทำนายของโมเดลโดยใช้ชุดข้อมูลที่ต่างกันที่ได้จากการคำนวณโดยใช้ค่า Mean Absolute Percentage Error	9



## สารบัญรูปภาพ

	หน้า
รูปที่ 4.2.1 กราฟผลลัพธ์ที่ได้จากชุดข้อมูลภาพที่ได้จากวิดีโอความรู้เรื่องสงครามอ่าวเปอร์เซีย (a), กราฟผลลัพธ์ที่ได้จากชุดข้อมูลภาพที่ได้จากวิดีโอความรู้เรื่องโรคทางจิตเวช (b), กราฟผลลัพธ์ที่ได้จากชุดข้อมูลภาพที่ได้จากทั้งสองวิดีโอ (c)	8
รูปที่ 4.2.3 กราฟ Feature Importance ที่ได้จากชุดข้อมูลภาพที่ได้จากวิดีโอความรู้เรื่องสงครามอ่าวเปอร์เซีย (a), กราฟ Feature Importance ที่ได้จากชุดข้อมูลภาพที่ได้จากวิดีโอความรู้เรื่องโรคทางจิตเวช (b), กราฟ Feature Importance ที่ได้จากชุดข้อมูลภาพที่ได้จากทั้งสองวิดีโอ (c)	9
รูปที่ 4.3.1 รูปภาพแสดงแผนที่โลกและธงชาติที่ไม่สามารถถูกจำแนกโดย Mask R-CNN ได้ (a), รูปภาพแสดงรูปโลกที่ถูกจำแนกโดย Mask R-CNN เป็นรูปว่าว ซึ่งไม่ตรงความเป็นจริง (b)	10
รูปที่ ก.1 ภาพจากวิดีโอความรู้สงครามอ่าวเปอร์เซียที่ยังไม่ถูกทำ Image Segmentation จำนวน 21 ภาพ	15
รูปที่ ก.2 ตัวอย่างภาพที่ถูกทำ Image Segmentation จากชุดข้อมูลภาพจากวิดีโอความรู้สงครามอ่าวเปอร์เซีย	15
รูปที่ ข.1 ภาพจากวิดีโอความรู้โรคทางจิตเวชที่ยังไม่ถูกทำ Image Segmentation จำนวน 24 ภาพ	16
รูปที่ ข.2 ตัวอย่างภาพที่ถูกทำ Image Segmentation จากชุดข้อมูลภาพจากวิดีโอความรู้โรคทางจิตเวช	16

## บทที่ 1

### บทนำ

#### 1.1 ที่มาและความสำคัญ

การศึกษาเป็นพื้นฐานที่สำคัญต่อการพัฒนาความสามารถทุกด้านของมนุษย์ ในปัจจุบันเทคโนโลยีก้าวหน้ามากขึ้น ทำให้มีบทบาทในการด้านการศึกษามากขึ้น ส่งผลให้เกิดสื่อการสอนที่หลายรูปแบบที่นำเสนอได้ผ่านเครือข่ายอินเทอร์เน็ต อินทราเน็ต หรือซีดี การเปลี่ยนแปลงและใช้สื่อการสอนรูปแบบดังกล่าวทำให้ผู้เรียนเกิดทักษะการเรียนรู้ที่จำเป็นในปัจจุบัน โดยเป็นสื่อการสอนประเภทมัลติมีเดียประกอบด้วยข้อความ รูปภาพ ภาพเคลื่อนไหว เสียง ทำให้ผู้เรียนสามารถจินตนาการภาพและเข้าใจเนื้อหาการเรียนได้อย่างมีประสิทธิภาพมากขึ้น

การผลิตสื่อการสอนที่ดี ผู้สอนหรือผู้ผลิตสื่อการสอนต้องคำนึงถึงผู้เรียนเป็นหลัก โดยจะต้องคิด จัดการ และวางแผนการใช้เทคนิคในการสร้างและจัดองค์ประกอบทุกส่วนในสื่อการสอน ทำให้ผู้เรียนตื่นตัวขณะเรียน สนใจในการเรียนรู้ ความเข้าใจในเนื้อหาที่เรียน และมีความสุขในการเรียน เพื่อเพิ่มประสิทธิภาพการเรียนรู้จากสื่อการสอนมากที่สุด

เนื่องด้วยสถานการณ์การแพร่ระบาดของเชื้อโคโรนาสายพันธุ์ใหม่ 2019 ทำให้การเรียนการสอนในปัจจุบันส่วนใหญ่ไม่สามารถจัดขึ้นที่โรงเรียนได้ จึงทำให้ผู้เรียนต้องเรียนรู้ผ่านระบบออนไลน์ การเรียนรู้ผ่านระบบอินเทอร์เน็ตจากสถานที่ส่วนตัวของผู้เรียน ซึ่งส่วนใหญ่มีบรรยากาศและสภาพแวดล้อมที่ไม่เหมาะสมต่อการเรียน ทำให้ผู้เรียนขาดสมาธิ ขาดความสนใจ ผู้เรียนส่วนใหญ่ไม่คุ้นชินกับการเรียนออนไลน์ ทำให้มีปัญหาในการควบคุมตนเองขณะเรียน ทำให้การเรียนมีประสิทธิภาพลดลงหากเทียบกับการเรียนที่โรงเรียนตามปกติ ซึ่งอาจส่งผลต่อการศึกษาของผู้เรียนในระยะยาวได้

สิ่งที่ปัญหาสำหรับการเรียนอย่างหนึ่ง คือ สื่อการสอนและวิธีการสอนที่ไม่ทำให้ผู้เรียนสนใจหรือกระตุ้นการเรียนรู้มากพอ ผู้วิจัยเห็นถึงปัญหาการขาดประสิทธิภาพในการเรียนรู้ จึงสนใจที่วิเคราะห์และแก้ไขปัญหาประสิทธิภาพจากการเรียนรู้ผ่านมัลติมีเดีย โดยเน้นที่ความสนใจด้านการจดบันทึกจากการเรียนรู้เป็นหลัก ผ่านการใช้องค์ความรู้ทางด้านปัญญาประดิษฐ์ ประเภทการเรียนรู้ของเครื่องในการประมวลผลข้อมูลที่เป็นองค์ประกอบของมัลติมีเดีย เช่น รูปภาพ ข้อความ และเสียง โดยใช้ชุดข้อมูลจากการทดลองจดบันทึกของนักเรียนโรงเรียนมหิดลวิทยานุสรณ์จากการดูมัลติมีเดีย เพื่อวิเคราะห์ความสอดคล้องและความเชื่อมโยงขององค์ประกอบในมัลติมีเดียที่ใช้ เพื่อนำไปประยุกต์ใช้ในการผลิตสื่อการสอนที่มีประสิทธิภาพต่อไป

#### 1.2 วัตถุประสงค์

1. เพื่อศึกษาและสร้างปัญญาประดิษฐ์ที่สามารถทำนายการจดบันทึกของผู้เรียนจากรูปภาพในมัลติมีเดีย
2. เพื่อหาความสัมพันธ์ขององค์ประกอบรูปภาพของมัลติมีเดียที่ส่งผลต่อการจดบันทึกของผู้เรียน

### 1.3 ขอบเขตการศึกษา

1. ข้อมูลที่ใช้เป็นข้อมูลจากการทดลองจัดบันทึกมัลติมีเดียจากนักเรียนโรงเรียนมหิดลวิทยานุสรณ์ชั้นมัธยมศึกษาปีที่ 4 และ 5 ปีการศึกษา 2563 จำนวน 61 คน
2. เนื้อหาของมัลติมีเดียเป็นเนื้อหาที่ผู้เรียนส่วนใหญ่ไม่มีสนใจหรือไม่มีพื้นฐานมาก่อน
3. โครงการนี้จะศึกษาเฉพาะความสนใจในการจัดบันทึกของผู้เรียนจากองค์ประกอบภาพของมัลติมีเดียเท่านั้น

### 1.4 ประโยชน์ที่คาดว่าจะได้รับ

ปัญญาประดิษฐ์ที่สร้างขึ้นสามารถหาความสัมพันธ์ขององค์ประกอบในมัลติมีเดียกับผลลัพธ์จากการจัดบันทึกของผู้เรียน และสามารถประเมินประสิทธิภาพของมัลติมีเดียในด้านความสนใจในการจัดบันทึกได้อย่างแม่นยำ

เพื่อให้ผู้ต้องการผลิตมัลติมีเดียสามารถนำองค์ความรู้หรือปัญญาประดิษฐ์นี้ไปใช้ในการประกอบการตัดสินใจเพื่อเลือกใช้หรือจัดองค์ประกอบของสื่อการสอนได้อย่างเหมาะสมต่อการเรียนรู้จากการจัดบันทึกผู้สอนสามารถหาวิธีหรือเทคนิคในการสร้างสื่อการสอนให้มีประสิทธิภาพ เพื่อเพิ่มประสิทธิภาพของการเรียนของผู้เรียน และส่งเสริมให้การศึกษาของประเทศไทยดีขึ้น

### 1.5 ระยะเวลาการดำเนินการ

ใช้เวลา 5 เดือน ตั้งแต่เดือนพฤษภาคม พ.ศ. 2564 ถึงเดือนกันยายน พ.ศ. 2564

กิจกรรมที่ทำ	ระยะเวลาปฏิบัติงาน				
	พ.ศ. 2564				
	พ.ค.	มิ.ย.	ก.ค.	ส.ค.	ก.ย.
ระบุปัญหาและหาหัวข้อโครงการ					
สืบค้นข้อมูลและวางแผนโครงการ					
เขียนโครงร่างโครงการ					
ทำการทดลอง					
วิเคราะห์ข้อมูลและสรุปผล					
จัดทำรูปแบบรายงาน					
นำเสนอโครงการ					

ตารางที่ 1.5 ตารางแสดงแผนการทำงานของผู้จัดทำโครงการเรียงตามลำดับเวลา

## 1.6 สถานที่ทำโครงการ

ทำโครงการผ่านระบบออนไลน์ ที่ที่พักอาศัยของผู้จัดทำ

## 1.7 นิยามเชิงปฏิบัติการ

ประสิทธิภาพการเรียนรู้	ความสัมพันธ์ระหว่างการวัดผลสัมฤทธิ์ทางการเรียน เช่น คะแนนการทดสอบหรือระยะเวลาที่ผู้เรียนใช้ในการทำงานอย่างถูกต้อง กับการรับรู้ความพยายามทางจิต (PME) ของผู้เรียน (Hearrington, 2010)
มาสก์	ภาพเลขฐานสองที่ประกอบด้วยค่าศูนย์และหนึ่ง บอกค่าความจริงของแต่ละพิกเซลในภาพ เพื่อบอกตำแหน่งของวัตถุในภาพ
มัลติมีเดีย	การนำองค์ประกอบของสื่อชนิดต่าง ๆ มาผสมผสานรวมกัน ซึ่งประกอบด้วยตัวอักษร ภาพนิ่ง ภาพเคลื่อนไหว เสียง และวิดีโอ โดยผ่านกระบวนการทางระบบคอมพิวเตอร์ เพื่อสื่อความหมายกับผู้ใช้อย่างมีประสิทธิภาพ และได้บรรลุตามวัตถุประสงค์การใช้งาน
Image Segmentation	กระบวนการจำแนกพิกเซลในภาพว่าเป็นพิกเซลของวัตถุใดในภาพ จะได้ผลออกมาเป็นแบ่งเป็นพื้นที่สีหลายสี ซึ่งแต่ละสีหมายความว่าลักษณะหรือวัตถุที่แตกต่างกัน
Features	ตัวแปรอิสระที่ทำหน้าที่เหมือนข้อมูลนำเข้าในโมเดลปัญญาประดิษฐ์ โดยขณะทำการทำนาย โมเดลจะใช้มันเพื่อทำนายผลลัพธ์โดยใช้กระบวนการและอัลกอริทึมเฉพาะของโมเดลที่ใช้
ผู้เรียน	ผู้ศึกษาค้นคว้าปฏิบัติด้วยตนเองในเรื่องตามที่ครูกำหนด รวม นักเรียน นิสิต นักศึกษา และผู้รับการศึกษาอบรมนอกระบบสถานศึกษาตามปกติ โดยอยู่ในบริบทของการเรียนรู้

## บทที่ 2

### งานวิจัย/เอกสารที่เกี่ยวข้อง

#### 2.3 ทฤษฎีการเรียนรู้ของมนุษย์

Mayer (2005) เสนอ Cognitive Load Theory of Multimedia Learning กล่าวไว้ว่า ช่องทางสำหรับการประมวลผลข้อมูลของสมองมีอยู่ 2 ช่องทาง ที่แยกกัน คือการจำข้อมูลเสียงผ่านการได้ยิน (Phonological Loop) และการจำข้อมูลภาพหรือการเคลื่อนไหวผ่านการมองเห็น (Visual-Spatial Sketch pad) โดยความจุของช่องทางประมวลผลแต่ละช่องมีขีดจำกัด ซึ่งข้อสมมติฐานเหล่านี้บ่งชี้ว่ามนุษย์สามารถประมวลผลข้อมูลในแต่ละช่องทางได้ในปริมาณที่จำกัดและไม่ถาวร ซึ่งจะถูเก็บไว้ในความทรงจำระยะสั้น ก่อนที่จะถูกนำเข้าความทรงจำระยะยาวต่อไป

จากการศึกษาของเมเยอร์ (Mayer, 2009) พบว่า การทำวิดีโอที่ดีต้องมีสี่ส่วนสำคัญ ได้แก่ ประการแรก ความสัมพันธ์ระหว่างเนื้อหา โดยต้องไม่มีรูปหรือข้อความที่ไม่จำเป็นใส่ลงไปในวิดีโอ ประการที่สอง คำพูดและภาพผสมผสานกันจะทำให้ผู้เรียนเรียนรู้ได้ดีขึ้น ประการที่สาม ภาพผู้บรรยายไม่จำเป็นต้องมาใส่ในวิดีโอ หากเป็นภาพที่มีความซับซ้อนในการทำความเข้าใจ รูปภาพของผู้บรรยายก็ไม่ควรนำมาใส่ สอดคล้องกับงานวิจัย Van Gog & Rummel (2010) ประการที่สี่ ผู้เรียนเรียนรู้ได้ดีขึ้นถ้ามีการพูดหรือมีบทสนทนาในวิดีโอ และการทำภาพเคลื่อนไหวอาจทำให้เสมือนว่ามีมนุษย์อยู่ในภาพนั้น

#### 2.4 การใช้ CNN เพื่อวิเคราะห์ภาพ

Jeon และคณะ (2020) นำเสนอข้อมูลการประยุกต์ใช้ปัญญาประดิษฐ์ประเภทการเรียนรู้เชิงลึกและการเรียนรู้ของเครื่องเพื่อวิเคราะห์ข้อมูลที่มีความเกี่ยวข้องกับมลพิษ โดยนำเสนอประเภทแบบจำลองประสิทธิภาพ ความแม่นยำ การปรับปรุงแบบจำลองและการประยุกต์ใช้ในชีวิตประจำวัน ยกตัวอย่างเช่น การใช้แบบจำลองโครงข่ายประสาทแบบคอนโวลูชันในการเปรียบเทียบและหาความเชื่อมโยงสอดคล้องจากภาพจากกล้องบันทึกภาพจากตำแหน่งที่ต่างกันโดยใช้ Convolutional layers สามมิติ และ Spatial Transformer Network (STN) ได้พบว่าแบบจำลองมีความแม่นยำสูงกว่าแบบจำลองอื่นในชุดข้อมูลเดียวกัน 4%

Kaiming และคณะ (2018) นำเสนอกรอบแนวคิดที่เรียบง่าย ยืดหยุ่น และทั่วไปสำหรับการแบ่งส่วนวัตถุในภาพได้อย่างมีประสิทธิภาพ ในขณะเดียวกันก็สร้างมาส์กการแบ่งส่วนคุณภาพสูง โมเดลนี้เรียกว่า Mask R-CNN ซึ่งถูกพัฒนาต่อจาก Faster R-CNN โดยการเพิ่มการทำนายมาส์กควบคู่ไปกับการจดจำขอบเขต Mask R-CNN นั้นฝึกได้ง่ายซึ่งทำงานที่ 5 fps นอกจากนี้ Mask R-CNN ยังง่ายต่อการสรุปรูปงานอื่นๆ เช่น ช่วยให้เราประเมินท่าทางของมนุษย์ในกรอบการทำงานเดียวกัน มันสามารถแสดงผลลัพธ์ที่ดีที่สุดในชุดทดสอบทั้งสามของ COCO รวมถึงการแบ่งและตรวจจับขอบกล่องของวัตถุ และการตรวจจับจุดสำคัญของบุคคล Mask R-CNN มีประสิทธิภาพเหนือกว่าผลงานรุ่นเดียวที่มีอยู่ทั้งหมดในทุกงาน รวมถึงชนะการแข่งขัน COCO 2016

## บทที่ 3

### วิธีการดำเนินการทดลอง

#### 3.1 วัสดุอุปกรณ์

- 3.1.1 คอมพิวเตอร์พกพา HP Pavilion x360 Convertible 14-dh0xxx
- 3.1.2 ซอฟต์แวร์ Jupyter Notebook
- 3.1.3 ซอฟต์แวร์ Microsoft Excel

#### 3.2 วิธีการทดลอง

##### 3.2.1 การเตรียมข้อมูล

รวบรวมข้อมูลจากข้อมูลที่ได้จากการทดลองของโครงงานวิทยาศาสตร์ที่ผู้จัดทำทำในรายวิชาโครงงานวิทยาศาสตร์ 1 2 และ 3 โดยแบ่งเป็นสองส่วนคือ ข้อมูลการจัดบันทึกของผู้เรียนและข้อมูลภาพที่ได้จากมัลติมีเดียที่ผู้เรียนใช้ในการเรียนรู้

ข้อมูลภาพที่ได้จากมัลติมีเดีย ได้จากการเก็บภาพหนึ่งทีอธิบายเฉพาะหนึ่งหัวข้อภายในช่วงเวลาทีวิดีโออธิบายหัวข้อนั้น จำนวนหนึ่งภาพต่อหัวข้อ ขนาดภาพ 1080 พิกเซล คูณ 1960 พิกเซล

ข้อมูลการจัดบันทึก ประกอบไปด้วยข้อมูลหัวข้อที่แบ่งตามเนื้อหาที่แตกต่างกันในมัลติมีเดียที่ใช้ คือ วิดีโอความรู้เรื่องสงครามอ่าวเปอร์เซียและวิดีโอความรู้เรื่องโรคทางจิตเวช โดยแบ่งเป็น 21 หัวข้อและ 24 หัวข้อจากทั้งสองวิดีโอตามลำดับ และข้อมูลการจัดบันทึกของผู้เรียนซึ่งเป็นนักเรียนโรงเรียนมหิดลวิทยานุสรณ์ ชั้นมัธยมศึกษาปีที่ 4 และ 5 รวมทั้งหมด 61 คน โดยบอกสถานะการจัดบันทึกหัวข้อเฉพาะหนึ่งของผู้เรียนคนหนึ่ง แบ่งเป็นสองสถานะคือจดและไม่จด

การเก็บข้อมูลการจัดบันทึกจะเก็บในรูปแบบไฟล์ใน Microsoft Excel แล้วหาค่าเฉลี่ยของสถานะการจัดของผู้เรียนในทุกหัวข้อของทั้งสองวิดีโอ โดยกำหนดให้ 1 แทนจด และ 0 แทนไม่จด และเก็บค่าเฉลี่ยที่ได้เป็นตัวแปร probability เพื่อนำไปใช้ในโมเดลปัญญาประดิษฐ์ต่อไป

##### 3.2.2 Image Segmentation

การใช้โมเดล Mask R-CNN ที่ถูกฝึกโดยชุดข้อมูลจาก COCO ในการทำ Image Segmentation กับข้อมูลภาพที่ได้จากมัลติมีเดียภายในซอฟต์แวร์ Jupyter Notebook เพื่อหามาสก์ของวัตถุที่อยู่ในภาพนั้น และจำแนกประเภทวัตถุจากทั้งหมด 81 ประเภทในชุดข้อมูล โดยได้ผลลัพธ์เป็นชุดข้อมูลของวัตถุทั้งหมดในภาพที่ได้จากการทำ Image Segmentation

ชุดข้อมูลที่ได้จากโมเดล Mask R-CNN ประกอบด้วยตำแหน่งกล่องที่ครอบวัตถุในภาพแทนด้วยตัวแปร  $x1\ y1\ x2\ y2$  เพื่อบอกพิกัดของขอบ ประเภของวัตถุแทนด้วยตัวแปร ClassID มาสก์ของวัตถุแทนด้วยอาเรย์สองมิติที่มีค่าจริงหรือเท็จตามตำแหน่งพิกเซลในภาพ และค่าความน่าจะเป็นที่ได้จากการทำนายที่จะเป็นวัตถุนั้นแทนด้วยตัวแปร score

ผู้จัดทำกำหนดตัวแปร cntmask จากข้อมูลมาสก์ของวัตถุทั้งหมดในภาพ คือ จำนวนค่าความจริงที่เป็นจริงทั้งหมดของวัตถุทุกวัตถุในภาพ เพื่อใช้ในการคำนวณ Features ในขั้นตอนถัดไป

### 3.2.3 Features

การกำหนด Features เพื่อใช้ในการทำนายโดยโดยใช้โมเดลปัญญาประดิษฐ์จากข้อมูลที่ได้จากการกระบวนการ Image Segmentation โดยสามารถ Features จากภาพที่ใช้ เพื่อเป็นข้อมูลนำเข้าสำหรับโมเดลปัญญาประดิษฐ์ได้ดังนี้

$n$  คือ จำนวนวัตถุที่อยู่ในภาพ

box\_area คือ อัตราส่วนของผลรวมพื้นที่ของกรอบที่ครอบวัตถุทั้งหมดในภาพ มีค่าเท่ากับ

$$\sum_{i=1}^n (x1_i - x2_i) \times (y1_i - y2_i) / (1080 \times 1960)$$

mask\_area คือ อัตราส่วนของผลรวมมาสก์ของวัตถุทั้งหมดในภาพโดยคำนวณจากตัวแปร cntmask มีค่าเท่ากับ  $cntmask / (1080 \times 1960)$

avg\_score คือ ค่าเฉลี่ยของ score ของวัตถุทั้งหมดในภาพ มีค่าเท่ากับ  $\sum_{i=1}^n score_i / n$

avg\_x1 คือ ค่าเฉลี่ยของ x1 ของวัตถุทั้งหมดในภาพ มีค่าเท่ากับ  $\sum_{i=1}^n x1_i / n$

avg\_y1 คือ ค่าเฉลี่ยของ y1 ของวัตถุทั้งหมดในภาพ มีค่าเท่ากับ  $\sum_{i=1}^n y1_i / n$

avg\_x2 คือ ค่าเฉลี่ยของ x2 ของวัตถุทั้งหมดในภาพ มีค่าเท่ากับ  $\sum_{i=1}^n x2_i / n$

avg\_y2 คือ ค่าเฉลี่ยของ y2 ของวัตถุทั้งหมดในภาพ มีค่าเท่ากับ  $\sum_{i=1}^n y2_i / n$

avg\_width คือค่าเฉลี่ยของความกว้างของวัตถุหรือ  $x2-x1$  ของวัตถุทั้งหมดในภาพ มีค่าเท่ากับ  $\sum_{i=1}^n (x2_i - x1_i) / n$

avg\_height คือ ค่าเฉลี่ยของความยาวของวัตถุหรือ  $y_2 - y_1$  ของวัตถุทั้งหมดในภาพ มีค่าเท่ากับ  $\sum_{i=1}^n (y_{2i} - y_{1i}) / n$

ข้อมูลการจดบันทึกสามารถแปลงเป็นเป็น Feature ที่ต้องการให้โมเดลทำนายจากข้อมูลนำเข้าหรือ Features ที่ได้จากข้อมูลรูปภาพได้ดังนี้

prob คือ ความน่าจะเป็นที่ผู้เรียนจะจดในหัวข้อนั้นโดยอ้างอิงจากตัวแปร probability ที่ได้จาก Microsoft Excel

### 3.2.4 Random Forest

ผู้จัดทำใช้โมเดลปัญญาประดิษฐ์ประเภท Random Forest จาก sklearn เพื่อใช้ทำนายผลลัพธ์ โดยใช้ Features จากข้อมูลภาพ ซึ่งสามารถแบ่งชุดข้อมูลได้สามชุดข้อมูลคือ ชุดข้อมูลที่ได้จากวิดีโอความรู้เรื่อง สงครามอ่าวเปอร์เซีย (Persian) จำนวน 21 ข้อมูล ชุดข้อมูลที่ได้จากวิดีโอความรู้เรื่อง โรคทางจิตเวช (Schizophrenia) จำนวน 24 ข้อมูล และชุดข้อมูลที่รวมข้อมูลจากทั้งสองวิดีโอ (Mixed) จำนวน 45 ข้อมูล

กระบวนการเริ่มต้นด้วยการสุ่มข้อมูลจากชุดข้อมูลที่ใช้ โดยแบ่งเป็น 75 เปอร์เซ็นต์สำหรับการฝึก โมเดล และ 25 เปอร์เซ็นต์สำหรับการทดสอบความแม่นยำของโมเดล หลังจากการแบ่งชุดข้อมูล ผู้จัดทำนำชุด ข้อมูลเหล่านั้นฝึกในโมเดล Random Forest ที่มีต้นไม้จำนวน 1000 ต้น โดยการสุ่มเงื่อนไขจาก Features ทั้งหมด

หลังจากที่ได้โมเดลที่ถูกฝึกแล้ว ผู้จัดทำนำชุดข้อมูลที่ใช้ทดสอบเข้าโมเดลที่ถูกฝึก เพื่อให้โมเดล ทำนายผลลัพธ์จากข้อมูลนำเข้าเหล่านั้น โดยผลลัพธ์เกิดจากการเฉลี่ยผลลัพธ์ที่ได้จากต้นไม้ทั้งหมด ซึ่งคือ ความน่าจะเป็นที่ผู้เรียนจะจดจากการเห็นภาพที่มี Features เฉพาะค่าหนึ่งซึ่งตรงกับข้อมูลที่ใช้ทดสอบ

โมเดล Random Forest สามารถหา Features Importance เพื่อบอกว่า Feature ไດส่งผลต่อการ เปลี่ยนแปลงของผลลัพธ์หรือความน่าจะเป็นที่ผู้เรียนจะจดมากที่สุด โดยการสุ่มเปลี่ยนเงื่อนไขของต้นไม้ทุกต้น ทีละ Feature ทำให้โมเดลสามารถสังเกตและวิเคราะห์การเปลี่ยนแปลงของผลลัพธ์ได้



## บทที่ 4

### ผลการทดลองและอภิปรายผลการทดลอง

#### 4.1 การคำนวณความแม่นยำ

ผู้จัดทำใช้ค่า Mean Absolute Percentage Error หรือ MAPE เพื่อคำนวณหาความแม่นยำของผลลัพธ์ที่ได้จากโมเดล ซึ่งเป็นไปตามสมการ

$$M = \frac{1}{n} \sum_{i=1}^n \frac{|Actual_i - Predict_i|}{Actual_i}$$

ความแม่นยำ (Accuracy) คำนวณได้จาก

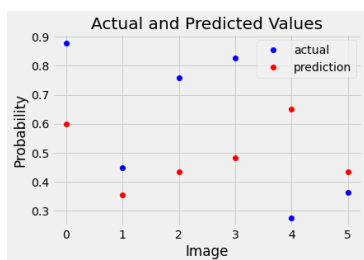
$$Accuracy = 1 - M$$

โดยค่าความแม่นยำที่ได้ดีมีความได้ว่า หากความแม่นยำมีค่าเท่ากับ 1 แปลว่าโมเดลมีความแม่นยำมากที่สุด และหากความแม่นยำมีค่าเท่ากับ 0 แปลว่าโมเดลไม่มีความแม่นยำ ซึ่งหากความแม่นยำมีค่ามากกว่า 0.8 แปลว่าโมเดลมีความแม่นยำในระดับสูง

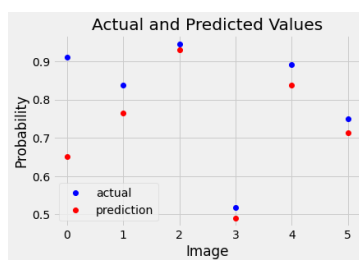
#### 4.2 ผลลัพธ์

##### 4.2.1 กราฟจากผลการทำนาย

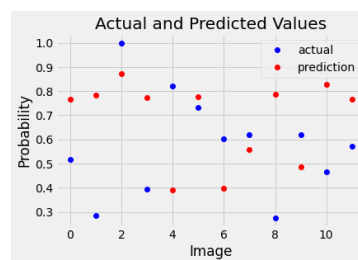
ผู้จัดทำนำผลลัพธ์ที่ได้จากการทำนายมาพล็อตเป็นกราฟเพื่อเปรียบเทียบผลที่ได้จากการทำนายและค่าจริงที่คาดหวัง โดยแกนนอนเป็นภาพที่ใช้ แกนตั้งเป็นค่าความน่าจะเป็น จุดสีแดงคือค่าที่ได้จากการทำนาย และจุดสีฟ้าคือค่าจริง



(a)



(b)



(c)

รูปที่ 4.2.1 กราฟผลลัพธ์ที่ได้จากชุดข้อมูลภาพที่ได้จากวิดีโอความรู้เรื่องสงครามอ่าวเปอร์เซีย (a), กราฟผลลัพธ์ที่ได้จากชุดข้อมูลภาพที่ได้จากวิดีโอความรู้เรื่องโรคทางจิตเวช (b), กราฟผลลัพธ์ที่ได้จากชุดข้อมูลภาพที่ได้จากทั้งสองวิดีโอ (c)

#### 4.2.2 ความแม่นยำของโมเดล

จากการใช้ MAPE ในการคำนวณความแม่นยำจากผลลัพธ์ของทั้งสามชุดข้อมูล ได้ว่า

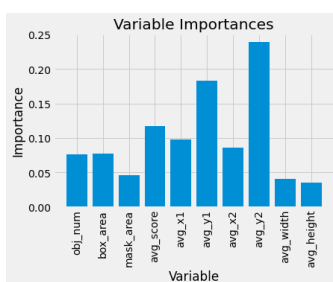
ชุดข้อมูล	Accuracy (1-MAPE)
Persian	0.5113
Schizophrenia	0.9084
Mixed	0.3998

ตารางที่ 4.2.2 ค่าความแม่นยำจากการทำนายของโมเดลโดยใช้ชุดข้อมูลที่ต่างกันที่ได้จากการคำนวณโดยใช้ค่า Mean Absolute Percentage Error

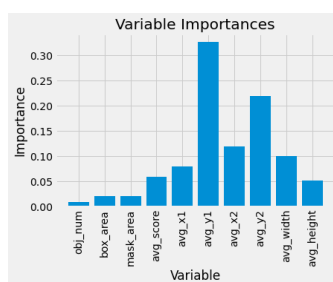
#### 4.2.3 Features Importance

จากการวิเคราะห์ของโมเดล Random Forest ทำให้ทราบ Features Importance ซึ่งบอกความสำคัญของ Feature หนึ่งต่อค่าผลลัพธ์ที่ได้จากการทำนายโดยโมเดล

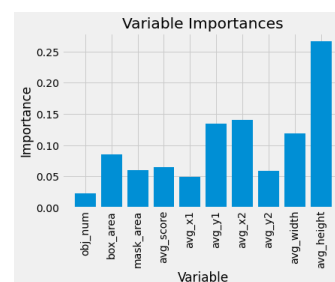
โมเดลแสดงกราฟเพื่อบ่งบอก Feature Importance โดยแกนนอนคือชื่อของ Feature จำนวน 10 ประเภทที่ใช้ในการทำนาย และแกนตั้งคือความสำคัญของ Feature (Importance) นั้น ซึ่งความสำคัญของ Feature แปรผันตามค่า Importance ที่ได้



(a)



(b)



(c)

รูปที่ 4.2.3 กราฟ Feature Importance ที่ได้จากชุดข้อมูลภาพที่ได้จากวิดีโอความรู้เรื่องสงครามอ่าวเปอร์เซีย (a), กราฟ Feature Importance ที่ได้จากชุดข้อมูลภาพที่ได้จากวิดีโอความรู้เรื่องโรคทางจิตเวช (b), กราฟ Feature Importance ที่ได้จากชุดข้อมูลภาพที่ได้จากทั้งสองวิดีโอ (c)

#### 4.3 อภิปรายผล

##### 4.3.1 ความแม่นยำของโมเดล

จากตารางที่ 4.2.2 ได้ว่า ความแม่นยำจากการใช้โมเดล Random Forest ทำนายโดยใช้ชุดข้อมูลที่แตกต่างกันมีค่าแตกต่างกัน คือ ความแม่นยำจากชุดข้อมูล Persian และ Mixed มีค่าน้อยกว่า 0.8 แปลว่า

โมเดลมีความแม่นยำในระดับต่ำ ซึ่งอาจเกิดจากความไม่สมบูรณ์ของข้อมูลและ Features ที่คำนวณได้จากข้อมูลภาพที่ได้จากการทำ Image Segmentation โดย Mask R-CNN เพราะชุดข้อมูล COCO ที่ใช้ฝึกโมเดล Mask R-CNN มีจำนวนไม่มากและหลากหลายพอที่จะวิเคราะห์รูปภาพที่ผู้จัดทำนำมาใช้ได้ ยกตัวอย่างเช่น การทำ Image Segmentation กับวัตถุที่ Mask R-CNN จะไม่สามารถจำแนกได้ เพราะไม่มีข้อมูลวัตถุนั้นในชุดข้อมูล COCO หรือ ความผิดพลาดจากการจำแนกที่ตรงความเป็นจริง เพราะวัตถุที่อยู่ในภาพมีความคล้ายกับวัตถุในชุดข้อมูล COCO แต่ไม่ใช่วัตถุเดียวกัน



(a)



(b)

**รูปที่ 4.3.1** รูปภาพแสดงแผนที่โลกและธงชาติที่ไม่สามารถถูกจำแนกโดย Mask R-CNN ได้ (a), รูปภาพแสดงรูปโลกที่ถูกจำแนกโดย Mask R-CNN เป็นรูปว่าว ซึ่งไม่ตรงความเป็นจริง (b)

ความแม่นยำข้างต้นตรงกันข้ามกับความแม่นยำจากชุดข้อมูล Schizophrenia ที่มีความแม่นยำมากกว่า 0.9 ซึ่งแปลว่าโมเดลมีความแม่นยำในระดับสูงมาก ซึ่งสามารถวิเคราะห์ได้ว่าชุดข้อมูลที่ใช้อาจเกิดจากการสุ่มแบ่งชุดข้อมูลที่ใช้ฝึกและทดสอบได้จำนวนไม่มากและหลากหลายพอทำให้เกิดการ overfit ของโมเดลทำให้ความแม่นยำมีค่ามากกว่าสองชุดข้อมูลข้างต้น

#### 4.3.2 Features Importance

จากรูปที่ 4.2.3 หากพิจารณา Feature ที่มีความสำคัญ (Importance) มากที่สุดสามอันดับแรก พบว่ามี Feature  $y_1$   $y_2$  และ  $avg\_height$  ที่ส่งผลมากที่สุด ซึ่ง Features เหล่านี้เป็นค่าที่ได้มาจากการคำนวณพิกัดในแกน  $y$  ของวัตถุที่ได้จากการทำ Image Segmentation ทำให้อาจวิเคราะห์ได้ว่าพิกัด ตำแหน่ง และความยาวในแกน  $y$  อาจส่งผลต่อความสนใจในการจดบันทึกของผู้เรียนมากกว่า Feature อื่นที่กำหนดไว้ โดยผู้จัดทำตั้งสมมติฐานว่าการมองโดยใช้ระดับสายตาในแนวตั้งส่งผลต่อความสนใจในการจดของผู้เรียน แต่ไม่สามารถหางานวิจัยเพื่อศึกษาหรือยืนยันสมมติฐานดังกล่าวได้

หากพิจารณา Feature ที่มีความสำคัญ (Importance) น้อยที่สุด ได้ว่าจำนวนวัตถุหรือ  $obj\_num$  ส่งผลต่อผลลัพธ์จากการทำนายน้อยที่สุด ทำให้อาจสรุปได้ว่าจำนวนของวัตถุที่ปรากฏบนรูปภาพมีผลต่อการจดบันทึกของผู้เรียนน้อยกว่า Feature อื่นที่ผู้จัดทำกำหนด

## บทที่ 5

### สรุปผลการทดลอง

#### 5.1 สรุปผล

ในการหาความสัมพันธ์ขององค์ประกอบรูปภาพของมัลติมีเดียที่ส่งผลต่อการจดบันทึกของผู้เรียนนั้น เพื่อตรวจวัดและเพิ่มประสิทธิภาพของมัลติมีเดียที่ใช้สอน โดยการสร้างโมเดลปัญญาประดิษฐ์ที่สามารถทำนายความสนใจในการจดของผู้เรียนจากองค์ประกอบในมัลติมีเดียได้ ผู้จัดทำใช้ข้อมูลจากการทดสอบนักเรียนโรงเรียนมหิดลวิทยานุสรณ์จำนวน 61 คน โดยผู้เรียนให้ดูวิดีโอสองวิดีโอคือเรื่อง Persian และ Schizophrenia แล้วจดตาม หลังจากนั้นแบ่งวิดีโอเป็นหัวข้อย่อยจำนวน 45 หัวข้อ (Persian 21 หัวข้อและ Schizophrenia 24 หัวข้อ) แล้วนับจำนวนคนจดในแต่ละหัวข้อ เพื่อหาความน่าจะเป็นที่ผู้เรียนจะจดในหัวข้อนั้น หลังจากนั้นเก็บข้อมูลภาพที่ใช้อธิบายหัวข้อนั้น แล้วทำ Image Segmentation ด้วย Mask R-CNN ที่ถูกฝึกโดยชุดข้อมูล COCO ได้เป็นข้อมูลของวัตถุในภาพ ผู้จัดทำสร้าง Features จากข้อมูลดังกล่าวจำนวน 10 Features แล้วแบ่งชุดข้อมูลเป็น 3 ชุดคือ Persian Schizophrenia และ Mixed โดยใช้โมเดล Random Forest เพื่อทำนายค่าความน่าจะเป็นโดยใช้ Features เหล่านั้นเป็นข้อมูลนำเข้า หลังจากที่ได้ผลจากโมเดลแล้ว วิเคราะห์ความแม่นยำของโมเดลด้วย Mean Absolute Percentage Error ได้ว่า ความแม่นยำของผลจากทั้งสามชุดข้อมูลเป็น 0.5113 0.9084 และ 0.3998 ตามลำดับ และ Feature y1 y2 และ avg\_height มี Feature Importance มากที่สุด

ทั้งนี้ผู้วิจัยไม่สามารถสร้างโมเดลปัญญาประดิษฐ์ที่ใช้ข้อมูลรูปภาพจากมัลติมีเดียในการทำนายความสนใจในการจดของผู้เรียนได้อย่างแม่นยำและมีประสิทธิภาพที่จะนำไปประยุกต์ใช้จริงได้ เพราะ ความไม่สมบูรณ์ของข้อมูล จำนวนข้อมูลที่ไม่เพียงพอ ข้อจำกัดด้านองค์ความรู้ด้านการวิเคราะห์องค์ประกอบของมัลติมีเดียที่ส่งผลต่อความสนใจของผู้เรียน ความคลาดเคลื่อนจากโมเดล Mask R-CNN ที่ถูกฝึกโดยชุดข้อมูล COCO ที่ใช้ทำ Image Segmentation และอาจมีองค์ประกอบอื่นในมัลติมีเดียที่ส่งผลต่อความสนใจของผู้เรียนมากกว่าองค์ประกอบด้านรูปภาพ เช่น เสียง คำพูด เพลงประกอบ ภาพเคลื่อนไหว การใช้สีที่แตกต่างกัน (Um, Plass, Hayward & Homer, 2012) การแสดงออกทางอารมณ์ของภาพและเสียง ภาษาที่ใช้ข้อความในรูปภาพ ประเภทของวัตถุ ฯลฯ

#### 5.2 ข้อเสนอแนะ

เนื่องจากความคลาดเคลื่อนจากการทำ Image Segmentation โดย Mask R-CNN ผู้ทำวิจัยควรใช้โมเดล Image Segmentation ตัวอื่นและฝึกโดยใช้ชุดข้อมูลที่ใหญ่และหลากหลายมากขึ้น เพื่อให้สามารถจำแนกวัตถุได้อย่างแม่นยำและละเอียด

ชุดข้อมูลที่ใช้ฝึกโมเดลปัญญาประดิษฐ์ควรมีจำนวนมากและหลากหลายกว่านี้ เพื่อความแม่นยำและลดโอกาสที่โมเดลจะ เกิดการ overfit จากการใช้ข้อมูลที่คล้ายกันเกินไป หากชุดข้อมูลมีจำนวนมากพอ ผู้วิจัยจะสามารถใช้โมเดลที่มีความซับซ้อนมากขึ้นได้ ทำให้สามารถทำนายได้แม่นยำกว่าเดิม

การศึกษาและประยุกต์ใช้องค์ประกอบอื่นในมัลติมีเดียเพื่อให้มี Features มากขึ้นเพื่อใช้ในการทำนาย

เป็นสิ่งที่จำเป็นและคู่ควรแก่การศึกษา เพื่อให้สามารถวิเคราะห์การเรียนรู้และการจัดบันทึกของผู้เรียนได้อย่างครอบคลุมหลายด้านมากขึ้น ซึ่งสามารถช่วยเพิ่มโอกาสในการเพิ่มความแม่นยำในการทำนายได้

## บรรณานุกรม

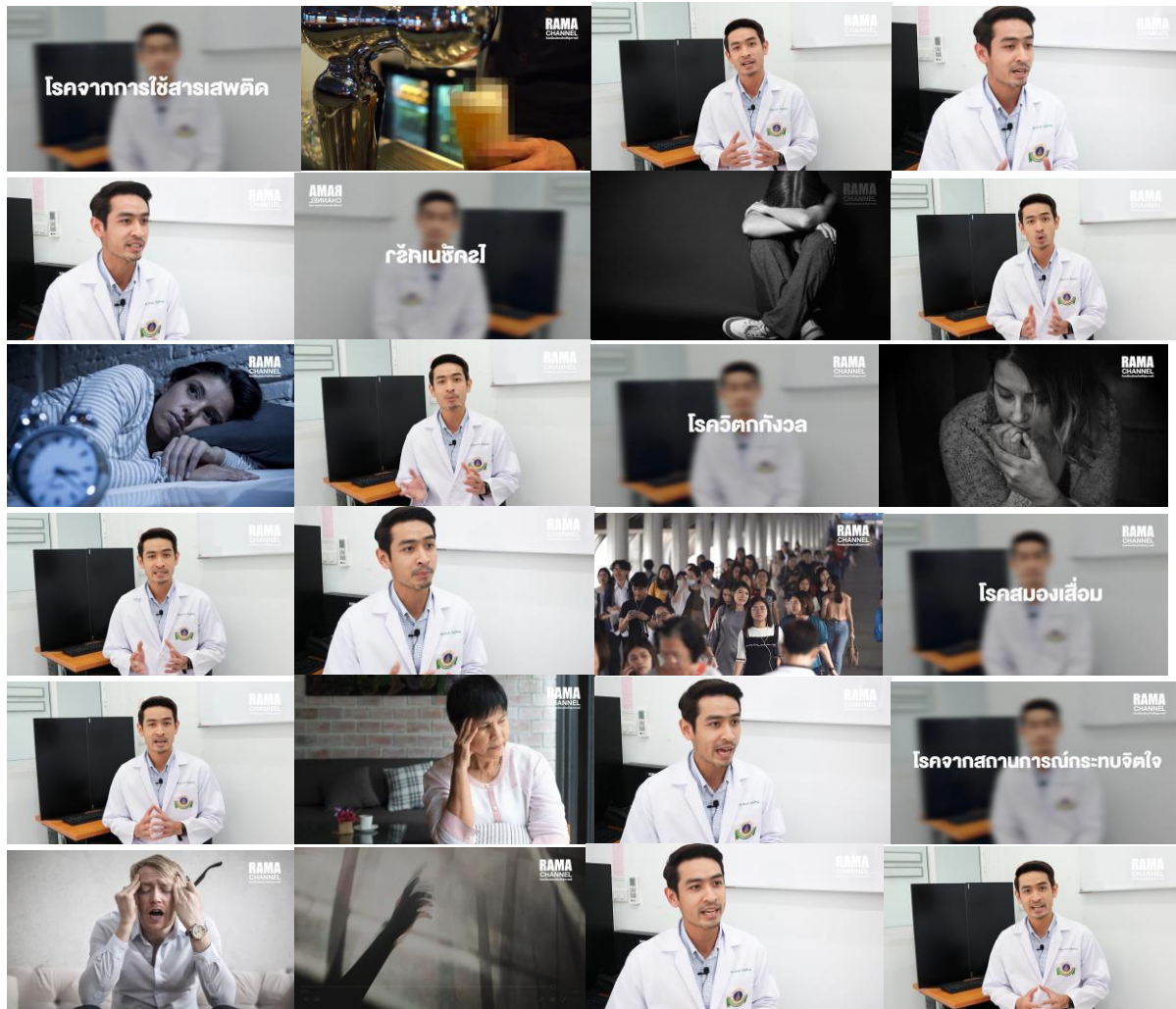
- Bohay, M., Blakely, D. P., Tamplin, A. K., & Radvansky, G. A. (2011). Note taking, review, memory, and comprehension. *The American Journal of Psychology*, 124, 63–73. doi:10.5406/amerjpsyc.124.1.0063
- Davydenko, A., Fildes, R.: Forecast error measures: critical review and practical recommendations. In: *Business forecasting: practical problems and solutions*. Wiley, Hoboken, NJ (2016)
- Herrington, D. (2010). Evaluation of learning efficiency and efficacy in a multi-user virtual environment. *Journal of Digital Learning in Teacher Education*, 27(2), 65–75. <https://doi.org/10.1080/21532974.2010.10784659>
- Jeon, G., Anisetti, M., Damiani, E. *et al.* Artificial intelligence in deep learning algorithms for multimedia analysis. *Multimed Tools Appl* 79, 34129–34139 (2020).
- Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick; Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2961-2969
- Mayer, Heiser, & Lonn. (2001). Cognitive Constraints on Multimedia Learning: When Presenting More Material Results in Less Understanding. *Journal of educational psychology.*, 93(1), 187.
- Mayer, & Moreno. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43 - 52.
- M. Al Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, and R. Yager. Deep learning approach for active classification of electrocardiogram signals. *Information Sciences*, 345:340–354, 2016. doi:10.1016/j.ins.2016.01.082
- Paula Arguello, David Morales, Yesid Fonseca, Henry Arguello, "Video-Tensor Completion using a Deep Learning approach", *Applications of Computational Intelligence 2020 IEEE Colombian Conference on*, pp. 1-6, 2020.
- P. Wlodarczak, J. Soar and M. Ally, "Multimedia data mining using deep learning," *2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC)*, 2015, pp. 190-196.
- Richard E. Mayer & Roxana Moreno (2003) Nine Ways to Reduce Cognitive Load in Multimedia Learning, *Educational Psychologist*, 38:1, 43-52, doi: 10.1207/S15326985EP3801\_6

- S. Malik and A. Agarwal. "Use of Multimedia as a New Educational Technology". Int. Journal of Information and Education Technology, Vol. 2, No. 5.
- Um, E. "R.", Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). Emotional design in multimedia learning. *Journal of Educational Psychology*, 104(2), 485–498. <https://doi.org/10.1037/a0026609>
- Van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review*, 22, 155– 174.
- Zhou, Z., Shin, J., Zhang, L., Gurudu, S., Gotway, M., Liang, J. (2017). Fine-tuning convolutional neural networks for biomedical image analysis: Actively and incrementally. In: CVPR. pp. 7340–7351 doi:10.1109/CVPR.2017.506





ภาคผนวก ข ภาพจากวิดีโอความรู้โรคทางจิตเวช



รูปที่ ข.1 ภาพจากวิดีโอความรู้โรคทางจิตเวชที่ยังไม่ถูกทำ Image Segmentation จำนวน 24 ภาพ



รูปที่ ข.2 ตัวอย่างภาพที่ถูกทำ Image Segmentation จากชุดข้อมูลภาพจากวิดีโอความรู้โรคทางจิตเวช

## ประวัติผู้ทำโครงการ

1. ชื่อ-นามสกุล: โชติพิสิฐ อุดลสีหัตต์

เลขประจำตัวนักเรียน: 07967

ระดับการศึกษา: มัธยมศึกษาปีที่ 6

สถานศึกษา: โรงเรียนมหิดลวิทยานุสรณ์

สถานที่ติดต่อ: 44/393 หมู่บ้านวิลลาจิโอพระราม 2 ซอยบางกระดี 35/1 ถนนบางกระดี

แขวงสามด้า เขตบางขุนเทียน กรุงเทพมหานคร 10150

โทรศัพท์: 096-189-6046

E-mail: s6207967@mwit.ac.th