

การตรวจจําบารมณัของผูัเรียนระหวางการเรียนออนไลน์ โดยใชัระบบปัญญาประดิษฐ์

ภูริชญา โชคเศรษฐกิจ

TNII

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

บัณฑิตศึกษา สาขาวิชาเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีไทย-ญี่ปุ่น

ปีการศึกษา 2566

DETECTING STUDENT EMOTIONS DURING ONLINE LEARNING
USING ARTIFICIAL INTELLIGENCE SYSTEMS

Phurichaya Choksettakit

TNII

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Information Technology

Gradute Studies

Thai-Nichi Institute of Technology

Academic Year 2023

หัวข้อวิทยานิพนธ์

การตรวจจับอารมณ์ของผู้เรียนระหว่างการเรียนออนไลน์
โดยใช้ระบบปัญญาประดิษฐ์

โดย

ภูริชญา โชคเศรษฐกิจ

สาขาวิชา

เทคโนโลยีสารสนเทศ

อาจารย์ที่ปรึกษาวิทยานิพนธ์

ดร.สรมย์พร เจริญพิทย์

บัณฑิตศึกษา สถาบันเทคโนโลยีไทย-ญี่ปุ่น อนุมัติให้บัณฑิตวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

..... รองอธิการบดีฝ่ายวิชาการ
(รองศาสตราจารย์ ดร.วรากร ศรีเชวงทรัพย์)

วันที่.....เดือน.....พ.ศ.....

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. ธงชัย แก้วกิริยา)

..... กรรมการ
(ดร.ภาสกร อภิรักษ์วรพินิต)

..... กรรมการ
(ดร.ณปภัช วิชัยดิษฐ์)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์
(ดร.สรมย์พร เจริญพิทย์)

ภริชญา โชคเศรษฐกิจ : การตรวจจ้บอารมณ์ของผู้เรียนระหว่างการเรียนออนไลน์ โดยใช้ระบบปัญญาประดิษฐ์. อาจารย์ที่ปรึกษา : ดร.สมรย์พร เจริญพิทย์, 120 หน้า.

เนื่องจากวิกฤตโรคระบาด COVID-19 ที่แพร่ระบาดในประเทศไทย และมีผลกระทบต่อการเรียนการสอนในชั้นเรียน เนื่องจากไม่สามารถรวมกลุ่มได้ จึงต้องเปลี่ยนรูปแบบในการเรียนการสอนเป็นการเรียนออนไลน์ ซึ่งเกิดปัญหาขึ้นเนื่องจากผู้สอนไม่สามารถทราบอารมณ์ของผู้เรียนแต่ละคนได้ว่ามีความสนใจเรียนมากน้อยเพียงใด งานวิจัยนี้จึงมุ่งเน้นไปที่การสร้างระบบตรวจจ้บอารมณ์ของผู้เรียนในระหว่างการเรียนออนไลน์ โดยใช้ระบบปัญญาประดิษฐ์ในการตรวจจ้บใบหน้าและอารมณ์ของผู้เรียน โดยใช้เทคโนโลยีโครงข่ายประสาทเทียมคอนโวลูชันเป็นหลัก

จากผลการทดลองภายในงานวิจัยนี้ ได้มีการจัดทำซอฟต์แวร์เพื่อตรวจจ้บภาพใบหน้าของผู้เรียนในระหว่างการเรียนออนไลน์ และสามารถตรวจจ้บใบหน้าและอารมณ์ของผู้เรียนในระหว่างการเรียน พร้อมทั้งจัดกลุ่มอารมณ์ของผู้เรียนออกเป็นสองกลุ่ม ได้แก่ กลุ่มที่มีความตั้งใจเรียน และกลุ่มที่ไม่ตั้งใจเรียน โดยในการประเมินประสิทธิภาพของงานวิจัยนี้ สามารถตรวจจ้บอารมณ์ของผู้เรียนที่แบ่งเป็นเจ็ดกลุ่มอารมณ์ ได้แก่ ปกติ , มีความสุข , ประหลาดใจ , โกรธ , ขยะแขยง , เศร้า และกลัว โดยมีค่าความถูกต้องเท่ากับ 69.2% และความถูกต้องในการจำแนกอารมณ์ที่แบ่งออกเป็นสองกลุ่มอารมณ์ คือ ตั้งใจเรียนและไม่ตั้งใจเรียน มีความถูกต้องเท่ากับ 84.65%.

บัณฑิตวิทยาลัย

สาขาวิชา เทคโนโลยีสารสนเทศ

ปีการศึกษา 2566

ลายมือชื่อนักศึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษา.....

PHURICHAYA CHOKSETTAKIT : DETECTING STUDENT EMOTIONS DURING
ONLINE LEARNING USING ARTIFICIAL INTELLIGENCE SYSTEMS. ADVISOR :
DR.SAROMPORN CHAROENPIT, 120 PP.

The COVID-19 pandemic has significantly impacted education in Thailand, forcing a shift from traditional classroom teaching to online learning. While online learning offers several advantages, it also presents challenges, particularly in assessing student engagement and attention levels. This research aims to develop an emotion detection system for online learners using artificial intelligence (AI). The system utilizes image processing and mostly using Convolutional Neural Network techniques to analyze students' facial expressions and emotions.

The researchers developed software that goes beyond attendance verification by incorporating emotion recognition. It can identify student faces and categorize their emotional state as attentive or inattentive during online sessions. While the system exhibits a 69.2% accuracy rate in recognizing seven emotions (angry, fear, disgust, happy, sadness, surprise, neutral). The research findings demonstrate a notable level of accuracy, with the two emotions classification 84.65%.

Graduate Studies

Field of Study Information Technology

Academic Year 2023

Student's Signature.....

Advisor's Signature.....

กิตติกรรมประกาศ

การทำวิทยานิพนธ์นี้จะสำเร็จลุล่วงไปได้ด้วยดีนั้น ผู้วิจัยขอกราบขอบพระคุณในความกรุณาของ ดร.สรมย์พร เจริญพิทย์ ซึ่งได้สละเวลาในการให้คำแนะนำ คำปรึกษา การสนับสนุนต่างๆ รวมถึงตรวจสอบข้อบกพร่องทุกขั้นตอน กำลังใจในการทำงาน และความเอาใจใส่อย่างดียิ่งตลอดระยะเวลาในการทำวิทยานิพนธ์ฉบับนี้ อีกทั้งคณาจารย์คณะเทคโนโลยีสารสนเทศ เจ้าหน้าที่ที่อำนวยความสะดวกต่าง ๆ ในการทำวิทยานิพนธ์

เหนือสิ่งอื่นใดผู้วิจัยขอกราบขอบพระคุณบิดา มารดา ของผู้วิจัยที่คอยให้กำลังใจ และสนับสนุนในทุก ๆ ด้านอย่างเสมอมา และเพื่อน ๆ ที่คอยให้การสนับสนุนและเป็นกำลังใจในเวลาที่ผ่านมา

ผู้วิจัยหวังว่างานวิทยานิพนธ์ฉบับนี้จะมีประโยชน์สำหรับผู้ที่ต้องการศึกษาการตรวจจับอารมณ์จากใบหน้า และผู้ที่มีความสนใจในเรื่องนี้ สามารถนำไปต่อยอดองค์ความรู้ เพื่อประยุกต์ใช้ให้เกิดประโยชน์ในการใช้งานจริงได้

ภุริชญา โชคเศรษฐกิจ

TNI

TAI

NICHI INSTITUTE OF TECHNOLOGY

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฌ
สารบัญรูป.....	ฎ
 บทที่	
1 บทนำ.....	1
1.1 ที่มาและปัญหา.....	1
1.2 วิธีการในการแก้ไขปัญหา.....	4
1.3 วัตถุประสงค์ของงานวิจัยและเป้าหมาย.....	4
1.4 ขอบเขตของงานวิจัย.....	4
1.4.1 ขอบเขตด้านผู้เรียน.....	4
1.4.2 ขอบเขตด้านเนื้อหาในการเรียน.....	4
1.4.3 ขอบเขตด้านอุปกรณ์ และสภาพแวดล้อมในการทดลอง.....	4
1.5 ประโยชน์ของงานวิจัย.....	4
1.6 นิยามศัพท์เฉพาะ.....	5
1.6.1 COVID-19.....	5
1.6.2 เทคโนโลยีตรวจจับใบหน้า (Face detection).....	5
1.6.3 Deep Learning Model.....	5
1.6.4 ปัญญาประดิษฐ์ (Artificial Intelligence : AI).....	5
1.6.5 การแสดงออกทางสีหน้า (Facial Expression).....	6
1.6.6 Region of Interest (ROI).....	6
2 ทฤษฎี หลักการ และ งานวิจัยที่เกี่ยวข้อง.....	7
2.1 เทคโนโลยีปัญญาประดิษฐ์.....	7

สารบัญ (ต่อ)

บทที่		หน้า
2	2.2 เทคโนโลยีการตรวจจับใบหน้า (Face Detection)	8
	2.2.1 เทคนิคแบบใช้ลักษณะเฉพาะ (feature-based detection).....	8
	2.2.2 เทคนิคแบบใช้การเรียนรู้ของเครื่อง (machine learning-based detection)	8
	2.3 การตรวจสอบลักษณะอารมณ์ทางใบหน้า (Facial Emotional Expression Detection)	8
	2.4 โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network: CNN).....	9
	2.5 Resnet50.....	11
	2.6 MTCNN.....	13
	2.7 การจำแนกลักษณะอารมณ์.....	15
	2.8 การขยับของกล้ามเนื้อใบหน้า.....	16
	2.9 การเปลี่ยนแปลงใบหน้าตามลักษณะอารมณ์.....	23
	2.10 FER-2013.....	26
	2.11 การวัดประสิทธิภาพของโมเดล.....	28
	2.12 เทคโนโลยีที่เกี่ยวข้อง.....	30
	2.12.1 Visual Studio Code.....	30
	2.12.2 Google API.....	31
	2.12.3 Python.....	31
	2.12.4 Simple Mail Transfer Protocol : SMTP.....	31
	2.12.5 Google Drive.....	32
	2.13 งานวิจัยที่เกี่ยวข้อง.....	32
3	ระเบียบและวิธีการดำเนินการวิจัย.....	39
	3.1 การวิเคราะห์ข้อมูลปัญหาของงานวิจัย.....	40
	3.2 การศึกษาข้อมูลและเทคโนโลยีที่นำมาใช้.....	41
	3.3 การออกแบบการทดลอง.....	42

สารบัญ (ต่อ)

บทที่		หน้า
3	3.4 การพัฒนาโมเดลเพื่อตรวจจ็บบรรณของผูเรียน.....	46
	3.5 ขั้นตอนในการดำเนินงานวิจัย.....	47
4	ผลการวิจัย.....	54
	4.1 ผลการตรวจสอบความถูกต้องของโมเดล.....	54
	4.2 ผลการพัฒนาซอฟต์แวร์ในการตรวจจ็บบรรณและบรรณของผูเรียน.....	56
	4.3 ผลการประเมินประสิทธิภาพซอฟต์แวร์ในการตรวจจ็บบรรณและ บรรณของผูเรียน.....	61
5	บทสรุป อภิปราย และข้อเสนอแนะ.....	76
	5.1 ข้อมำกัตในงานวิจัย.....	77
	5.2 ปัญหาที่พบในงานวิจัย.....	77
	5.3 ข้อเสนอแนะเพิ่มเติม.....	78
	บรรณานุกรม.....	79
	ภาคผนวก.....	85
	ภาคผนวก ก. โค้ดโปรแกรม (System Code).....	86
	ภาคผนวก ข. ผลงานที่ได้รับการตีพิมพ์.....	113
	ประวัติย่อผูวิจัย.....	121

สารบัญตาราง

ตาราง	หน้า
2.1 แสดงการขยับของกล้ามเนื้อใบหน้า.....	17
2.2 แสดงจำนวนข้อมูลภาพในชุดข้อมูล FER-2013.....	27
2.3 ตารางสรุปวิธีการหรือเทคนิคที่ใกล้เคียงกับวิธีการของงานวิจัย.....	34
3.1 แผนในการดำเนินงานวิจัย.....	40
3.2 แสดงหัวข้อในบทเรียนแต่ละส่วน.....	44
4.1 จำนวนวินาทีที่สามารถตรวจจับใบหน้าของผู้เรียนได้.....	62
4.2 เปอร์เซ็นต์ที่สามารถตรวจจับใบหน้าของผู้เรียนได้.....	63
4.3 จำนวนวินาทีที่สามารถตรวจจับใบหน้าที่ไม่เคยตรวจสอบว่าตั้งใจเรียนของผู้เรียน...	66
4.4 เปอร์เซ็นต์ความตั้งใจเรียนตามที่ไม่เคยตรวจสอบได้ของผู้เรียน.....	67
4.5 จำนวนวินาทีที่สามารถตรวจจับใบหน้า ที่ไม่เคยตรวจสอบว่าผู้เรียนไม่ตั้งใจเรียน...	71
4.6 เปอร์เซ็นต์ที่ไม่เคยตรวจสอบได้ว่าผู้เรียนไม่ตั้งใจเรียน.....	72



สารบัญรูป

รูป	หน้า
1.1 นักเรียนที่ไม่จดจ่อกับบทเรียน.....	3
1.2 นักเรียนที่ตั้งใจเรียน.....	3
2.1 แสดงการทำงานของ Residual Network.....	11
2.2 แสดงโครงสร้างของ ResNet-34.....	12
2.3 แสดงสถาปัตยกรรมของ MTCNN.....	14
2.4 แสดงการหากรอบพื้นที่ของใบหน้าภายในขั้นตอน P-Net , R-Net และ O-Net.....	15
2.5 แสดงโมเดลของ Russell.....	16
2.6 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์โกรธ.....	23
2.7 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์กลัว.....	24
2.8 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์ขะแยง.....	24
2.9 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์มีความสุข.....	25
2.10 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์เป็นกลาง.....	25
2.11 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์เศร้า.....	26
2.12 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์เศร้า.....	26
2.13 แสดงรูปตัวอย่างของ dataset.....	28
2.14 แสดง Confusion Matrix.....	30
3.1 แสดงการออกแบบระบบ.....	43
3.2 การตรวจจับใบหน้า.....	45
3.3 แสดงการปรับตัวแปรในการเรียนรู้ของโมเดล.....	46
3.4 แสดงขั้นตอนในการดำเนินงานวิจัยทั้งหมด.....	47
3.5 แสดงขั้นตอนย่อยในการทำการทดลอง.....	49
3.6 แสดงการนั่งประจำจุดการทดลองของผู้ทำการทดลอง.....	50
3.7 แสดงการเปิดวิดีโอเพื่อเตรียมทำการทดลอง.....	50
3.8 แสดงการเข้าสู่ระบบ.....	51
3.9 หน้าจอแจ้งเตือนจาก Google.....	51
3.10 แสดงการขอสิทธิ์ในการอัปโหลดไฟล์.....	52

สารบัญรูป (ต่อ)

รูป	หน้า
3.11 หน้าจอแสดงการเข้าสู่ระบบสำเร็จ.....	52
3.12 แสดงอีเมลที่ได้รับจากการรันโปรแกรม.....	53



บทที่ 1

บทนำ

1.1 ที่มาและปัญหา

การแพร่ระบาดของไวรัสโคโรนาสายพันธุ์ใหม่ 2019 (novel coronavirus 2019, 2019-nCoV) [1] ที่ก่อโรคปอดอักเสบ (pneumonia) ในเมืองอู่ฮั่น (Wuhan) มณฑลหูเป่ย์ (Hubei) ประเทศจีน เริ่มต้นในช่วงปลายปี ค.ศ. 2019 จนถึงปัจจุบัน ในช่วงแรกคาดว่า เป็นการติดเชื้อจากสัตว์สู่คน โดยเมื่อการแพร่ระบาดของโรคเพิ่มขึ้นอย่างรวดเร็ว ส่งผลให้ประเทศต่างๆ ทั่วโลกมีการเฝ้าระวังโรคระบาดนี้อย่างใกล้ชิด เนื่องจากเหตุการณ์โรคระบาดทางเดินหายใจร้ายแรงที่เกิดจากไวรัสโคโรนา ได้แก่ โรคซาร์ (severe acute respiratory syndrome, SARS) ซึ่งเคยระบาดในช่วงปี ค.ศ. 2002-2003 ซึ่งมีสาเหตุจากไวรัสโคโรนาสายพันธุ์ SARS-CoV ที่เป็นไวรัสโคโรนาข้ามสปีชีส์จากค้างคาวผ่าน civet cat (ชะมด) มาติดเชื้อในมนุษย์ โดยเริ่มต้นการระบาดจากประเทศจีน และแพร่กระจายไปทั่วโลก โดยมีผู้ติดเชื้อมากกว่าแปดพันคน และมีอัตราการเสียชีวิตร้อยละ 10 และเพิ่มเป็นร้อยละ 50 ในผู้สูงอายุ ต่อมาในปีค.ศ. 2012-2014 ก็มีการระบาดของไวรัสโคโรนาสายพันธุ์ใหม่ คือ Middle East respiratory syndrome coronavirus (MERS-CoV) ซึ่งเป็นไวรัสโคโรนาข้ามสปีชีส์จากค้างคาว ผ่านอูฐ และติดเชื้อในมนุษย์เป็นลำดับต่อมา โดยเริ่มต้นการระบาดจากผู้ป่วยในประเทศซาอุดีอาระเบีย โดยเหตุการณ์ในครั้งนี้มีผู้ติดเชื้อรวม 1,733 คน และมีอัตราการเสียชีวิตมากถึงร้อยละ 36

สำหรับประเทศไทย [2] พบผู้ป่วยหญิงจากเมืองอู่ฮั่น ประเทศจีน ติดเชื้อโควิด-19 ในไทยคนแรกกระทรวงสาธารณสุขของไทยประกาศเมื่อวันที่ 12 มกราคม พ.ศ.2563 ว่าพบนักท่องเที่ยวหญิงวัย 61 ปี สัญชาติจีน ซึ่งมีภูมิลำเนาอยู่ที่เมืองอู่ฮั่น ประเทศจีน ติดเชื้อโควิด-19 ซึ่งเป็นการพบผู้ติดเชื้อคนแรกนอกประเทศจีน และจากข้อมูลของกระทรวงสาธารณสุข ตั้งแต่มีรายงานพบผู้ติดเชื้อชาวจีนคนแรกที่เดินทางมาจากเมืองอู่ฮั่นในวันที่ 12 มกราคม พ.ศ.2563 เป็นต้นมา ช่วงระหว่างวันที่ 12-31 มกราคม พ.ศ.2563 ยังพบว่าในประเทศไทยมีผู้ป่วยชาวจีนจากเมืองอู่ฮั่น รวม 16 คน แบ่งเป็นในพื้นที่กรุงเทพมหานคร 13 คน จังหวัดประจวบคีรีขันธ์ ชลบุรี และภูเก็ต จังหวัดละ 1 คน จนเมื่อวันที่ 25 มีนาคม พ.ศ.2563 รัฐบาลไทยได้ประกาศสถานการณ์ฉุกเฉินทั่วประเทศ จึงส่งผลให้ประชาชนต้องปรับเปลี่ยนวิถีชีวิตการทำงาน ความเป็นอยู่ รวมถึงผู้เรียนนักศึกษาที่ต้องเปลี่ยนรูปแบบในการเรียน เป็นการใช้ชีวิตในรูปแบบปกติใหม่ (New Normal) ซึ่งเกิดจากคำแนะนำในการป้องกันการติดเชื้อและแพร่เชื้อให้ผู้อื่นมาถือปฏิบัติในชีวิตประจำวัน ไม่ว่าจะเป็นการล้างมือเมื่อเหมาะสม หลีกเลี่ยงการ เว้นระยะห่างทางสังคม (Social Distance) สวมหน้ากากอนามัยในที่สาธารณะ และตรวจวัดอุณหภูมิร่างกายตามสถานที่ต่างๆ ลดสัมผัสและการเดินทางที่ไม่จำเป็น ซึ่งในการลดการ

เดินทางนี้ ส่งผลให้วัยทำงานต้องมีการปรับเปลี่ยนวิถีชีวิตเป็นการทำงานที่บ้าน และในวัยเรียน เปลี่ยนเป็นการเรียนการสอนออนไลน์ทดแทน

การศึกษาเป็นกลไกสำคัญในการพัฒนาทรัพยากรมนุษย์เพื่อขับเคลื่อนประเทศ [3] และเป็นช่องทางสำคัญในการเปิดโอกาสให้กับประชาชนทุกคน โดยเฉพาะอย่างยิ่งกับเยาวชนไทย เพื่อให้มีความรู้เพื่อนำไปประกอบอาชีพในอนาคตได้ แต่การศึกษาไทยนั้นเกิดปัญหาในหลายด้านด้วยกัน ทั้งในแง่ของคุณภาพและปริมาณ โดยสามารถอธิบายได้ดังต่อไปนี้

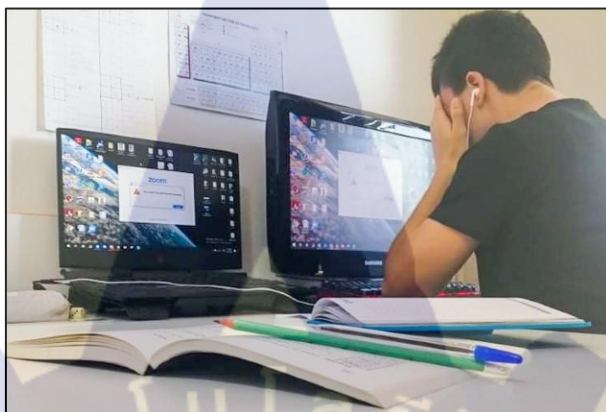
ในแง่ของคุณภาพนั้น สามารถพิจารณาได้จากผลคะแนนและผลประเมิน PISA ที่น้อยกว่าประเทศในภูมิภาค ASEAN ด้วยกัน ซึ่งรวมไปถึงในแง่ของปริมาณ เช่น ความเหลื่อมล้ำในการศึกษาไทยระหว่างสถาบันการศึกษาของเอกชนและสถาบันการศึกษาของรัฐบาล รวมทั้งการขาดโอกาสทางการศึกษาของเยาวชนในหลายพื้นที่ เช่น ภูมิภาคตะวันออกเฉียงเหนือที่ห่างไกลจากความช่วยเหลือของหลายภาคส่วน หรือบางสถานศึกษาที่ขาดแคลนบุคลากรหรือทรัพยากรการเรียนการสอนที่ควรเป็นปัจจัยขั้นพื้นฐานของทุกสถานศึกษา

ด้วยสถานการณ์วิกฤตจากโรคระบาด COVID-19 ที่เริ่มแพร่กระจายเข้ามาในประเทศไทย ส่งผลให้การเรียนการสอนที่สถานศึกษาเกิดการหยุดชะงักไปตั้งแต่ต้นปีการศึกษา 2563 และส่งผลให้ผู้เรียนและนักศึกษาทั่วประเทศมีความจำเป็นต้องรับการศึกษาผ่านช่องทางอื่นเป็นหลัก เช่น การเรียนทางออนไลน์ โทรทัศน์ดาวเทียม ซึ่งจากวิกฤตนี้ ส่งผลให้ปัญหาด้านคุณภาพของการศึกษาไทยนั้นลดระดับลงอย่างเห็นได้ชัด

ผลกระทบในด้านการศึกษาที่เกิดขึ้นจากการเรียนผ่านรูปแบบออนไลน์นั้น สามารถประเมินได้อย่างชัดเจนจากผลประเมิน PISA ในปี 2022 [4] โดยผู้เรียนไทยมีคะแนนเฉลี่ยด้านคณิตศาสตร์ 394 คะแนน ด้านวิทยาศาสตร์ 409 คะแนน และด้านการอ่าน 379 คะแนน ซึ่งเมื่อเทียบกับ PISA 2018 พบว่า คะแนนเฉลี่ยของประเทศไทยทั้งสามด้านลดลง โดยด้านคณิตศาสตร์มีคะแนนเฉลี่ยลดลง 25 คะแนน ส่วนด้านวิทยาศาสตร์และการอ่าน มีคะแนนเฉลี่ยลดลง 17 คะแนน และ 14 คะแนน ตามลำดับ ทั้งนี้ ผลการประเมินของประเทศไทยตั้งแต่ PISA 2000 จนถึง PISA 2022 พบว่า คะแนนเฉลี่ยด้านคณิตศาสตร์และการอ่านมีแนวโน้มลดลง ส่วนด้านวิทยาศาสตร์ถือว่าไม่เปลี่ยนแปลงทางสถิติ โดยภายในช่วงปีการศึกษา 2563-2564 นั้น เป็นช่วงที่ประเทศไทยใช้ระบบการเรียนการสอนออนไลน์เป็นหลัก และส่งผลอย่างชัดเจนต่อประสิทธิภาพในการเรียนรู้

เนื่องจากสภาพแวดล้อมในการเรียนออนไลน์ จะขาดการสื่อสารแบบ face-to-face ซึ่งแตกต่างจากการเรียนแบบดั้งเดิม ซึ่งผู้สอนสามารถรับรู้อารมณ์ของผู้เรียนในชั้นเรียนได้เช่นเดียวกับที่มีการปฏิสัมพันธ์และการตอบโต้กันระหว่างผู้สอนและผู้เรียน ในการเรียนการสอนออนไลน์มีข้อจำกัดในด้านสภาพแวดล้อม ซึ่งทำให้เกิดความเข้าใจที่ไม่สมบูรณ์ได้ [5] โดยรูปที่ 1.1 แสดงรูปพฤติกรรมของ

ผู้เรียนในชั้นเรียนออนไลน์ที่ไม่ตั้งใจเรียน และรูปที่ 1.2 แสดงรูปที่ได้จากการเรียนออนไลน์ และผู้เรียนมีพฤติกรรมที่ตั้งใจเรียน



รูปที่ 1.1 นักเรียนที่ไม่จดจ่อกับบทเรียน



รูปที่ 1.2 นักเรียนที่ตั้งใจเรียน

เนื่องจากอารมณ์ของผู้เรียนส่งผลกระทบอย่างมากต่อกระบวนการเรียนรู้ โดยอารมณ์ของผู้เรียน จะส่งผลต่อความตั้งใจเรียน แรงจูงใจในการเรียน วิธีในการเรียนรู้ [6] ในขณะที่ระบบการเรียนรู้ออนไลน์ในปัจจุบันไม่สามารถตรวจจับอารมณ์ของผู้เรียนได้ จึงเป็นเรื่องยากสำหรับผู้เรียนที่จะมีสมาธิกับบทเรียนอย่างต่อเนื่องเป็นเวลานาน และไม่สามารถควบคุมสมาธิในการเรียนรู้ได้

1.2 วิธีการในการแก้ไขปัญหา

ในงานวิจัยนี้ เราได้กล่าวถึงแนวทางแก้ไขปัญหาที่กล่าวมาข้างต้น โดยการออกแบบระบบการเรียนรู้ออนไลน์ที่เน้นด้านอารมณ์ วิธีอธิบายอารมณ์ของผู้เรียน วิธีประเมินอารมณ์ของผู้เรียนด้วยสีหน้า โดยเป้าหมายของงานวิจัยนี้ คือ การออกแบบระบบที่สามารถตรวจจับอารมณ์ของผู้เรียน โดยจำแนกออกเป็นสองกลุ่มอารมณ์ โดยแบ่งเป็นกลุ่มอารมณ์ในเชิงบวก (ตั้งใจเรียน) และอารมณ์ในเชิงลบ (ไม่ตั้งใจเรียน) รวมทั้งสามารถบันทึกวิดีโอพฤติกรรมของผู้เรียนในขณะที่เรียนออนไลน์ได้

1.3 วัตถุประสงค์ของงานวิจัยและเป้าหมาย

- 1) เพื่อพัฒนาซอฟต์แวร์ในการตรวจจับใบหน้าและอารมณ์ของผู้เรียน
- 2) เพื่อประเมินประสิทธิภาพซอฟต์แวร์ในการตรวจจับใบหน้าและอารมณ์ของผู้เรียน

1.4 ขอบเขตของงานวิจัย

1.4.1 ขอบเขตด้านผู้เรียน

- กลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญาตรีของสถาบันเทคโนโลยีไทย-ญี่ปุ่น

1.4.2 ขอบเขตด้านเนื้อหาในการเรียน

- เนื้อหาที่ใช้ในการเรียน คือ คลิปการสอนวิชา data science ในหัวข้อ Machine Learning ความยาว 20.00 นาที
- ผู้สอนคืออาจารย์ในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีไทย-ญี่ปุ่น

1.4.3 ขอบเขตด้านอุปกรณ์ และสภาพแวดล้อมในการทดลอง

- ใช้งานคอมพิวเตอร์ตั้งโต๊ะ หรือโน้ตบุ๊กคอมพิวเตอร์ ซึ่งมีรายละเอียดดังนี้
 - 1) สามารถเชื่อมต่ออินเทอร์เน็ตได้
 - 2) มีกล้องที่สามารถจับภาพผู้เรียนได้
 - 3) มีพื้นที่ในการจัดเก็บบนคอมพิวเตอร์เหลือมากกว่า 1 GB
- ในขณะที่ทำการทดลอง ต้องมีผู้เรียนเพียงคนเดียวอยู่ในขอบเขตการจับภาพของกล้อง

1.5 ประโยชน์ของงานวิจัย

- 1) สามารถตรวจสอบอารมณ์ของผู้เรียนในระหว่างเรียนออนไลน์ได้
- 2) สามารถตรวจจับใบหน้าและอารมณ์ของผู้เรียนโดยใช้ซอฟต์แวร์ที่จัดทำขึ้นได้

- 3) ผู้สอนสามารถปรับเนื้อหาในการสอนได้สะดวกขึ้น เมื่อทราบอารมณ์ของผู้เรียนในช่วงเวลาที่ผ่านมา
- 4) สามารถประเมินความตั้งใจและความเข้าใจของผู้เรียนในระหว่างการเรียนรู้ออนไลน์ได้
- 5) สามารถนำข้อมูลให้ผู้เรียนไปใช้ในการจัดการกับอารมณ์ด้านลบของตนเองที่เกิดจากชั้นเรียน เช่น ความเบื่อ ความวิตกกังวล หรือความโกรธ ตลอดจนการให้ความสนใจในชั้นเรียน และสามารถเพิ่มแรงจูงใจ และประสิทธิภาพในการเรียนรู้

1.6 นิยามศัพท์เฉพาะ

1.6.1 COVID-19 คือ โรคที่มีสาเหตุจากเชื้อไวรัสโคโรนาสายพันธุ์ใหม่ซึ่งพบเป็นครั้งแรกในเมืองอู่ฮั่น ประเทศจีน มีชื่อว่า โรคติดเชื้อไวรัสโคโรนา 2019 หรือ โควิด-19 (COVID19) – ‘CO’ มาจากคำว่า Corona, ‘VI’ มาจาก Virus, และ ‘D’ มาจาก Disease ที่แปลว่า ‘โรค’ ไวรัสโควิด-19 เป็นไวรัสชนิดใหม่ซึ่งมีความเชื่อมโยงกับตระกูลของไวรัสที่เป็นต้นเหตุของโรคซาร์ส (Severe Acute Respiratory Syndrome – SARS) หรือโรคทางเดินหายใจเฉียบพลันรุนแรง รวมทั้งโรคหัดธรรมดาบางประเภท

1.6.2 เทคโนโลยีตรวจจับใบหน้า (Face detection) คือ ระบบวิเคราะห์ใบหน้า ซึ่งถือว่าเป็นหนึ่งในระบบที่ใช้ในการพิสูจน์ยืนยันตัวตนบุคคลโดยใช้คุณลักษณะจำเพาะทางสรีระ โดยระบบรู้จำใบหน้าจะทำงานโดยการเปรียบเทียบใบหน้าจากภาพถ่ายดิจิทัล หรือภาพจากกล้องวิดีโอของบุคคลที่เราสนใจกับฐานข้อมูลใบหน้าที่มีอยู่ และเมื่อเปรียบเทียบเสร็จก็จะแสดงผลใบหน้าที่อยู่ในฐานข้อมูลที่มีใบหน้าเหมือนกับภาพที่นำมาเปรียบเทียบออกมา

1.6.3 Deep Learning Model คือ วิธีการเรียนรู้แบบอัตโนมัติด้วยการเลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neurons) โดยนำระบบโครงข่ายประสาท (Neural Network) มาซ้อนกันหลายชั้น (Layer) และทำการเรียนรู้ข้อมูลตัวอย่างซึ่งข้อมูลดังกล่าวจะถูกนำไปใช้ในการตรวจจับรูปแบบ (Pattern) หรือจัดหมวดหมู่ข้อมูล (Classify the Data)

1.6.4 ปัญญาประดิษฐ์ (Artificial Intelligence : AI) คือ เทคโนโลยีที่พัฒนาขึ้น และสามารถทำความเข้าใจ เรียนรู้องค์ความรู้ต่างๆ ได้ สามารถเรียนรู้สิ่งต่างๆ เพิ่มเติมได้

1.6.5 การแสดงออกทางสีหน้า (Facial Expression) คือ การขยับของกล้ามเนื้อใบหน้า ซึ่งจะแสดงสีหน้าแตกต่างกันไปตามอารมณ์ของผู้พูด โดยจะจัดประเภทอารมณ์ที่แสดงออกทางสีหน้าของมนุษย์ เช่น เมื่อเกิดอารมณ์โกรธจะมีหน้าตาบึ้งตึง มีความสุขจะยิ้ม

1.6.6 Region of Interest (ROI) คือ บริเวณที่เราสนใจ อาจจะเป็นบริเวณใดภายในภาพก็ได้โดยการตีกรอบล้อมรอบบริเวณที่สนใจด้วยวงกลม กรอบสี่เหลี่ยม หรือกรอบรูปเหลี่ยมใดๆ เพื่อ

นำภาพเฉพาะส่วนดังกล่าวมาประมวลผล หรือเปลี่ยนแปลงภาพตามต้องการ โดยไม่มีผลกระทบต่อส่วนอื่นๆ ซึ่งภายในหนึ่งภาพ สามารถกำหนดได้หลายๆ บริเวณที่สนใจ



บทที่ 2

ทฤษฎี หลักการ และ งานวิจัยที่เกี่ยวข้อง

หลังจากสถานการณ์ COVID-19 ในปัจจุบันจึงมีการนำเทคโนโลยีเข้ามาเกี่ยวข้องกับการเรียนการสอนมากขึ้น เช่น หลักสูตรปริญญาโทที่มีการเรียนการสอนแบบออนไลน์ เพื่อดึงดูดกลุ่มผู้เรียนในวัยทำงานหรือมีข้อจำกัดในการเข้าเรียนในชั้นเรียน [7] หรือการพัฒนากระบวนการเรียนการสอนให้แบบออนไลน์ร่วมด้วย [8-10] เนื่องจากการเรียนการสอนออนไลน์นั้นส่งผลให้ประสิทธิภาพในการเรียน ความเข้าใจของผู้เรียน และการมีปฏิสัมพันธ์ระหว่างผู้เรียนและผู้สอนลดน้อยลง ภายใต้งานวิจัยนี้จึงมีการนำทฤษฎี หลักการ และงานวิจัยที่เกี่ยวข้องต่างๆ เพื่อนำมาสร้างโปรแกรมที่จะแก้ไขปัญหาภายในงานวิจัยนี้

2.1 เทคโนโลยีปัญญาประดิษฐ์

ปัญญาประดิษฐ์ (Artificial Intelligence : AI) เป็นการรวมความหมายของคำศัพท์ 2 คำ คือ Artificial ซึ่งมีความหมายว่า การประดิษฐ์ขึ้น และ Intelligence ซึ่งแปลว่าความฉลาด เมื่อรวมความหมายกันแล้ว จึงหมายถึงความฉลาดเทียมที่สร้างขึ้นให้กับสิ่งที่ไม่มีชีวิต ปัญญาประดิษฐ์เป็นสาขาหนึ่งในด้านวิทยาการคอมพิวเตอร์ และวิศวกรรม ซึ่งสาขาปัญญาประดิษฐ์เป็นการเรียนรู้เกี่ยวกับกระบวนการการคิด การกระทำ การให้เหตุผล การปรับตัว หรือการอนุมาน และการทำงานของสมอง โดยคำนิยามแรกของปัญญาประดิษฐ์ มีที่มาจาก J. McCarthy [11] ในปีคริสต์ศักราช 1996 ซึ่งได้กล่าวว่า The Goal of AI is to Develop Machines That Behave as Though They were Intelligent ซึ่งมีความหมายว่า เป้าหมายของปัญญาประดิษฐ์ คือการประดิษฐ์เครื่องจักรซึ่งสามารถกล่าวได้ว่า เครื่องจักรเหล่านั้นมีความเป็นอัจฉริยะ [12-15] ได้อธิบายถึงความหมายของปัญญาประดิษฐ์ว่า ปัญญาประดิษฐ์คือการศึกษารื่องของการได้รับความรู้จากสิ่งแวดล้อม การประมวลผลข้อมูลเหล่านั้น และดำเนินการที่ส่งผลกระทบต่สิ่งแวดล้อมนั้น ซึ่งจะแก้ไขปัญหาที่เป็นปัญหาโดยทั่วไปในการวิจัยด้านปัญญาประดิษฐ์ รวมถึงการคิด, การแสดงความรู้, การวางแผน, การเรียนรู้, การประมวลผลภาษาธรรมชาติ (Natural Language Processing), การรับรู้, และความสามารถในการเคลื่อนไหวและจัดการวัตถุ

จากคำนิยามที่ได้กล่าวมาข้างต้นแล้ว สามารถสรุปความหมายโดยรวมของคำว่าปัญญาประดิษฐ์ได้ว่า เป็นเครื่องจักรที่มีความสามารถในการทำความเข้าใจ เรียนรู้องค์ความรู้ต่าง สามารถมีการรับรู้ การเรียนรู้ การให้เหตุผล และการแก้ปัญหาต่างๆ ได้ โดยเทคโนโลยีปัญญาประดิษฐ์ที่มีการ

นำมาใช้งานภายในงานวิจัยนี้ คือ ปัญญาประดิษฐ์ที่สามารถเรียนรู้ และแยกส่วนของใบหน้าออกจากรูปภาพได้ ได้แก่ เทคโนโลยีการตรวจจับใบหน้า

2.2 เทคโนโลยีการตรวจจับใบหน้า (Face Detection)

เทคโนโลยีการตรวจจับใบหน้า หรือ Face Detection [16-21] คือกระบวนการในการตรวจสอบใบหน้า ซึ่งจะทำการตรวจสอบเพื่อหาลักษณะประกอบต่างๆบนใบหน้า ได้แก่ คิ้ว ดวงตา จมูก ริมฝีปาก โดยเทคโนโลยีตรวจจับใบหน้านั้นสามารถแบ่งได้เป็น 2 หมวดหมู่ คือ เทคนิคแบบใช้ลักษณะเฉพาะ และเทคนิคแบบใช้การเรียนรู้ของเครื่อง

2.2.1 เทคนิคแบบใช้ลักษณะเฉพาะ (feature-based detection)

การทำงานของงานเทคนิคนี้ จะเป็นการวิเคราะห์ลักษณะเฉพาะบนใบหน้า เช่น ตา จมูก ปาก เป็นต้น โดยใช้เทมเพลตหรือแบบจำลองทางสถิติเพื่อระบุลักษณะเฉพาะเหล่านี้ โดยอัลกอริทึมที่ได้รับความนิยม ได้แก่ Viola-Jones และ Haar cascades

2.2.2 เทคนิคแบบใช้การเรียนรู้ของเครื่อง (machine learning-based detection)

การทำงานของงานเทคนิคนี้ ทำงานโดยการสร้างโมเดลการเรียนรู้ของเครื่องที่ได้รับการฝึกฝนบนชุดข้อมูลภาพที่มีป้ายกำกับ โมเดลจะเรียนรู้รูปแบบและสามารถแยกแยะใบหน้าออกจากวัตถุอื่นๆ ได้ เทคนิคนี้ให้ผลลัพธ์ที่แม่นยำและทนทานต่อการเปลี่ยนแปลงของสภาพแวดล้อมมากกว่า เทคนิคแบบใช้ลักษณะเฉพาะ ตัวอย่าง ได้แก่ Convolutional Neural Networks (CNNs) และ YOLO (You Only Look Once) ซึ่งในเทคนิคนี้จะมีการทำงานของเทคโนโลยีปัญญาประดิษฐ์มาใช้งานในการเรียนรู้ข้อมูล

เทคโนโลยีการตรวจจับใบหน้านักนิยมนำไปพัฒนาต่อเป็นเทคโนโลยีตรวจจับใบหน้า และมีการประยุกต์ใช้ในด้านต่างๆ เช่น ด้านความปลอดภัย เพื่อติดตามบุคคล, ด้านการตลาด โดยการวิเคราะห์ข้อมูลประชากร เป็นต้น

การใช้เทคโนโลยีการรู้จำใบหน้าจำเป็นต้องระมัดระวังในด้านความเป็นส่วนตัว และความเสี่ยงที่อาจถูกนำไปใช้ในทางที่ผิด ซึ่งเมื่อมีการนำเทคโนโลยีการจดจำใบหน้ามาใช้งาน ควรแจ้งให้ผู้ที่ถูกตรวจจับใบหน้าทราบด้วย

2.3 การตรวจสอบลักษณะอารมณ์ทางใบหน้า (Facial Emotional Expression Detection)

การแสดงออกทางสีหน้า (Facial Expression) [22] คือ รูปแบบการแสดงสีหน้าที่บ่งบอกถึงอารมณ์ในขณะนั้น ซึ่งจะมีการเปลี่ยนแปลงลักษณะของกล้ามเนื้อบนใบหน้า จึงมีการนำเทคโนโลยี

เข้ามาใช้งาน การตรวจสอบลักษณะอารมณ์ทางใบหน้า (Facial Emotional Expression Detection) [23-26] เป็นเทคโนโลยีที่ใช้อัลกอริทึมคอมพิวเตอร์และเทคนิคเรียนรู้ของเครื่อง เพื่อวิเคราะห์รูปหน้า และระบุสถานะอารมณ์ของบุคคล ซึ่งมีเป้าหมายเพื่อให้คอมพิวเตอร์เข้าใจ ดีความ และตอบสนอง ต่ออารมณ์ของมนุษย์ได้ โดย W. Mellouk และ W. Handouzi [27] ได้มีการทบทวนวรรณกรรมใน หัวข้อ “Facial emotion recognition using deep learning: review and insights” ซึ่งจากการ ทบทวนวรรณกรรมจะพบว่า ในการจำแนกอารมณ์นั้น มักจะถูกจำแนกให้เป็น 6 อารมณ์พื้นฐาน คือ 1) โกรธ 2) ขยะแขยง 3) กลัว 4) มีความสุข 5) เศร้า 6) ประหลาดใจ และภายในงานวิจัยมักมีการ ตรวจสอบลักษณะอารมณ์ทางใบหน้าตามอารมณ์ทั้ง 6 อารมณ์ และใบหน้าที่ไม่แสดงอารมณ์ โดยมี เทคโนโลยีที่นิยมใช้งานในการตรวจสอบคือ โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network: CNN) ซึ่งมีอัตราความถูกต้องมากกว่า 90% โดยกระบวนการในการตรวจสอบ ลักษณะอารมณ์ทางใบหน้า แบ่งเป็นขั้นตอนย่อยได้ดังนี้

1) การรวบรวมข้อมูล (Data Collection) : ขั้นตอนนี้จะเกี่ยวข้องกับการรวบรวมชุดข้อมูล ของรูปภาพใบหน้าที่มีป้ายกำกับด้วยอารมณ์ รูปภาพเหล่านี้จะมาจากแหล่งที่มาต่าง ๆ เช่น รูปภาพ หรือคลิปวิดีโอของผู้คนที่แสดงอารมณ์ต่าง ๆ

2) Feature Extraction : คัดแยกจุดสำคัญและลักษณะของใบหน้า เช่น ตำแหน่งของตา คิ้ว จมูก และปาก ซึ่งจะถูกแยกวิเคราะห์จากรูปภาพที่เก็บไว้ โดยจะเป็นฐานข้อมูลเพื่อการเรียนรู้ของ เครื่อง

3) การฝึกโมเดลการเรียนรู้ของเครื่อง (Machine Learning Model Training) : โมเดล เครื่องเรียนรู้ที่ใช้หลักสูตรการเรียนรู้เชิงลึก เช่น เครือข่ายประสาทเทียมเชิงลึก (CNNs) ถูกฝึกใน ลักษณะที่ได้รับจากลักษณะที่แยกวิเคราะห์ด้วยชุดข้อมูลที่มีป้ายกำกับ ขณะฝึก โมเดลจะเรียนรู้ รูปแบบและความสัมพันธ์ระหว่างลักษณะของใบหน้าและอารมณ์ที่เกี่ยวข้อง

4) การรู้จำอารมณ์ (Emotion Recognition) : เมื่อโมเดลได้รับการฝึก จะสามารถวิเคราะห์ ใบหน้า และสามารถทำนายอารมณ์ของบุคคลในรูปภาพนั้น โดยโมเดลจะจัดกลุ่มข้อมูลนำเข้าเป็น หลาย ๆ หมวดหมู่ที่แตกต่างกัน คือ 1) โกรธ 2) ขยะแขยง 3) กลัว 4) มีความสุข 5) อารมณ์เป็นกลาง 6) เศร้า 7) ประหลาดใจ

2.4 โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network: CNN)

โครงข่ายประสาทเทียมคอนโวลูชัน [28-32] เป็นหนึ่งในลักษณะของโครงข่ายประสาท เทียม โดยมักมีการใช้งานในการตรวจจับใบหน้า เนื่องจากสามารถทำงานได้ทั้งสกัดคุณลักษณะพิเศษ (Feature Extraction) และจำแนกประเภท (Classification) ซึ่งถือได้ว่าเป็นจุดเด่นของโครงข่ายนี้ เนื่องจากการทำงานโดยใช้ Convolution layer ซึ่งมีหน้าที่สกัดคุณลักษณะพิเศษของภาพ

ออกมา โดยที่ Convolution layer จะจำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นที่ย่อย ๆ และนำกลุ่มของพื้นที่ย่อย ๆ มาผสมกัน เช่น เส้นขอบของวัตถุต่าง เพื่อให้โมเดลสามารถเรียนรู้ลักษณะของภาพได้อย่างมีประสิทธิภาพและแม่นยำ โดยจะมีการนำ Convolution layer ไปทำงานร่วมกับ layer อื่นๆ เช่น Pooling Layer , Fully-Connected Layer ซึ่งเป็นจุดที่ทำให้โครงข่ายประสาทเทียมคอนโวลูชันแตกต่างจากโครงข่ายประสาทเทียมอื่นๆ

การใช้งาน CNN เหมาะกับการทำ Face Recognition [33] เนื่องจากโครงสร้างของ CNN ที่มีตัวกรองหลายชั้นจะช่วยปรับปรุงคุณภาพในการตรวจจับใบหน้าได้ดี ถึงแม้ว่าจะมีการปรับเปลี่ยนอัตราส่วนของภาพก็ตาม นอกจากนี้ CNN ยังช่วยในการปรับปรุงประสิทธิภาพและประสิทธิภาพของการทำ Face Recognition และช่วยลดจำนวนกลุ่มภาพที่คาดว่าจะเป็นส่วนหนึ่งของใบหน้าลงด้วย

จากงานวิจัยเรื่อง “A Convolutional Neural Network Cascade for Face Detection” [34] พบว่าการใช้ CNN จะช่วยจัดการกับปัญหาของการตรวจจับใบหน้า และสามารถรักษาประสิทธิภาพการทำงานที่สูงได้ โดยการทำงานที่ความละเอียดหลายระดับ สามารถแยกพื้นหลังได้อย่างรวดเร็ว และหาพื้นที่ที่เป็นใบหน้าได้อย่างแม่นยำในขั้นตอนที่มีความละเอียดสูง

ประโยชน์ในการนำ CNN มาใช้ในงาน Face Detection มีด้วยกันหลายข้อ ดังนี้

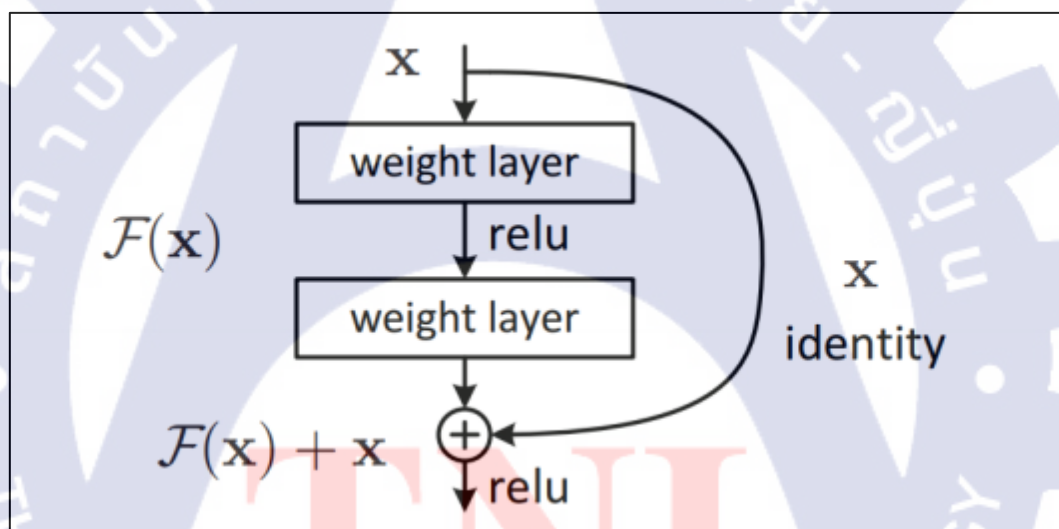
1. CNN สามารถเรียนรู้คุณลักษณะโดยอัตโนมัติเพื่อจับภาพในรูปแบบต่างๆ ที่ซับซ้อนได้ โดยอาศัยข้อมูลที่ได้จากการ training โดยเมื่อข้อมูลในการ training มีจำนวนมาก จะส่งผลให้ CNN สามารถแยกส่วนของใบหน้าออกจากภาพพื้นหลังในสถานการณ์จริงได้แม่นยำมากขึ้น โดยการเรียนรู้ภาพที่มีความหลากหลาย เช่น ท่าทาง การแสดงออก และแสงของภาพ จะช่วยในการปรับปรุงประสิทธิภาพโมเดลได้มากขึ้น
2. เมื่อเปรียบเทียบกับวิธีอื่นๆ แล้ว CNN มีความแม่นยำสูงในการตรวจจับใบหน้า
3. ขั้นตอนการทดสอบของ CNN สามารถทำงานคู่ขนานกันบน GPU ได้ จึงทำให้มีความเร็วเพิ่มขึ้น ซึ่งหมายความว่าโมเดลการตรวจจับใบหน้าที่ใช้ CNN สามารถทำงานได้เร็วและมีประสิทธิภาพมากกว่าโมเดลอื่นๆ
4. โมเดลที่ใช้ CNN สามารถปรับได้สูงและสามารถปรับแต่งสำหรับงานหรือชุดข้อมูลเฉพาะได้ ซึ่งหมายความว่าผู้วิจัยและพัฒนาสามารถปรับแต่งโมเดลเพื่อให้ได้ประสิทธิภาพที่ดีขึ้นในงานของตนเอง

เมื่อก้าวโดยสรุปแล้ว การใช้ CNN สำหรับการตรวจจับใบหน้าได้รับการพิสูจน์แล้วว่าเป็นแนวทางที่มีประสิทธิภาพสูง เนื่องจากความสามารถในการตรวจจับภาพที่มีความซับซ้อนและมีความแม่นยำสูง

2.5 Resnet50

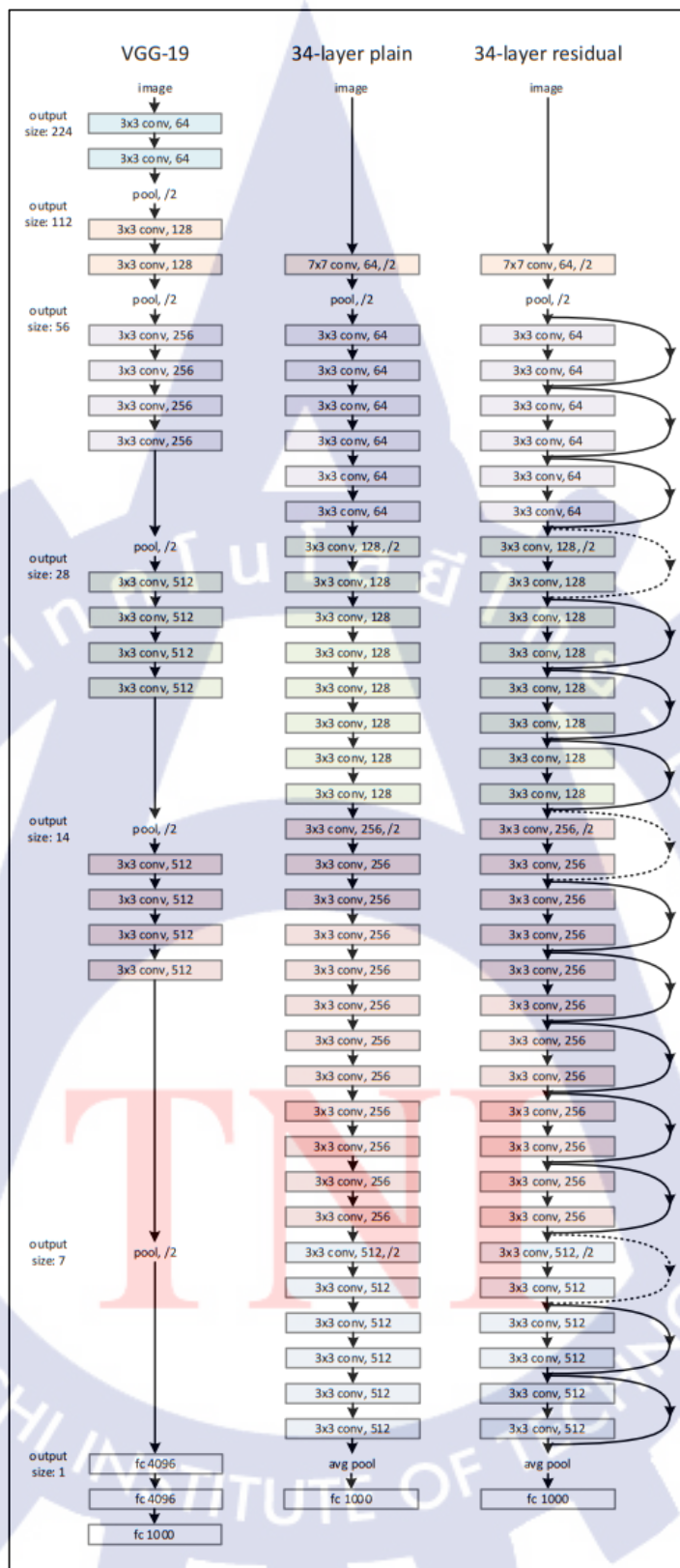
คำว่า "ResNet" มีที่มาจาก "Residual Network" ซึ่งเป็นหมวดหมู่หนึ่งของโครงข่ายประสาทเทียมคอนโวลูชัน (CNNs) ดังที่ได้ถูกนำเสนอในบทความที่มีชื่อ "Deep Residual Learning for Image Recognition" ซึ่งเขียนโดย K. He และคณะ ในปี 2015 ซึ่ง CNNs ได้ถูกนำมาใช้อย่างแพร่หลายในแอปพลิเคชันต่าง ๆ [35]

ResNet ได้ถูกพัฒนาขึ้น เพื่อลดปัญหาในการพัฒนาโมเดลเพื่อพัฒนาโครงข่ายประสาทเทียมที่มีจำนวน layer จำนวนมากในการรู้จำภาพ ซึ่งมักพบปัญหาเกี่ยวกับ vanishing/exploding gradients เมื่อมีจำนวน layer ของโมเดลเพิ่มมากขึ้น จึงแก้ปัญหาโดยใช้ Residual Network ซึ่งมีการทำงานโดยการส่งค่า parameter ข้ามไปยังอีก layer เพื่อไม่ให้ค่าของตัวแปรใน layer ก่อนหน้าหายไป เมื่อผ่านการทำงานของ layer อีกชั้นหนึ่ง โดยการส่งค่า parameter จะส่งข้ามไป 1 layer ดังรูปที่ 2.1



รูปที่ 2.1 แสดงการทำงานของ Residual Network

โมเดล ResNet มีประเภทย่อยหลากหลายประเภท ซึ่งแบ่งตามจำนวนของ layer ของโมเดล โดยโมเดลเริ่มต้นของ ResNet คือ ResNet-34 ซึ่งประกอบไปด้วย layer ต่างๆ จำนวน 34 layer ซึ่งประกอบไปด้วย 7 ส่วน คือ Input Layer, Initial Convolutional Layer, Residual Blocks, Intermediate Layers, Global Average Pooling Layer, Fully Connected Layer (Dense Layer), Output Layer ซึ่งมีการเปรียบเทียบกับโมเดล Convolution Neural Network ดังรูปที่ 2.2



รูปที่ 2.2 แสดงโครงสร้างของ ResNet-34

ResNet-50 เป็นหนึ่งในประเภทของ ResNet ซึ่งเป็นโครงข่ายประสาทเชิงคอนโวลูชัน โดยที่ Resnet50 จะประกอบด้วย 50 ชั้น ซึ่งประกอบด้วยชั้นคอนโวลูชัน 48 ชั้น MaxPool layer 1 ชั้น และ average pool layer 1 ชั้น Residual neural networks เป็นส่วนหนึ่งของโครงข่ายประสาทเทียม (ANNs) ซึ่งมีลักษณะในการสร้างโดยการนำบล็อกที่เชื่อมต่อต่อกันโดยใช้ residual blocks

ในงานวิจัยนี้ ResNet-50 เป็นตัวเลือกเหมาะสม เนื่องจากจำนวนข้อมูลมีไม่มากและน้อยเกินไป และทำงานได้ดีสำหรับวัตถุประสงค์ในการวิจัย เมื่อพิจารณาโดยขนาดของข้อมูล

ResNet-50 มีความเหมาะสม เนื่องจากมีความสมดุลระหว่างประสิทธิภาพและการใช้ทรัพยากรในการประมวลผลข้อมูล โมเดลขนาดใหญ่ เช่น ResNet-101 และ ResNet-152 ต้องการพื้นที่จัดเก็บและพลังการประมวลผลที่มากขึ้น ซึ่งเกินความต้องการของงานวิจัยนี้

2.6 MTCNN

MTCNN หรือ Multi-task Cascaded Convolutional Networks เป็นเทคโนโลยีในการตรวจจับใบหน้า ซึ่งถูกนำเสนอโดย K. Zhang et al. [36] ในงานวิจัยที่ชื่อว่า Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks ในปี 2016 ซึ่งได้รับความนิยมเป็นอย่างมาก เพราะมีความแม่นยำสูง และใช้เวลาในการประมวลผลน้อย

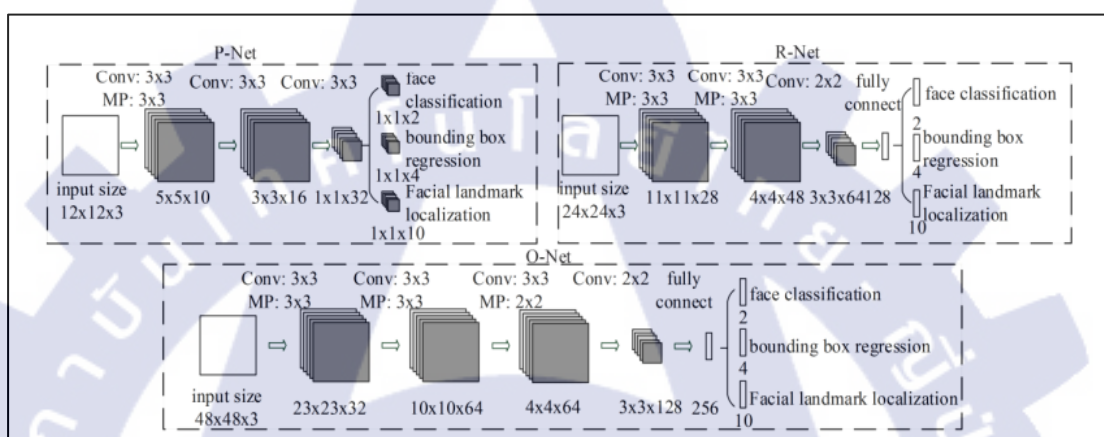
การทำงานของ MTCNN คือการทำ Face Detection และ Face Alignment แบบ Multi-Tasking โดยใช้ Neural Network เพียงหนึ่งตัว เพื่อทำงานทั้ง 2 อย่างไปพร้อมกัน ซึ่งมีความแตกต่างจากวิธีอื่นๆ ที่สามารถทำงานได้เพียงอย่างใดอย่างหนึ่ง โดยที่ประโยชน์ของการทำทั้ง 2 Tasks ใน Network เดียวกัน จะช่วยลดเวลา เนื่องจากทำงานทั้งสองอย่างพร้อมกันได้ และยังไม่ได้มีแค่ทำให้เราสามารถทำทั้งสองอย่างได้พร้อมกัน แต่ยังช่วยเพิ่มความแม่นยำให้การทำงานทั้งสองกระบวนการอีกด้วย

Face Alignment คือ การจัดภาพใบหน้าที่จับภาพได้ ให้อยู่ในทิศทางที่ถูกต้อง ซึ่งวิธีที่เป็นที่นิยมในการทำ Face Alignment นั้นคือการหา Keypoint บนใบหน้า เช่น ตาซ้าย ตาขวา จมูก ปาก และตัดภาพส่วนที่ไม่เกี่ยวข้องออกไป และจัดภาพให้จมูกอยู่ตรงกลางภาพ จากนั้นก็ปรับปรุงภาพตามขั้นตอนต่างๆ เช่น การหมุนภาพเพื่อให้ตาซ้ายและขวาอยู่ในระดับเดียวกัน เป็นต้น

การทำงานของ MTCNN นั้นแตกต่างกับกระบวนการ face detection และ Face Alignment โดยทั่วไป ที่มักจะแยกการทำงานของ face detection และ face alignment ออกจากกัน ซึ่งภายในงานวิจัยของ K. Zhang และคณะ ได้มีการปรับการทำงานให้การทำงานของ face detection และ face alignment ทำงานพร้อมกันภายในโมเดลเดียวกัน เนื่องจากมีการทำงานที่คล้ายคลึงกัน และส่งผลให้โครงข่ายประสาทเทียมสามารถเรียนรู้ได้ดีขึ้น เนื่องจากการทำงานของโครงข่ายประสาทเทียมคอนโวลูชันที่มีการทำงานต่อเนื่องกันระหว่าง layer ซึ่งภายใน layer ที่เป็น

ลำดับต้นๆ ของโครงข่ายประสาทเทียมคอนโวลูชัน มักจะเป็นส่วนที่เรียนรู้ Low-Level Features เช่น การหาขอบของวัตถุในภาพ ซึ่ง Features เหล่านี้หากต้องการผลลัพธ์ที่คล้ายคลึงกันแล้ว มักจะมีการทำงานที่คล้ายกันด้วยเช่นกัน ซึ่งสามารถช่วยในการเรียนรู้ของโมเดลให้เรียนรู้ทั้งจาก 2 งาน และเป็นการเพิ่มข้อจำกัดเพื่อให้เรียนรู้เฉพาะเนื้อหาที่สำคัญของงานที่ต้องการตรวจสอบ

MTCNN มีการแบ่งขั้นตอนการทำงานออกเป็น 3 ขั้นตอน ซึ่งประกอบไปด้วย P-Net, R-Net และ O-Net ดังรูปที่ 2.3 โดยที่แต่ละโมเดลจะทำงานต่อกัน โดยที่ผลลัพธ์จากโมเดลก่อนหน้า จะเป็นข้อมูลนำเข้าของโมเดลถัดไป



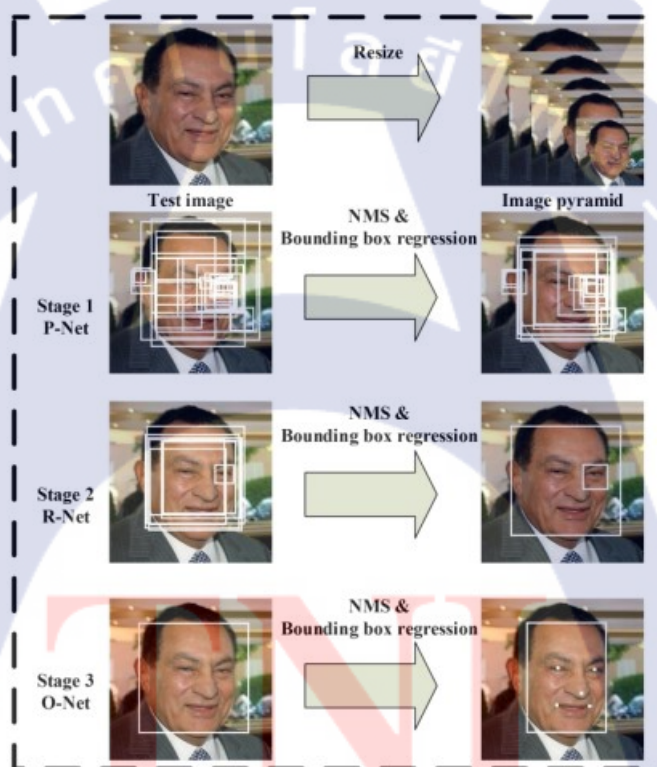
รูปที่ 2.3 แสดงสถาปัตยกรรมของ MTCNN

Proposal-Network (P-Net) เป็นโมเดลที่หนึ่งในการทำงานของ MTCNN ซึ่งมีความซับซ้อนน้อยที่สุดจากทั้ง 3 โมเดล โดยมีหน้าที่คือการหา Bounding Box ของภาพใบหน้า โดยมีขั้นตอนการทำงาน คือ เมื่อรับภาพเข้ามาแล้ว จะทำการย่อขนาดรูปลงตามลำดับทั้งหมด 12 ครั้ง เนื่องจาก P-Net มีจำนวน Filters ของ CNN น้อยมาก และเนื่องจาก CNN ไม่สามารถจับภาพเดียวกันที่มีขนาดแตกต่างกันได้ จึงลดขนาดของภาพ และเพิ่มจำนวนการทำงานของโมเดล และเมื่อโมเดลตรวจพบ Bounding Box ของภาพใบหน้าในขนาดภาพใดแล้ว จึงจะคำนวณย้อนกลับมาเป็นตำแหน่งในภาพขนาดจริง ซึ่ง Output ของ P-Net จะประกอบไปด้วย Bounding Box จำนวนมาก โดยจะเลือกเฉพาะ Bounding Box ที่มีจำนวนใบหน้ามากกว่าค่า Threshold ที่กำหนดไว้ และการ Non Maximum Suppression (NMS) ซึ่งจะทำให้เราได้เฉพาะ Bounding Box ที่ไม่ซ้ำกัน

Refine-Network (R-Net) เป็นโมเดล CNN ที่มีความซับซ้อนเพิ่มมากกว่า P-Net โดยจะรับข้อมูลภาพที่อยู่ใน Bounding Box ซึ่งได้จาก P-Net ทีละภาพ จากนั้นมันจึงหา Bounding Box ของใบหน้าในภาพเหล่านั้นอีกครั้ง เพื่อเพิ่มความแม่นยำในการหา Bounding Box และคัดเลือก

Bounding Box ที่มีจำนวนใบหน้ามากกว่าค่า Threshold ที่กำหนดไว้ และทำ NMS เพื่อตัด Bounding Box ที่ซ้ำกัน เช่นเดียวกับภายใน P-Net แต่ความแตกต่างของชั้น R-Net คือ จะมีการกำหนดค่า Threshold ให้สูงกว่าชั้น P-Net

Output Net (O-Net) เป็น Neural Network ที่มีความซับซ้อนมากที่สุด โดย O-Net จะใช้ผลลัพธ์ของ R-Net เพื่อหา Bounding Box และความน่าจะเป็นของใบหน้าอีกครั้ง โดยมีขั้นตอนคล้ายคลึงกับการทำงานภายใน R-Net แต่ O-Net จะทำการหา Face Alignment ของแต่ละหน้าที่พบด้วย ซึ่ง Face Alignment นั้นประกอบไปด้วยตำแหน่งจำนวน 5 ตำแหน่ง ได้แก่ ตาสองข้าง, จมูก, และขอบซ้ายขวาของริมฝีปาก

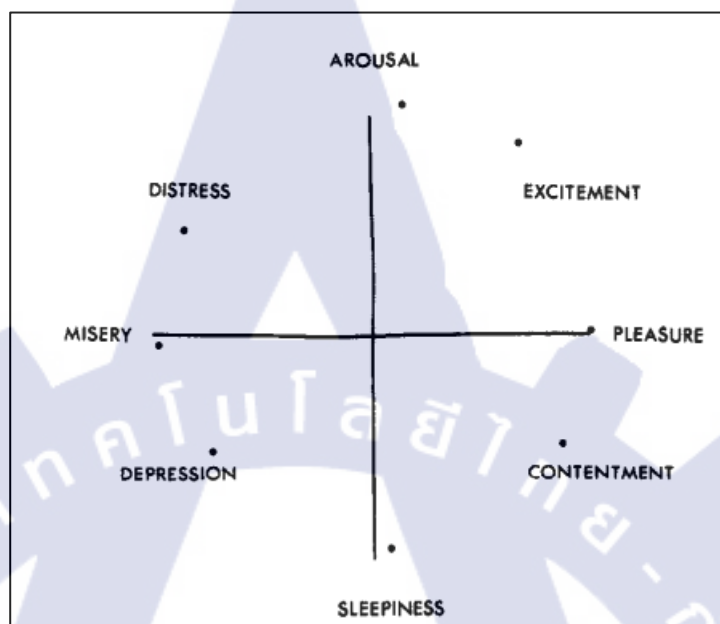


รูปที่ 2.4 แสดงการหากรอบพื้นที่ของใบหน้าภายในขั้นตอน P-Net , R-Net และ O-Net

2.7 การจำแนกลักษณะอารมณ์

J. A. Russell [37] ได้ทำงานวิจัยเรื่อง “A circumplex model of affect” ซึ่งมีการสร้างโมเดลเกี่ยวกับลักษณะทางอารมณ์ โดยจะมีการสร้างโมเดลเป็นวงกลมขึ้นมา โดยลักษณะอารมณ์หลักๆ จะแบ่งเป็นสี่ช่วง คือ พึงพอใจ (pleasure) ที่มุม 0 องศา กระตือรือร้น (Arousal) ที่ 90 องศา

ไม่พอใจ (misery) ที่ 180 องศา และง่วงนอน (Sleepiness) ที่ 270 องศา โดยจะมีการจำแนกอารมณ์ต่างๆ อยู่ในโมเดล ดังรูปที่ 2.5



รูปที่ 2.5 แสดงโมเดลของ Russell

2.8 การขยับของกล้ามเนื้อใบหน้า

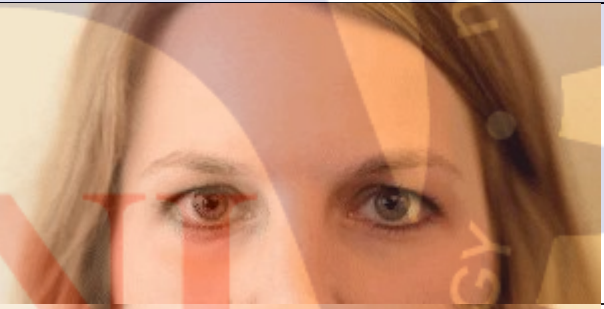
บนใบหน้าของมนุษย์จะมีกล้ามเนื้อประมาณ 20 มัด โดยกล้ามเนื้อบนใบหน้าจะมีหน้าที่หลัก 2 อย่าง คือ 1) ใช้สำหรับเคี้ยวอาหาร และ 2) เพื่อสร้างการแสดงออกบนใบหน้า โดยกล้ามเนื้อแต่ละมัด จะทำหน้าที่ดังนี้ [38]

- Auriculars เพื่อให้บางคนสามารถขยับใบหูได้
- Corrugator supercilii เป็นกล้ามเนื้อที่อยู่บริเวณคิ้ว ทำให้สามารถขมวดคิ้วได้
- Depressor anguli oris อยู่บริเวณคางทั้งสองข้าง มีส่วนช่วยในการขมวดคิ้ว
- Depressor labii inferioris อยู่บริเวณคาง มีส่วนช่วยในการขยับริมฝีปากล่าง
- Levator labii superioris alaeque nasi ช่วยในการขยับรูจมูกและริมฝีปากบน
- Mentalis กล้ามเนื้อบริเวณตรงกลางคาง ช่วยในการควบคุมริมฝีปากล่าง
- Nasalis ช่วยในการเปิดรูจมูก
- Occipitofrontalis กล้ามเนื้อที่อยู่บริเวณเหนือคิ้ว ไปจนถึงกะโหลก ทำหน้าที่ในการยกคิ้วขึ้นและทำให้ริมฝีปากยกขึ้นได้
- Orbicularis oculi ทำหน้าที่ในการควบคุมเปลือกตา

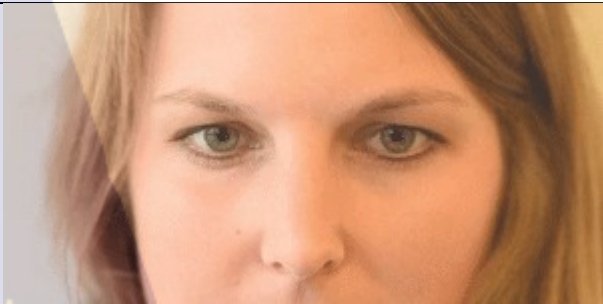
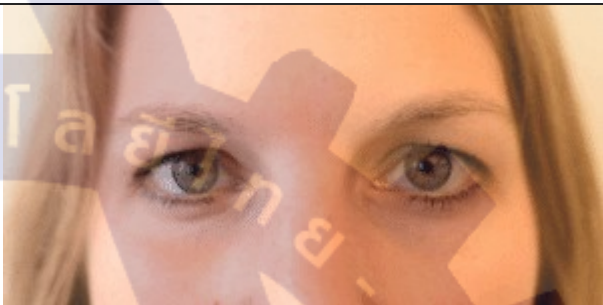
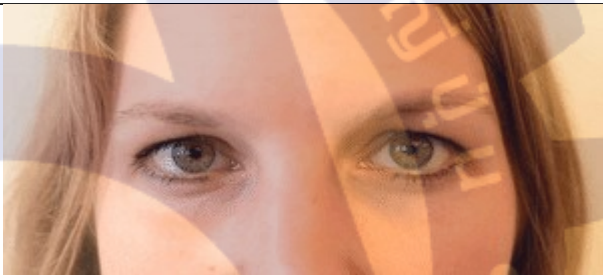
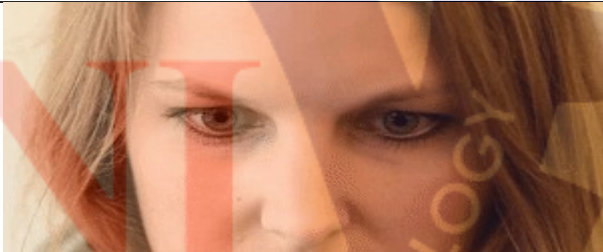
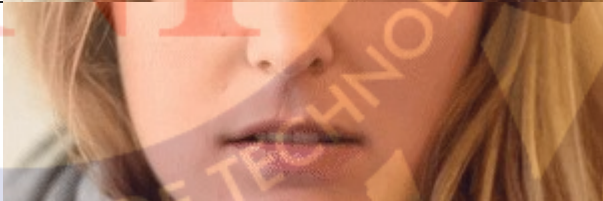
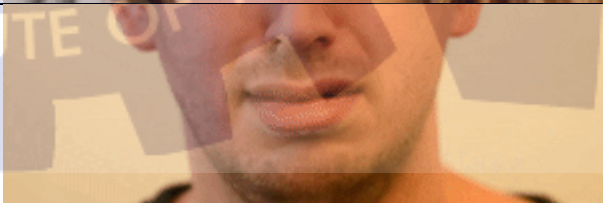
- Orbicularis oris กล้ามเนื้อรอบริมฝีปาก เพื่อควบคุมการเปิดปิดของริมฝีปาก
- Procerus กล้ามเนื้อระหว่างคิ้ว ทำงานในการควบคุมคิ้วและรูดมูก
- Risorius อยู่บริเวณด้านข้างของปากทั้งสองข้าง และมีส่วนช่วยในการยิ้ม
- Zygomaticus major and minor ช่วยในการยิ้ม

นอกจากนี้ B. Farnsworth [39] ได้ทำการรวบรวมเกี่ยวกับการขยับของกล้ามเนื้อใบหน้า โดยมีการนำข้อมูลมาจาก Carl-Herman Hjortsjö ซึ่งมีการรวบรวมเกี่ยวกับการขยับของกล้ามเนื้อบนใบหน้า 23 จุด ในปี 1970 โดยมีการรวบรวมข้อมูลเพิ่มเติม เป็น 46 จุด โดยแสดงดังตารางที่ 2.1

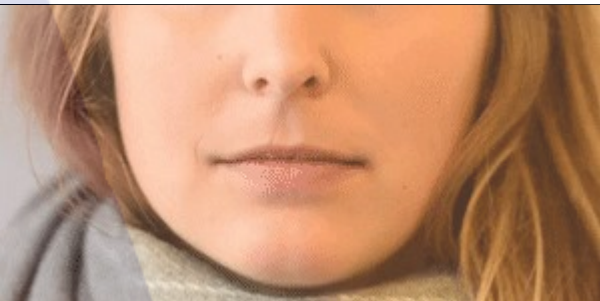
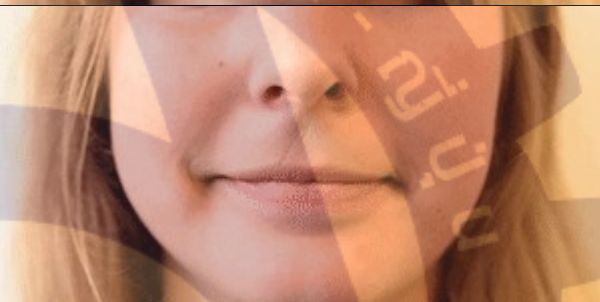
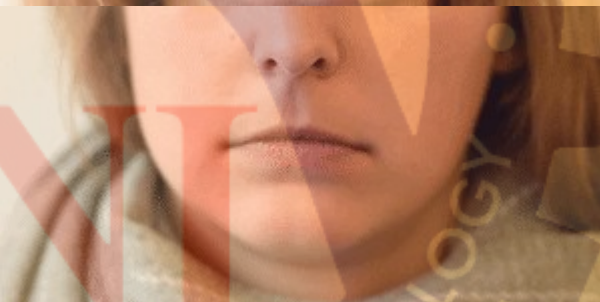
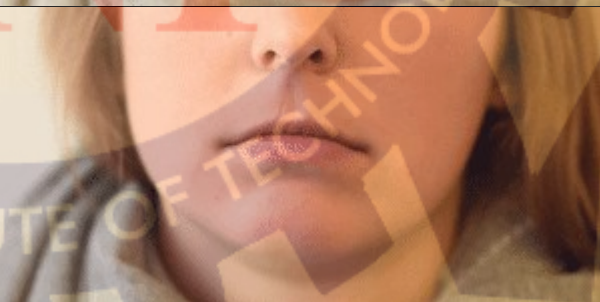
ตารางที่ 2.1 แสดงการขยับของกล้ามเนื้อใบหน้า

คำอธิบาย	กล้ามเนื้อที่ขยับ	ภาพตัวอย่าง
Inner Brow Raiser	<i>Frontalis, pars medialis</i>	
Outer Brow Raiser (unilateral, right side)	<i>Frontalis, pars lateralis</i>	
Brow Lowerer	<i>Depressor Glabellae, Depressor Supercilli, Currugator</i>	

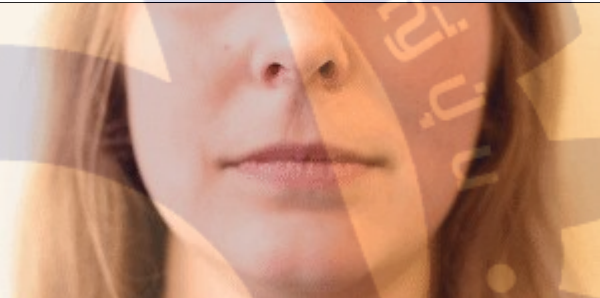
ตารางที่ 2.1 แสดงการขยับของกล้ามเนื้อใบหน้า (ต่อ)

คำอธิบาย	กล้ามเนื้อที่ขยับ	ภาพตัวอย่าง
Upper Lid Raiser	<i>Levator palpebrae superioris</i>	
Cheek Raiser	<i>Orbicularis oculi, pars orbitalis</i>	
Lid Tightener	<i>Orbicularis oculi, pars palpebralis</i>	
Nose Wrinkler	<i>Levator labii superioris alaeque nasi</i>	
Upper Lip Raiser	<i>Levator Labii Superioris, Caput infraorbitalis</i>	
Nasolabial Deepener	<i>Zygomatic Minor</i>	

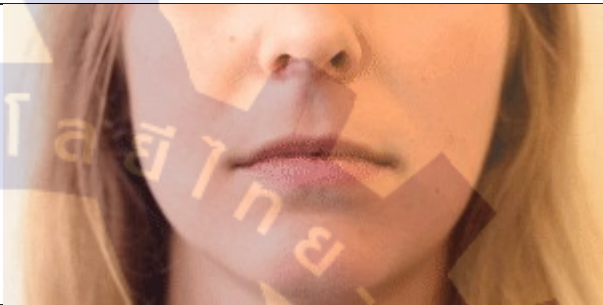
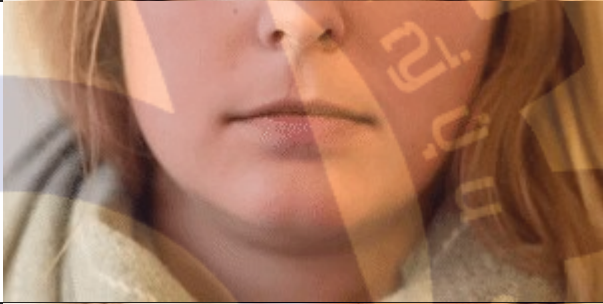
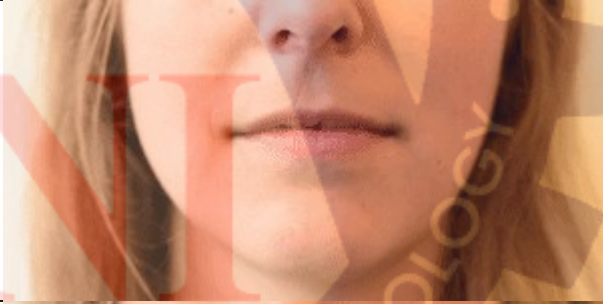
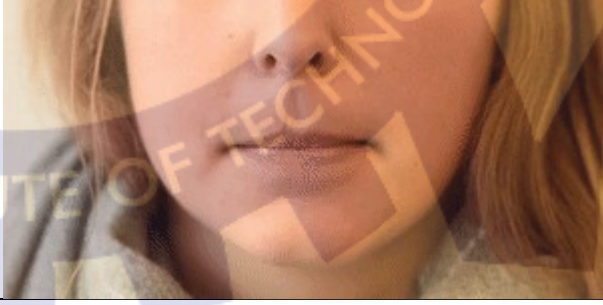
ตารางที่ 2.1 แสดงการขยับของกล้ามเนื้อใบหน้า (ต่อ)

คำอธิบาย	กล้ามเนื้อที่ขยับ	ภาพตัวอย่าง
Lip Corner Puller	<i>Zygomatic Major</i>	
Cheek Puffer	<i>Levator anguli oris (Caninus)</i>	
Dimpler	<i>Buccinator</i>	
Lip Corner Depressor	<i>Depressor anguli oris (Triangularis)</i>	
Lower Lip Depressor	<i>Depressor labii inferioris</i>	

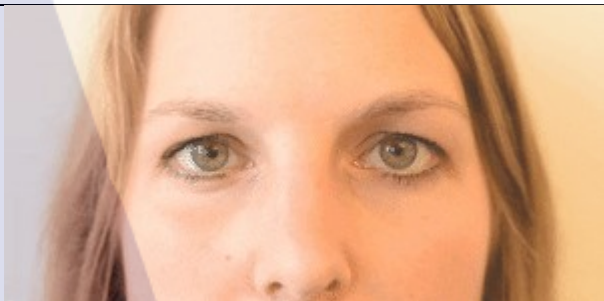
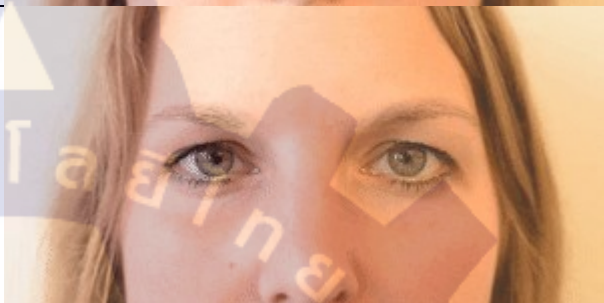
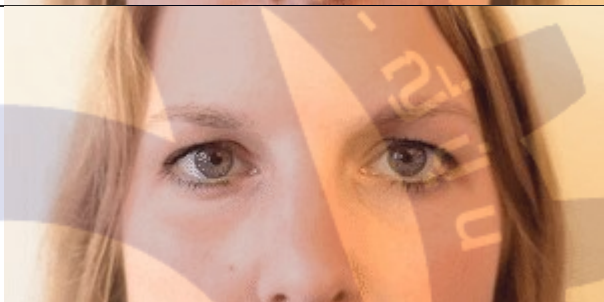
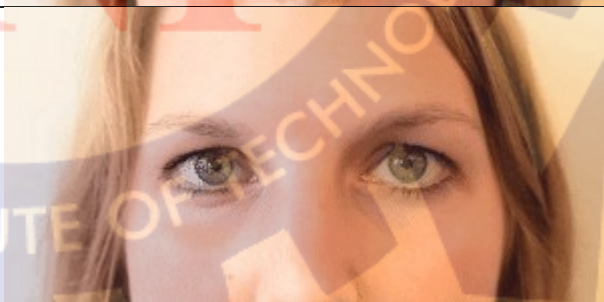
ตารางที่ 2.1 แสดงการขยับของกล้ามเนื้อใบหน้า (ต่อ)

คำอธิบาย	กล้ามเนื้อที่ขยับ	ภาพตัวอย่าง
Chin Raiser	<i>Mentalis</i>	
Lip Pucker	<i>Incisivii labii superioris and Incisivii labii inferioris</i>	
Lip stretcher	<i>Risorius</i>	
Lip Funneler	<i>Orbicularis oris</i>	
Lip Tightener	<i>Orbicularis oris</i>	

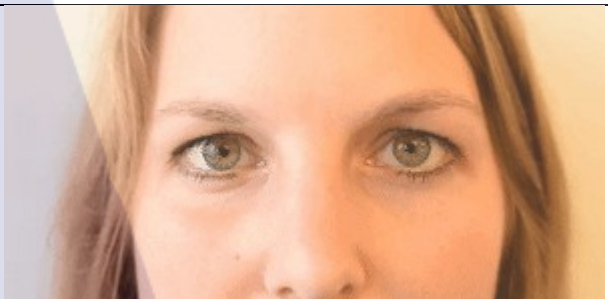
ตารางที่ 2.1 แสดงการขยับของกล้ามเนื้อใบหน้า (ต่อ)

คำอธิบาย	กล้ามเนื้อที่ขยับ	ภาพตัวอย่าง
Lip Pressor	<i>Orbicularis oris</i>	
Lips part	<i>Depressor Labii, Relaxation of Mentalis (AU17), Orbicularis Oris</i>	
Jaw Drop	<i>Maseter; Temporal and Internal Pterygoid relaxed</i>	
Mouth Stretch	<i>Pterygoids, Digastric</i>	
Lip Suck	<i>Orbicularis oris</i>	

ตารางที่ 2.1 แสดงการขยับของกล้ามเนื้อใบหน้า (ต่อ)

คำอธิบาย	กล้ามเนื้อที่ขยับ	ภาพตัวอย่าง
Lid droop	Relaxation of Levator <i>Palpebrae Superioris</i>	
Slit	<i>Orbicularis oculi</i>	
Eyes Closed	Relaxation of Levator <i>Palpebrae Superioris</i>	
Squint	<i>Orbicularis oculi,</i> <i>pars palpebralis</i>	
Blink	Relaxation of Levator <i>Palpebrae</i> and Contraction of <i>Orbicularis Oculi,</i> <i>Pars Palpebralis.</i>	

ตารางที่ 2.1 แสดงการขยับของกล้ามเนื้อใบหน้า (ต่อ)

คำอธิบาย	กล้ามเนื้อที่ขยับ	ภาพตัวอย่าง
Wink	<i>Levator palpebrae superioris; Orbicularis oculi, pars palpebralis</i>	

2.9 การเปลี่ยนแปลงใบหน้าตามลักษณะอารมณ์

การแสดงออกทางสีหน้าเป็นส่วนหนึ่งในการสื่อสารของมนุษย์ [40-41] ซึ่งเป็นวิธีการสื่อสารโดยไม่ใช้คำพูดที่สำคัญที่สุดอย่างหนึ่ง โดยอารมณ์ที่ต่างกัน จะมีการแสดงออกทางใบหน้าที่แตกต่างกัน ซึ่งสามารถสังเกตและแยกแยะได้โดยง่าย และมีบทบาทในการสื่อสารด้านจิตวิทยา ซึ่งการเปลี่ยนแปลงของใบหน้าจะให้ข้อมูลเกี่ยวกับอารมณ์, การแสดงความตั้งใจและเป้าหมาย , และการปฏิสัมพันธ์ระหว่างมนุษย์ การรู้จักและเข้าใจการแสดงสีหน้า จะมีส่วนช่วยเหลือในกระบวนการสื่อสาร โดยกระบวนการในการแยกประเภทการแสดงสีหน้าประกอบไปด้วย 3 ขั้นตอนหลัก คือ Face Detection , Feature Extraction , และ Classification โดยภายในการวิจัยนี้ แบ่งการแสดงอารมณ์ของมนุษย์ออกเป็น 7 อารมณ์หลัก ได้แก่

1) อารมณ์โกรธ (Angry) : คิ้วตกลง เปลือกตาล่างเปิดกว้าง ขอบริมฝีปากม้วนเข้า ริมฝีปากเหี้ยมดัด



รูปที่ 2.6 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์โกรธ

2) อารมณ์กลัว (Fear) : คิ้วขมวดเข้าหากัน เปลือกตาบนขยับขึ้น เปลือกตาล่างเลื่อนต่ำลง ปากอ้าออกและเหยียดตรง



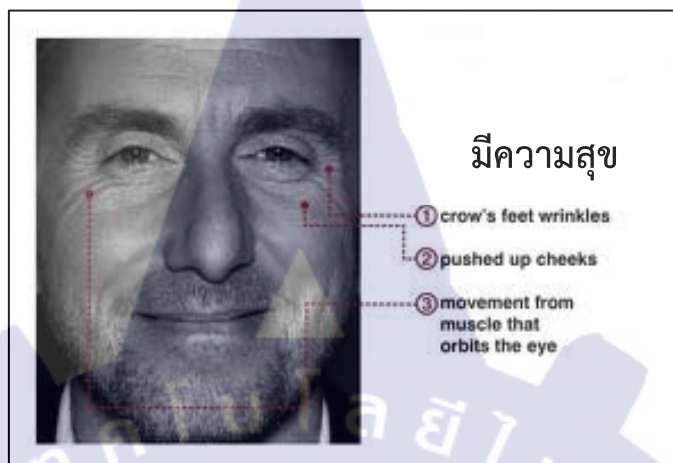
รูปที่ 2.7 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์กลัว

3) อารมณ์ขยะแขยง (Disgust) คิ้วตกลง จมูกย่น ริมฝีปากบนขยับขึ้น ปากอ้าออก



รูปที่ 2.8 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์ขยะแขยง

4) มีความสุข (Happy) : กล้ามเนื้อรอบดวงตากระชับขึ้น เห็นริ้วรอยรอบดวงตา แก้มยกขึ้น มุมริมฝีปากยกขึ้นตามแนวทแยงมุม



รูปที่ 2.9 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์มีความสุข

5) อารมณ์เป็นกลาง (Neutral) : ริมฝีปากเหยียดตรง และมุมปากยกขึ้นเพียงข้างเดียว



รูปที่ 2.10 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์เป็นกลาง

6) ความเศร้า (Sad) : เปลือกตาบนลดต่ำลง ดวงตาไม่โฟกัส มุมริมฝีปากลดต่ำลง



รูปที่ 2.11 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์เศร้า

7) ประหลาดใจ (Surprise) : คิ้วยกขึ้น ตาเบิกกว้าง ริมฝีปากอ้าออก



รูปที่ 2.12 แสดงการเปลี่ยนแปลงของใบหน้าที่มีอารมณ์เศร้า

2.10 FER-2013

FER-2013 หรือ Facial Expression Recognition 2013 [42-43] คือ ชุดข้อมูลที่รวบรวมข้อมูลภาพใบหน้าของมนุษย์ โดยมีการรวบรวมข้อมูลบนเว็บไซต์ Kaggle และเปิดให้เข้าใช้งานได้ฟรี

โดยมีการนำเสนอชุดข้อมูลนี้ที่ International Conference on Machine Learning (ICML) ในปี 2013 โดย Pierre-Luc Carrier และ Aaron Courvil

ภายในชุดข้อมูลนี้ รูปใบหน้าแต่ละรูปได้ถูกจัดกลุ่มตามอารมณ์ต่างๆ 7 กลุ่มอารมณ์ โดยภาพภายในชุดข้อมูล FER-2013 จะเป็นภาพ grayscale และมีขนาด 48 x 48 และมีรูปจำนวน 35,887 รูป โดยมีจำนวนรูปดังตารางที่ 2.2

ตารางที่ 2.2 แสดงจำนวนข้อมูลภาพในชุดข้อมูล FER-2013

ลักษณะอารมณ์	ข้อมูลสำหรับเทรน โมเดล (รูป)	ข้อมูลสำหรับ ทดสอบโมเดล (รูป)	จำนวนใบหน้า ทั้งหมด (รูป)
อารมณ์โกรธ (Angry)	3,995	958	4,953
อารมณ์ขยะแขยง (Disgust)	436	111	547
อารมณ์กลัว (Fear)	4,097	1,024	5,121
มีความสุข (Happy)	7,215	1,774	8,989
อารมณ์เป็นกลาง (Neutral)	4,965	1,233	6,198
เศร้า (Sad)	4,830	1,247	6,077
ประหลาดใจ (Surprise)	3,171	831	4,002
รวม	28,709	7,178	35,887



รูปที่ 2.13 แสดงภาพตัวอย่างของ dataset [42-43]

2.11 การวัดประสิทธิภาพของโมเดล

การวัดประสิทธิภาพของโมเดล [44] สามารถวัดได้หลายวิธี โดยวิธีที่นิยมใช้งานจะมี 1) ค่าความถูกต้อง (Accuracy) 2) ค่าความแม่นยำ (Precision) 3) ค่าความระลึก (Recall) 4) ค่าประสิทธิภาพโดยรวม (F-measure) และ 5) Confusion Matrix โดยในการวัดค่าข้อมูลเหล่านี้ จะมีการแบ่งข้อมูลที่ผ่านการทดสอบจากโมเดลออกเป็น 4 กลุ่ม คือ

- ค่าความถูกต้องเชิงบวก (True Positive : TP) คือ ข้อมูลที่อยู่ในกลุ่ม A และแบบจำลองทำนายว่าอยู่ในกลุ่ม A
- ค่าความผิดพลาดเชิงบวก (False Positive : FP) คือ ข้อมูลที่ไม่อยู่ในกลุ่ม A แต่แบบจำลองทำนายว่าอยู่ในกลุ่ม A
- ค่าความผิดพลาดเชิงลบ (False Negative : FN) คือ ข้อมูลที่อยู่ในกลุ่ม A แต่แบบจำลองทำนายว่าไม่อยู่ในกลุ่ม A
- ค่าความถูกต้องเชิงลบ (True Negative : TN) คือ ข้อมูลที่ไม่อยู่ในกลุ่ม A และแบบจำลองทำนายว่าไม่อยู่ในกลุ่ม A

ค่าความถูกต้อง (Accuracy) เป็นการวัดประสิทธิภาพของอัลกอริทึมการจำแนกประเภท โดยพิจารณาจากอัตราส่วนของจำนวนข้อมูลที่จำแนกได้ถูกต้อง กับข้อมูลทั้งหมดที่ได้มีการกำกับค่าไว้

แล้ว ค่าความถูกต้องจะแสดงให้เห็นถึงความน่าจะเป็นที่อัลกอริทึมนั้นจะทำนายถูก ดังนั้นยิ่งค่าความถูกต้องสูง แสดงว่ามีการจำแนกที่ดีกว่า ดังสมการที่ 2.1

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.1)$$

ค่าความแม่นยำ (Precision) คือ ค่าความสามารถในการจัดข้อมูลที่ไม่เกี่ยวข้องออกไปเป็นอัตราส่วนของการค้นพบข้อมูลที่ถูกต้องจากจำนวนข้อมูลทั้งหมดที่ทำการค้นคืนมาได้ ถ้าค่าความแม่นยำเท่ากับ 1 หมายถึง ระบบสามารถค้นคืนข้อมูลได้ถูกต้อง โดยที่ไม่มีข้อมูลที่ไม่เกี่ยวข้องปะปนอยู่ โดยมีวิธีการคำนวณดังสมการที่ 2.2

$$Precision = \frac{TP}{TP + FP} \quad (2.2)$$

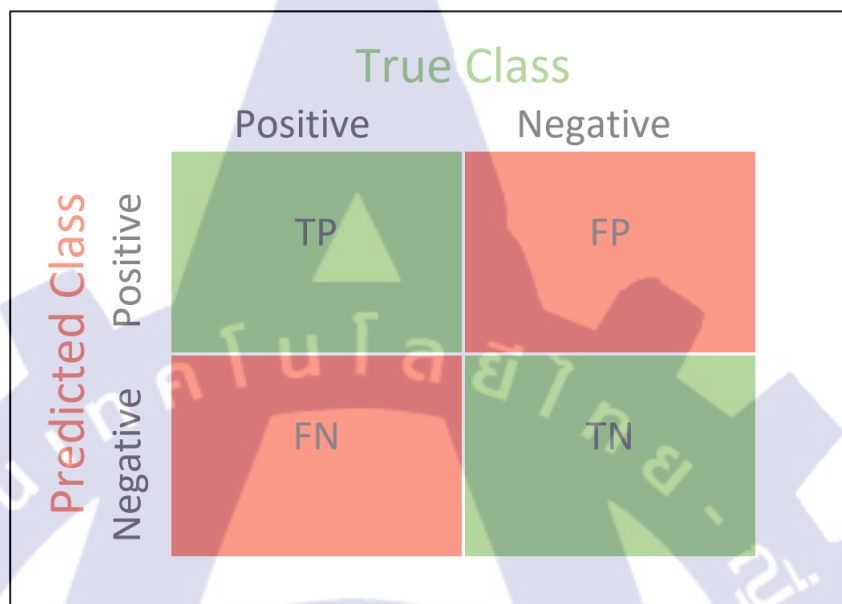
ค่าความระลึก (Recall) คือ ค่าที่วัดความถูกต้องของโมเดลเมื่อเทียบกับค่าความเป็นจริง ซึ่งจะสนใจค่าข้อมูลของกลุ่ม False Negative เป็นหลัก คำนวณได้โดยใช้สมการที่ 2.3

$$Recall = \frac{TP}{TP + FN} \quad (2.3)$$

ค่าประสิทธิภาพโดยรวม (F1-score) ค่าที่แสดงประสิทธิภาพของโมเดล เป็นค่าที่ได้จากการนำค่า precision และ recall มาคำนวณรวมกัน หากมีค่ามากถือว่ามีประสิทธิภาพที่ดี คำนวณได้จากสมการที่ 2.4

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2.4)$$

Confusion Matrix เป็นเครื่องมือที่มีความสำคัญ และนิยมใช้งานในการประเมินผลลัพธ์ของการทำนาย โดยจะอยู่ในรูปของ Matrix ที่เป็นแนวนอน (Row) และแนวตั้ง (Column) ซึ่งจะมีลักษณะข้อมูลเป็น 2 มิติ ดังรูปที่ 2.14



รูปที่ 2.14 แสดง Confusion Matrix [45]

2.12 เทคโนโลยีที่เกี่ยวข้อง

2.12.1 Visual Studio Code

Visual Studio Code [46] หรือ VS Code จากบริษัทไมโครซอฟต์ เป็นโปรแกรมประเภท Editor ใช้ในการแก้ไขโค้ดที่มีขนาดเล็ก แต่มีประสิทธิภาพสูง เป็น OpenSource โปรแกรมจึงสามารถนำมาใช้งานได้โดยไม่มีค่าใช้จ่าย เหมาะสำหรับนักพัฒนาโปรแกรมที่ต้องการใช้งานหลายแพลตฟอร์ม รองรับการใช้งานทั้งบน Windows , macOS และ Linux รองรับหลายภาษาทั้ง JavaScript, TypeScript และ Node.js ในตัว และสามารถเชื่อมต่อกับ Git ได้ง่าย สามารถนำมาใช้งานได้ง่ายไม่ซับซ้อน มีเครื่องมือและส่วนขยายต่าง ๆ ให้เลือกใช่มากมาย รองรับการใช้งานภาษาอื่น ๆ ทั้ง ภาษา C++ , C# , Java , Python , PHP หรือ Go สามารถปรับเปลี่ยน Themes ได้ มีส่วน Debugger และ Commands เป็นต้น

2.12.2 Google API

Google API หรือ Google Application Programming Interface [47] เป็นชุดคำสั่งที่ Google เผยแพร่ให้นักพัฒนาซอฟต์แวร์ใช้เพื่อเข้าถึงและเรียกใช้บริการต่างๆ ของ Google ผ่านโปรแกรมของตนเอง โดย API ของ Google มีบริการหลากหลายประเภท ครอบคลุมการใช้งานในหลากหลายด้าน เช่น Google Search API, Google Maps API เพื่อใช้ในการค้นหา , Google Cloud Platform API, Google Cloud Storage API เพื่อเก็บข้อมูลบน Cloud , การปรับแต่งซอฟต์แวร์สำนักงาน ได้แก่ Google Sheets API, Google Docs API หรือ Google Drive API เพื่อใช้งานในการ upload-download ข้อมูลจาก Google Drive

การให้บริการ Google API นั้น นักพัฒนาจะต้องลงทะเบียนใช้งานและสร้างบัญชี Google Cloud Platform เพื่อรับคีย์ API จากนั้นจึงใช้คีย์ API ในการเรียกใช้ API ผ่านโปรแกรมของตนเอง

API ของ Google ช่วยให้นักพัฒนาสามารถเข้าถึงและเรียกใช้บริการต่างๆ ของ Google ได้อย่างสะดวกและรวดเร็ว ซึ่งช่วยลดเวลาและค่าใช้จ่ายในการพัฒนาแอปพลิเคชันหรือบริการใหม่ๆ ได้เป็นอย่างดี

2.12.3 Python

Python [48-51] เป็นภาษาที่ใช้ในการเขียนโปรแกรม ที่ได้รับความนิยมอย่างแพร่หลายในการทำงานด้าน Machine Learning , วิทยาศาสตร์ข้อมูล และการพัฒนาซอฟต์แวร์ และนอกจากนี้ยังสามารถพัฒนาแอปพลิเคชันบนคอมพิวเตอร์ได้อีกด้วย สามารถนำไปใช้งานได้หลายรูปแบบ และเป็นภาษาที่เรียนรู้และทำความเข้าใจได้ง่าย เนื่องจากเป็นภาษาโปรแกรมที่มีความคล้ายคลึงกับภาษาอังกฤษ จึงทำให้ผู้ที่เริ่มต้นศึกษาการเขียนโปรแกรม นิยมเริ่มต้นศึกษาจากภาษา Python

Python มีการทำงานเป็นแบบ Interpreter คือการแปลคำสั่งการทำงานทีละบรรทัด เพื่อให้คอมพิวเตอร์ทำงานได้ และนิยมใช้งานในด้าน Machine Learning เนื่องจากมี library ที่เกี่ยวข้องกับงาน Machine Learning ให้ใช้งานจำนวนมาก เช่น tensorflow , numpy , scikit learn

2.12.4 Simple Mail Transfer Protocol : SMTP

Simple Mail Transfer Protocol (SMTP) [52-53] คือ มาตรฐานที่ใช้ในการส่ง e-mail ในเครือข่ายอินเทอร์เน็ต เพื่อไปยังเครื่องอื่นๆ ซึ่งสามารถส่งอีเมลไปหาผู้ใช้งานที่ใดก็ได้ทั่วโลก ซึ่งมีการทำงานโดยทำงานผ่านชุดคำสั่ง ซึ่งจะใช้ส่ง email ระหว่าง mail server โดยใช้ port 587

ขั้นตอนในการทำงานของ SMTP จะเริ่มต้นจากการ server ของผู้ส่งอีเมล จะสร้างการเชื่อมต่อกับ server SMTP ของผู้รับและให้ข้อมูลที่จำเป็น เช่น ที่อยู่อีเมลของผู้รับ จากนั้น server SMTP จะประมวลผลข้อมูลนี้และยืนยันที่อยู่ของผู้รับเพื่อตัดสินใจว่าจะยอมรับอีเมลหรือไม่ หากที่อยู่

ของผู้รับถูกต้อง อีเมลจะอยู่ในคิวเพื่อจัดส่ง จากนั้น server ของผู้รับจะพยายามส่งอีเมลไปยังกล่องขาเข้าอีเมลของผู้รับหรือโพลเดอร์ที่กำหนด

นอกจากนี้ในปัจจุบัน ยังมี SMTPs หรือ SMTPS (Simple Mail Transfer Protocol Secure) ซึ่งเป็นวิธีการรักษาความปลอดภัยของ SMTP ในเลเยอร์การขนส่ง เพื่อรับรองความถูกต้องและความสมบูรณ์ของการสื่อสาร โดยใช้งาน SSL (Secure Sockets Layer) หรือ TLS (Transport Layer Security) เพื่อสร้างความปลอดภัย

2.12.5 Google Drive

กูเกิล ไดรฟ์ (Google Drive) [54] เป็นบริการออนไลน์ประเภทเทคโนโลยีคลาวด์ ที่มีพื้นที่ให้ผู้ใช้งานการจัดเก็บข้อมูลได้หลากหลายรูปแบบ เช่น ไฟล์เอกสาร ไฟล์รูปภาพ ไฟล์วิดีโอ หรือไฟล์ประเภทอื่นๆ โดยสามารถจัดเก็บไฟล์ได้ในหลากหลายรูปแบบ โดยมีการใช้งานโดยไม่จัดเก็บค่าใช้จ่ายในพื้นที่ 15 จิกะไบต์แรก โดยหลังจากนั้นจะมีตัวเลือกเพื่อเลือกพื้นที่จัดเก็บที่ต้องการซื้อเพิ่ม โดยผู้ที่ใช้งานกูเกิล ไดรฟ์ จำเป็นจะต้องมีบัญชีผู้ใช้งานของจีเมล (Gmail) ก่อน และนอกจากนี้ กูเกิล ไดรฟ์ ยังสามารถแบ่งปันไฟล์ที่อยู่บนคลาวด์ให้ผู้อื่นสามารถเข้าใช้งาน แก้ไข หรือดาวน์โหลดไฟล์ได้ โดยการแชร์ข้อมูลผ่านบัญชีผู้ใช้งานจีเมลได้

การเข้าสู่ข้อมูลบนกูเกิล ไดรฟ์ สามารถเข้าใช้งานได้ในหลากหลายอุปกรณ์ ทั้งบนคอมพิวเตอร์ โทรศัพท์มือถือ หรือแท็บเล็ต ก็สามารถเข้าใช้งาน อัปโหลด ดาวน์โหลด และเปิดดูไฟล์ได้ในตลอดเวลาที่มีการเชื่อมต่ออินเทอร์เน็ต

2.13 งานวิจัยที่เกี่ยวข้อง

พิชญ์สินี เสถียรธราดล [55] ได้มีการศึกษาเรื่อง “ผลกระทบของการเรียนออนไลน์ภายใต้สถานการณ์การแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 : กรณีศึกษานิสิตหลักสูตรการศึกษาศาสตรบัณฑิตและศิลปศาสตรบัณฑิต สาขาวิชาภาษาอังกฤษ มหาวิทยาลัยพะเยา” โดยมีการเปรียบเทียบผลกระทบระหว่างการเรียนการสอนภายในห้องเรียน เทียบกับการเรียนการสอนแบบออนไลน์ ซึ่งจากงานวิจัย พบว่า รูปแบบการเรียนการสอนส่งผลกระทบต่อพฤติกรรมในการเรียน ทำให้การเรียนออนไลน์มีประสิทธิภาพที่น้อยกว่าการเรียนในชั้นเรียน จึงเป็นเหตุผลให้งานวิจัยนี้ทำขึ้น เพื่อแก้ไขปัญหาเรื่องประสิทธิภาพในการเรียนออนไลน์ที่ต่ำกว่าในชั้นเรียน โดยให้มีระบบตรวจสอบอารมณ์เพื่อให้ผู้เรียนสนใจบทเรียนได้มากขึ้น

ตริรัตน์ เสริมทรัพย์ และคณะ [56] ได้ศึกษาวิจัยเชิงสำรวจการรู้จำใบหน้า โดยเป็นการศึกษาวิธีการรู้จำใบหน้า 5 วิธีซึ่งประกอบไปด้วย วิธีใบหน้าไอแกน (Eigen Faces) วิธีโครงข่ายประสาทเทียม (Neural networks) วิธีโมเดลสี RGB-HSV-YCbCr รวมกับมอร์โฟโลยี (Color Model

and Morphology Technique) วิธี Local probabilistic subspace และ วิธีการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis : PCA) โดยพบว่าวิธีโมเดลสี RGB-HSV-YCbCr รวมกับมอร์โฟโลยี (Color Model and Morphology Technique) ได้ผลลัพธ์ที่ดีที่สุด แต่มีข้อจำกัดในด้านสีและแสงค่อนข้างมาก รองลงมาคือวิธีโครงข่ายประสาทเทียม (Neural networks) ภายในงานวิจัยนี้จึงเลือกใช้งานวิธีโครงข่ายประสาทเทียม เพื่อลดข้อจำกัดในด้านสีและแสงของภาพเนื่องจากแสงและเงาที่สะท้อนจากแต่ละพื้นที่มีความแตกต่างกันค่อนข้างมาก

ศศิชา บุญเก่า และวิฑิต ฉัตรรัตนกุลชัย [57] ได้ศึกษาวิจัยการจับความรู้สึกของคนจากใบหน้าด้วยเทคนิคปัญญาประดิษฐ์ โดยการใช้ Pretrained CNN สำหรับการจำแนกอารมณ์จากการแสดงออกทางสีหน้าโดยใช้วิธีการ Transfer Learning เพื่อระบุอารมณ์ของมนุษย์ทั้งหมด 7 อารมณ์ ได้แก่ อารมณ์โกรธ (Angry) อารมณ์ขยะแขยง (Disgust) อารมณ์กลัว (Fear) มีความสุข (Happy) ความเศร้า (Sad) ประหลาดใจ (Surprise) และอารมณ์เป็นกลาง (Neutral) โดยมีความเกี่ยวข้องกับงานวิจัยนี้ เนื่องจากเป็นการจำแนกอารมณ์ของคนจากใบหน้าออกเป็น 7 อารมณ์เหมือนกัน และใช้ Pretrained CNN และ Transfer Learning เพื่อพัฒนาโมเดลในการตรวจจับอารมณ์เช่นเดียวกัน

วัชรชัย คงศิริวัฒนา และคณะ [58] ได้ศึกษาวิจัยเกี่ยวกับระบบตรวจสอบการเข้าชั้นเรียนและประเมินความสนใจผ่านลักษณะอารมณ์ทางใบหน้าด้วยกล้องเว็บแคม ซึ่งมีการสร้างระบบเพื่อตรวจสอบการเข้าชั้นเรียน โดยเทคนิคการรู้จำใบหน้า (Face Recognition) ด้วยเทคนิคโครงข่ายประสาทเทียมคอนโวลูชัน และการประเมินความสนใจในชั้นเรียน ซึ่งใช้โมเดล mini-Xception ในการจำแนกอารมณ์ โดยเป็นการตรวจสอบการเข้าเรียนและอารมณ์ของผู้เรียนในการเรียนแบบ on-site โดยมีการจำแนกลักษณะอารมณ์ทั้ง 6 ออกเป็นสองกลุ่ม คือ มีความสนใจเรียน ได้แก่ อารมณ์เป็นกลาง (Neutral) , มีความสุข (Happy) และประหลาดใจ (Surprise) อารมณ์ที่ไม่สนใจเรียน คืออารมณ์โกรธ (Angry) , เศร้า (Sad) และกลัว (Fear) โดยภายในงานวิจัยนี้ได้มีการแบ่งอารมณ์ออกเป็นสองกลุ่ม คือตั้งใจเรียนและไม่ตั้งใจเรียน โดยอารมณ์เชิงบวก จะแสดงถึงความตั้งใจเรียน และอารมณ์เชิงลบแสดงถึงความไม่ตั้งใจเรียน

ฐิติพงศ์ รักษาภิรมณ์ และธนภัทร พังคะจิตร [59] ได้ศึกษาวิจัยเกี่ยวกับการจำแนกอารมณ์จากใบหน้าแบบ Real-time บนอุปกรณ์ฝังตัว ด้วยแบบจำลองการเรียนรู้เชิงลึก โดยมีการจำแนกออกเป็น 7 และ 2 กลุ่มอารมณ์ตามลำดับ โดยใช้ข้อมูลในการ trained model เป็น FER-2013 dataset

ภูบัติ ศิวาวงศ์ และคณะ [60] ได้ศึกษาและวิจัยภายในหัวข้อ “ระบบตรวจสอบเพื่อยืนยันตัวบุคคลด้วยใบหน้าบนแอนดรอยด์” โดยใช้เทคนิคการรู้จำใบหน้า เพื่อตรวจสอบภาพใบหน้าของบุคคลว่าอยู่ในฐานข้อมูลที่กำหนดไว้หรือไม่ โดยมีขั้นตอนที่ใช้ในการรู้จำใบหน้า โดยใช้ AdaBoost เป็นกระบวนการหา Feature ที่มีลักษณะใกล้เคียงและแตกต่างของภาพตัวอย่าง โดยกำหนดน้ำหนักให้

Haar – like feature เพื่อตรวจสอบภาพ โดยแยกภาพที่เราต้องการและไม่ต้องการด้วย Cascade classifiers โดยมีวิธีการทดสอบประสิทธิภาพโดยการใช้รูปภาพจากผู้เข้าทดสอบจำนวน 7 คน คนละ 10 รูป โดยนำไปทดสอบกับภาพที่ใช้ในการเรียนรู้ 140 ภาพ จาก 7 คน คนละ 20 ภาพ โดยรู้จำใบหน้าถูกต้อง 60 รูป จาก 70 รูป คิดเป็น 85.71%

P. Claudio และคณะ [61] ได้ทำงานวิจัยภายในหัวข้อ “Face and eye tracking algorithm based on Digital image processing” โดยเป็นการตรวจสอบภาพ โดยการใช้งานการประมวลผลภาพ เพื่อหาตำแหน่งของใบหน้าและดวงตาภายในภาพ โดยในการทำงานจะแบ่งออกเป็น 5 ขั้นตอน ประกอบด้วย 1) การหาพื้นที่ของใบหน้าอย่างคร่าวๆ โดยการแปลงภาพให้เป็นภาพขาวดำ และแปลงภาพให้เป็นเวกเตอร์ เพื่อหาพื้นที่ที่มีลักษณะของเวกเตอร์คล้ายวงรี 2) การหาพื้นที่ของใบหน้า จากพื้นที่ที่หาในข้อแรก จึงนำมาตรวจสอบและสร้างเป็นรูปร่างรี 3) หาพื้นที่ที่คาดว่าจะป็นดวงตา โดยการเทียบทางสถิติว่าพื้นที่ของดวงตาจะอยู่ภายในช่วงใดบ้างของใบหน้า 4) การตรวจสอบพื้นที่ที่ม่านตาและรูม่านตา และ 5) การตรวจจับดวงตาและรูม่านตา โดยภายในงานวิจัยนี้มีการตรวจสอบประสิทธิภาพโดยการทำการทดสอบโดยการใช้ฐานข้อมูลของรูปภาพจำนวน 2 ฐานข้อมูล โดยมีจำนวน 102 และ 897 ภาพ และนำมาตรวจสอบความถูกต้อง โดยมีความถูกต้อง 99 และ 100% ตามลำดับ

ตารางที่ 2.3 สรุปวิธีการหรือเทคนิคที่ใกล้เคียงกับวิธีการของงานวิจัย

ลำดับ	ชื่อหัวข้อวิจัย	วัตถุประสงค์	เทคนิค	ผลลัพธ์ของงานวิจัย
1	ผลกระทบของการเรียนออนไลน์ภายใต้สถานการณ์การแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 : กรณีศึกษานิสิตหลักสูตรการศึกษานิเทศศาสตร์และศิลปศาสตรบัณฑิต สาขาวิชาภาษาอังกฤษ มหาวิทยาลัยพะเยา [55]	ศึกษาผลกระทบที่เกิดขึ้นจากการเรียนออนไลน์	การทำแบบสอบถาม, ค่าเฉลี่ย, ส่วนเบี่ยงเบนมาตรฐาน	รูปแบบการเรียนการสอนส่งผลกระทบต่อพฤติกรรมในการเรียน โดยมีค่าเฉลี่ยรวม 3.06 คะแนน (จาก 5 คะแนน) และส่วนเบี่ยงเบน

ตารางที่ 2.3 สรุปวิธีการหรือเทคนิคที่ใกล้เคียงกับวิธีการของงานวิจัย (ต่อ)

ลำดับ	ชื่อหัวข้อวิจัย	วัตถุประสงค์	เทคนิค	ผลลัพธ์ของงานวิจัย
				มาตรฐานรวมเท่ากับ 1.91
2	วิจัยเชิงสำรวจการรู้จำใบหน้า [56]	สร้างระบบเพื่อตรวจสอบการเข้าชั้นเรียนและการประเมินความสนใจในชั้นเรียน โดยเป็นการตรวจสอบการเข้าเรียนและอารมณ์ของผู้เรียนในการเรียนแบบ on-site	วิธีใบหน้าไอแกน, วิธีโครงข่ายประสาทเทียม, วิธีโมเดลสี RGB-HSV-YCbCr รวมกับมอร์โฟโลยี, วิธี Local probabilistic subspace และวิธีการวิเคราะห์องค์ประกอบหลัก	ค่าเฉลี่ยในการรู้จำใบหน้าของแต่ละโมเดล มีค่าดังนี้ วิธีโมเดลสี RGB-HSV-YCbCr รวมกับมอร์โฟโลยี ร้อยละ 97.64, Neural networks ร้อยละ 94.94, Eigen ร้อยละ 91.83, Local probabilistic subspace ร้อยละ 82.30
3	การจับความรู้สึกของคนจากใบหน้าด้วยเทคนิคปัญญาประดิษฐ์ [57]	สามารถตรวจจับและแบ่งประเภทอารมณ์บนใบหน้าออกเป็น 7 กลุ่มอารมณ์ ได้แก่ อารมณ์โกรธ, อารมณ์	Transfer Learning GoogLeNet Inception-v3 Transfer Learning	ให้ผู้เข้าร่วมการทดลอง 30 คน ทำใบหน้าตามอารมณ์ต่างๆ และวัดประสิทธิภาพในการตรวจจับ โดย

ตารางที่ 2.3 สรุปวิธีการหรือเทคนิคที่ใกล้เคียงกับวิธีการของงานวิจัย (ต่อ)

ลำดับ	ชื่อหัวข้อวิจัย	วัตถุประสงค์	เทคนิค	ผลลัพธ์ของงานวิจัย
		ขยะแขยง, อารมณ์กลัว, มีความสุข, ความเศร้า, ประหลาดใจ, เป็นกลาง		โมเดล GoogLeNet มีประสิทธิภาพสูงที่สุด
4	ระบบตรวจสอบการเข้าชั้นเรียนและประเมินความสนใจผ่านลักษณะอารมณ์ทางใบหน้าด้วยกล้องเว็บแคม [58]	รู้จำใบหน้าเพื่อใช้ตรวจสอบในการเข้าชั้นเรียน, วัดความสนใจในการเรียนผ่านอารมณ์บนใบหน้า	CNN , แบบจำลองการเรียนรู้ mini-Xception	ทดสอบประสิทธิภาพโดยใช้รูปภาพจากผู้เข้าทดสอบจำนวน 7 คน คนละ 10 รูป โดยนำไปทดสอบกับภาพที่ใช้ในการเรียนรู้ 140 ภาพ จาก 7 คน คนละ 20 ภาพ โดยรู้จำใบหน้าถูกต้อง 60 รูป จาก 70 รูป คิดเป็น 85.71%
5	การจำแนกอารมณ์จากใบหน้าแบบ Real-time บนอุปกรณ์ฝังตัว [59]	ตรวจสอบและจำแนกอารมณ์แบบ Real-time ออกเป็น 7 กลุ่มอารมณ์	Convolution Neural Network Xception AlexNet MobileNet	โมเดลที่ได้ประสิทธิภาพสูงที่สุดคือโมเดลที่ผู้วิจัยพัฒนาขึ้นเอง โดยมีค่า

ตารางที่ 2.3 สรุปวิธีการหรือเทคนิคที่ใกล้เคียงกับวิธีการของงานวิจัย (ต่อ)

ลำดับ	ชื่อหัวข้อวิจัย	วัตถุประสงค์	เทคนิค	ผลลัพธ์ของงานวิจัย
		ได้บนอุปกรณ์ฝังตัว		accuracy อยู่ที่ 86.07%
6	ระบบตรวจสอบเพื่อยืนยันตัวบุคคลด้วยใบหน้าบนแอนดรอยด์ [60]	ใช้ AdaBoost เป็นกระบวนการหา Feature ที่มีลักษณะใกล้เคียงและแตกต่างของภาพตัวอย่าง โดยกำหนดน้ำหนักให้ Haar – like feature เพื่อตรวจสอบภาพ โดยแยกภาพที่เราต้องการและไม่ต้องการด้วย Cascade classifiers	Harr-like Feature AdaBoost Algorithm Cascade Classifiers	ใช้รูปภาพจาก 7 คน คนละ 10 รูป โดยนำไปทดสอบกับภาพที่ใช้ในการเรียนรู้ 140 ภาพ จาก 7 คน คนละ 20 ภาพ โดยรู้จำใบหน้าถูกต้อง 60 รูป จาก 70 รูป คิดเป็น 85.71% [59]
7	Face and Eye Tracking Algorithm Based on Digital Image Processing [61]	เพื่อตรวจสอบหาพื้นที่ของใบหน้าและดวงตา โดย	Find Face Location Algorithm	ทดสอบโดยการใช้ฐานข้อมูลของรูปภาพจำนวน 2 ฐานข้อมูล โดยมี

ตารางที่ 2.3 สรุปวิธีการหรือเทคนิคที่ใกล้เคียงกับวิธีการของงานวิจัย (ต่อ)

ลำดับ	ชื่อหัวข้อวิจัย	วัตถุประสงค์	เทคนิค	ผลลัพธ์ของงานวิจัย
		การประมวลผลภาพ		จำนวน 102 และ 897 ภาพ และนำมาตรวจสอบความถูกต้อง โดยมีความถูกต้อง 99 และ 100% ตามลำดับ

จากทฤษฎีทั้งหมดที่กล่าวมาข้างต้น จึงสามารถสรุปได้ว่า โมเดลที่เหมาะสมกับการทำงานภายในงานวิจัย เรื่อง การตรวจจับอารมณ์บนใบหน้าด้วยระบบปัญญาประดิษฐ์ คือโมเดล Convolution Neural Network (CNN) แต่เนื่องจากการ training โมเดลเองทั้งหมด จะต้องใช้ทรัพยากรคอมพิวเตอร์ และเวลามาก จึงมีการเลือกใช้ Pre-trained Model โดยที่เลือกใช้งาน ResNet-50 เนื่องจากเป็นโมเดลที่มีการพัฒนาต่อจนมีความแม่นยำสูง ใช้ทรัพยากรคอมพิวเตอร์อยู่ในระดับปานกลาง โดยชุดข้อมูลที่น่ามาฝึกฝนโมเดลเพิ่มเติมคือ ชุดข้อมูล FER-2013 ซึ่งเป็นชุดข้อมูลที่ถูกรวบรวมอยู่บนเว็บไซต์ Kaggle ซึ่งเป็นเว็บไซต์ที่มีการรวบรวมฐานข้อมูลหลากหลายประเภทอยู่ด้วยกัน และสามารถใช้งานได้โดยที่ไม่มีค่าใช้จ่าย โดยชุดข้อมูล FER-2013 เป็นชุดข้อมูลที่ได้รับความนิยมในการนำมาจำแนกอารมณ์บนใบหน้า และมีการจำแนกอารมณ์ออกเป็น 7 กลุ่มอารมณ์ดังที่งานวิจัยต้องการ จึงเลือกใช้ฐานข้อมูลนี้

นอกจากนี้ จากงานวิจัยที่ได้กล่าวถึงได้ข้างต้นนี้ งานวิจัยของวัชรชัย คงศิริวัฒนา และคณะ [58] และฐิติพงศ์ รักษาภิรมณ์ และธนภัทร ชังคะจิตร [59] ได้กล่าวถึงการแบ่งอารมณ์บนใบหน้า จาก 6 และ 7 กลุ่มอารมณ์ เป็น 2 กลุ่มอารมณ์ โดยงานวิจัยที่กล่าวมาข้างต้นทั้งหมด มีการแบ่งกลุ่มอารมณ์เช่นเดียวกันทั้งหมด คือ กลุ่มที่มีความสนใจ มีความสนใจ ประกอบด้วย “ปกติ (Neutral)” “มีความสุข (Happy)” และ “ประหลาดใจ (Surprised)” มีความไม่สนใจ ประกอบด้วย “โกรธ (Angry)” “เศร้า (Sad)” และ “กลัว (Fear)” โดยในงานวิจัยนี้ ได้ใช้ภาษาไพทอนในการพัฒนาซอฟต์แวร์เพื่อตรวจจับภาพใบหน้า โดยมีการส่งข้อมูลผ่านทาง Gmail และอัปโหลดไฟล์วิดีโอผ่าน Google Drive เพื่อส่งผลการทดสอบไปยังอาจารย์ผ่านเครือข่ายอินเทอร์เน็ต

บทที่ 3

ระเบียบและวิธีการดำเนินการวิจัย

จากการศึกษางานวิจัย แนวคิด และทฤษฎีที่เกี่ยวข้อง ผู้ศึกษาได้นำความรู้มาประยุกต์ใช้เพื่อเปรียบเทียบกับวิธีการต่างๆ ที่ใช้งานในการตรวจจับใบหน้า และการตรวจจับอารมณ์ เพื่อนำมาใช้งานในระบบการตรวจจับอารมณ์ของผู้เรียนระหว่างการเรียนรู้ออนไลน์ โดยใช้ระบบปัญญาประดิษฐ์ โดยมีขั้นตอนการดำเนินงาน ดังต่อไปนี้

ขั้นตอนที่ 1 การวิเคราะห์ข้อมูลปัญหาของงานวิจัย

ขั้นตอนที่ 2 การศึกษาข้อมูลและเทคโนโลยีที่นำมาใช้

ขั้นตอนที่ 3 การออกแบบการทดลอง

ขั้นตอนที่ 4 การดำเนินการทดลอง

ขั้นตอนที่ 5 การสรุปผลการวิจัย

ในการดำเนินงานวิจัย ตั้งแต่เดือนสิงหาคม พ.ศ.2565 จนถึงเดือนตุลาคม พ.ศ.2566 มีขั้นตอนและระยะเวลาในการดำเนินงาน โดยแบ่งแยกย่อยเป็น 5 ขั้นตอนที่ได้กล่าวมาข้างต้น ตามตารางที่ 3.1

ตารางที่ 3.1 : แผนในการดำเนินงานวิจัย

การดำเนินงาน	ระยะเวลา															
	พ.ศ. 2565					พ.ศ. 2566										
	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	
ขั้นตอนที่ 1 : การวิเคราะห์ข้อมูลปัญหาของงานวิจัย																
- ศึกษาปัญหา และค้นหาหัวข้อในการวิจัย																
ขั้นตอนที่ 2 : การศึกษาข้อมูลและเทคโนโลยีที่นำมาใช้																
การทบทวนวรรณกรรม																
ร่างเค้าโครงวิทยานิพนธ์																
นำเสนอเค้าโครงวิทยานิพนธ์																
ขั้นตอนที่ 3 : การออกแบบการทดลอง																
ศึกษาขั้นตอนการสร้างโมเดล																
- เขียนโปรแกรมเพื่อใช้ในการทดลอง																
ขั้นตอนที่ 4 การดำเนินการทดลอง																
- เก็บข้อมูลจากกลุ่มตัวอย่าง																
ขั้นตอนที่ 5 : การสรุปผลการวิจัย																
- สรุปผลการวิจัย																
- ตีพิมพ์งานวิจัย																
- เขียนวิทยานิพนธ์เล่มสมบูรณ์																
- สอบป้องกันวิทยานิพนธ์																

3.1 การวิเคราะห์ข้อมูลปัญหาของงานวิจัย

ผู้วิจัยได้ทำการศึกษาปัญหาที่เกิดขึ้นในสถานการณ์ปัจจุบันในการศึกษา ซึ่งมีผลเกี่ยวเนื่องมาจากโรคระบาด COVID-19 และในปัจจุบันมีการนำระบบปัญญาประดิษฐ์ (AI: Artificial

Intelligence) มาประยุกต์ใช้ได้อย่างกว้างขวางตั้งแต่ระดับบุคคลทั่วไปจนถึงระดับอุตสาหกรรม สังคม เศรษฐกิจ รวมถึงในด้านการศึกษา เนื่องจากในขณะที่ยังเริ่มต้นทำงานวิจัย ผู้วิจัยได้ประสบปัญหาในการเรียนออนไลน์ ซึ่งบรรยากาศ ลักษณะการเรียนการสอนไม่เหมือนกับการเรียนแบบออฟไลน์ จึงส่งผลกระทบต่ออารมณ์ในการเรียน ทำให้ความสนใจเรียนลดลง และมีผลให้ความเข้าใจในชั้นเรียนลดลง จึงศึกษาเทคโนโลยีด้านปัญญาประดิษฐ์เพื่อค้นหาวิธีที่จะจัดการเกี่ยวกับความไม่ตั้งใจเรียน และทำให้การสื่อสารภายในเรียนในรูปแบบออนไลน์คล้ายคลึงกับการเรียนแบบออฟไลน์ให้ได้มากที่สุด

จากการศึกษาปัญหาที่เกิดขึ้นรวมถึงแนวทางในการแก้ไขปัญหา ทำให้ผู้วิจัยปรึกษาและเสนอหัวข้อกับอาจารย์ที่ปรึกษาที่จะเลือกที่จะเลือกทำงานวิจัยในหัวข้อ "การตรวจจับอารมณ์ของผู้เรียนระหว่างการเรียนออนไลน์ โดยใช้ระบบปัญญาประดิษฐ์"

3.2 การศึกษาข้อมูลและเทคโนโลยีที่นำมาใช้

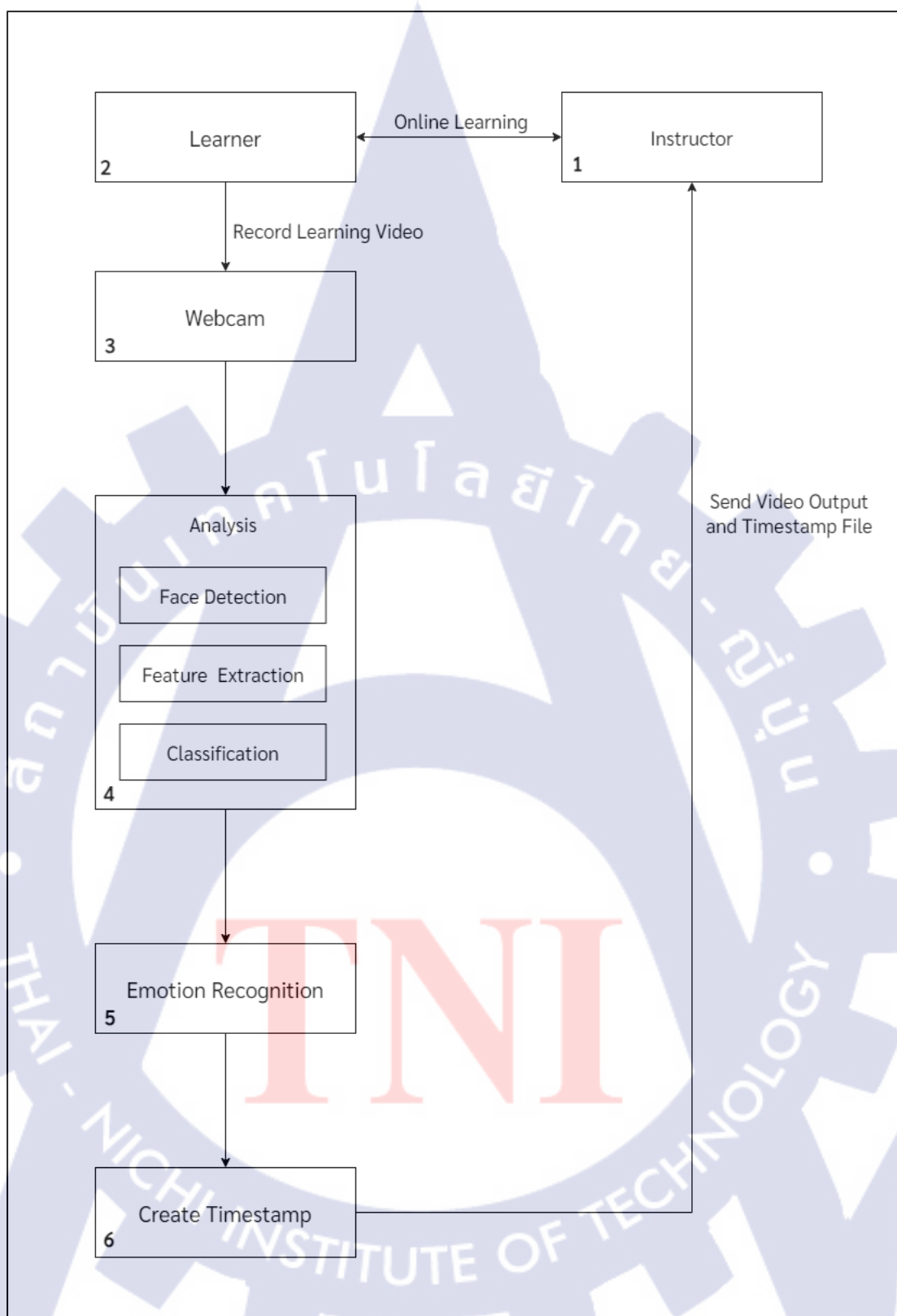
ผู้วิจัยได้ทำการศึกษางานวิจัยเกี่ยวกับเทคนิคการตรวจจับใบหน้า และตรวจจับอารมณ์บนใบหน้า ซึ่งเป็นส่วนหนึ่งของเทคโนโลยีปัญญาประดิษฐ์ (AI) โดยเทคนิคการตรวจสอบลักษณะใบหน้าของบุคคลที่เป็นที่นิยมในปัจจุบันนี้มีหลากหลายเทคโนโลยี แต่เทคโนโลยีที่เลือกใช้งานในงานวิจัยนี้คือ โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network) เนื่องจากในงานวิจัยนี้จะแบ่งการตรวจจับออกเป็น 2 ส่วน คือ 1) การตรวจจับใบหน้า และ 2) การตรวจจับอารมณ์ ซึ่งเทคโนโลยีที่เหมาะสมกับการตรวจจับใบหน้า และเหมาะสมกับการตรวจจับอารมณ์ด้วย คือ โครงข่ายประสาทเทียมคอนโวลูชัน เนื่องจากเทคโนโลยีโครงข่ายประสาทเทียมคอนโวลูชันจะมีชั้นคอนโวลูชันซึ่งสามารถสกัดแยกคุณลักษณะออกมาได้ จึงเหมาะกับการตรวจจับใบหน้า และการตรวจจับอารมณ์บนใบหน้า ที่จะต้องมีการตรวจสอบลักษณะต่างๆ บนใบหน้าแยกกัน ทั้งริมฝีปาก ดวงตา คิ้ว เพื่อตรวจจับอารมณ์ที่ได้ ซึ่งจากการศึกษางานวิจัยในการตรวจสอบอารมณ์บนใบหน้าแล้วพบว่า การสร้างโมเดลเพื่อตรวจสอบอารมณ์บนใบหน้าจำเป็นต้องใช้ทรัพยากรของเครื่องคอมพิวเตอร์สูง ผู้วิจัยจึงเลือกใช้ pre-trained model เพื่อสร้างโมเดลในการตรวจสอบอารมณ์บนใบหน้า และใช้โมเดล MTCNN (Multi-task Cascaded Convolutional Networks) ซึ่งเป็นโมเดลที่ถูกสร้างโดยใช้หลักการของ CNN และสามารถสกัดพื้นที่ในแต่ละส่วนของใบหน้าได้ โดยชุดข้อมูลที่เลือกนำมาใช้เพื่อพัฒนาโมเดล คือ ชุดข้อมูล FER-2013 เนื่องจากเป็นชุดข้อมูลที่มีการแบ่งอารมณ์ออกเป็น 7 กลุ่มอารมณ์ และมีการนำไปใช้งานในงานวิจัยจำนวนมาก โดยชุดข้อมูล FER-2013 เป็นชุดข้อมูลที่ถูกรวบรวมจากบนเว็บไซต์ Kaggle และเปิดให้เข้าใช้งานได้ฟรี โดยมีการนำเสนอชุดข้อมูลนี้ที่ International Conference on Machine Learning (ICML) ในปี 2013 โดย Pierre-Luc Carrier และ Aaron Courvil

3.3 การออกแบบการทดลอง

จากการศึกษาข้อมูลและเทคโนโลยีที่นำมาใช้ จึงได้มีการออกแบบระบบการเรียนรู้ออนไลน์ที่จะนำมาช่วยแก้ปัญหาข้างต้น ดังแสดงในรูปที่ 3.1 โดยในขณะที่ผู้เรียนกำลังอยู่ในชั้นเรียน ระบบจะมีการตรวจจับใบหน้าเพื่อวิเคราะห์และตรวจจับอารมณ์ของผู้เรียนได้ ซึ่งจะทำให้ผู้สอนสามารถประเมินความพึงพอใจของผู้เรียนได้

โดยรายละเอียดของระบบ จะประกอบด้วย 6 ส่วน ดังนี้ 1) ผู้สอน (Instructor) 2) ผู้สอน (Learner) 3) กล้องเว็บแคม (Webcam) 4) การวิเคราะห์ข้อมูลของระบบ (Analysis) 5) การรู้จำอารมณ์ (Emotion Recognition) และ 6) การบันทึกเวลา (Create Timestamp) ซึ่งจะระบุเวลาที่เกิดการเปลี่ยนแปลงอารมณ์ในวิดีโอ ซึ่งแต่ละส่วนมีรายละเอียด ดังนี้





รูปที่ 3.1 แสดงการออกแบบระบบ

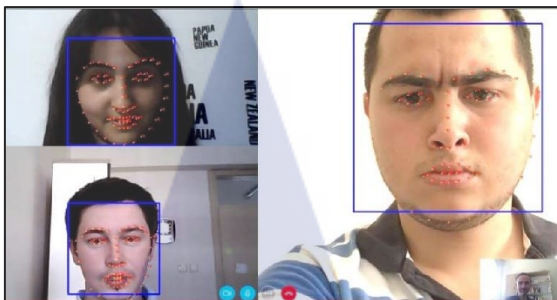
1) ผู้สอน (Instructor) : เป็นองค์ประกอบสำคัญของระบบนี้ โดยมีหน้าที่สร้างและออกแบบหลักสูตรเนื้อหา แบบทดสอบ และการประเมินผลในการเรียนรู้ออนไลน์ โดยเนื้อหาการเรียนภายใน การทดลองนี้ จะใช้วิดีโอบันทึกการสอนวิชา Data Science ความยาว 20.00 นาที ซึ่งเนื้อหาจะถูก แบ่งออกเป็น 6 ส่วน ได้แก่ เนื้อหาส่วนที่ 1 (ช่วง 0.00 – 3.00 นาที) : บทนำ เนื้อหาส่วนที่ 2 (ช่วง 3.01 – 8.45 นาที) อธิบายความหมายและความแตกต่างของ Data Science และ Machine Learning เนื้อหาส่วนที่ 3 (ช่วง 8.46 – 11.30 นาที) อธิบายประเภทของข้อมูล ที่สามารถนำไปใช้งาน ใน Machine Learning เนื้อหาส่วนที่ 4 (ช่วง 11.31 – 15.53 นาที) อธิบาย AI Company และ ยกตัวอย่าง เนื้อหาส่วนที่ 5 (ช่วง 15.54 - 19.21 นาที) อธิบาย Machine Learning และยกตัวอย่าง เนื้อหาส่วนที่ 6 (19.22 - 20.00 นาที) เกริ่นนำสู่เนื้อหาบทต่อไป ดังแสดงในตารางที่ 3.2

ตารางที่ 3.2 แสดงหัวข้อในบทเรียนแต่ละส่วน

เนื้อหาส่วนที่	หัวข้อ	เวลา (นาที)	เวลา (วินาที)
1	บทนำ	0.00 – 3.00	180
2	อธิบายความหมายและความแตกต่างของ Data Science และ Machine Learning	3.01 – 8.45	344
3	อธิบายประเภทของข้อมูล ที่สามารถนำไปใช้งานใน Machine Learning	8.46 – 11.30	164
4	อธิบาย AI Company และยกตัวอย่าง	11.31 – 15.53	262
5	อธิบาย Machine Learning และยกตัวอย่าง	15.54 - 19.21	207
6	เกริ่นนำสู่เนื้อหาบทต่อไป	19.22 - 20.00	38

2) ผู้เรียน (Learner) : บุคคลที่ลงทะเบียนเรียนออนไลน์ และทำการทดลองผ่านการเรียนผ่านวิดีโอที่กำหนดให้

3) กล้องเว็บแคม (Webcam) : วิดีโอจะถูกบันทึกโดยใช้เว็บแคม โดยกล้องจะแสดงผลมุมมองเฉพาะของผู้เรียนที่เข้าเรียน ระบบจะพยายามตรวจหาเนื้อหาส่วนที่ต้องการตรวจสอบ (Region of Interest : ROI) เช่น ใบหน้า ท่าทาง ท่าทาง หรือตา ของผู้เรียนระหว่างการเรียนรู้ ดังแสดงในรูปที่ 3.2



รูปที่ 3.2 การตรวจจับใบหน้า

4) การวิเคราะห์อารมณ์ของผู้เรียน (Analysis) : เพื่อทำความเข้าใจว่าอารมณ์ของผู้เรียนมีลำดับอย่างไรในระหว่างกระบวนการเรียนรู้ เพื่อพัฒนาระบบการเรียนรู้ที่รับรู้และตอบสนองอย่างเหมาะสมต่อการเปลี่ยนแปลงทางอารมณ์ของผู้เรียน โดยในขั้นตอนนี้จะถูกแบ่งออกเป็น 3 ส่วน คือ

1) การตรวจจับใบหน้า (Face Detection) จากภาพวิดีโอที่รับข้อมูลเข้ามา ซึ่งวิดีโอเหล่านั้นคือภาพนิ่งที่เรียงต่อกันจำนวนมาก จะนำภาพนิ่งที่ได้มาตรวจหาพื้นที่ที่คาดว่าจะจะเป็นบริเวณใบหน้า โดยในขั้นตอนนี้จะทำงานโดยใช้เทคโนโลยี MTCNN (Multi-task Cascaded Convolutional Networks) ในการทำงาน โดยจะเป็นการทำงานในช่วง P-Net และ R-Net ของ MTCNN

2) การสกัดคุณลักษณะ (Feature Extraction) หลังจากที่ได้หาบริเวณที่คาดว่าจะจะเป็นใบหน้าได้แล้ว จะมีกระบวนการถัดไป คือ การตรวจสอบคุณลักษณะต่างๆบนใบหน้า ว่าบริเวณที่คาดว่าจะจะเป็นใบหน้านั้น มีลักษณะส่วนประกอบของใบหน้าถูกต้องหรือไม่ มีตา จมูก คิ้ว ริมฝีปากครบถ้วนหรือไม่ หากไม่ จะจำแนกว่าภาพนี้ไม่ใช่ภาพของใบหน้า โดยภายในขั้นตอนนี้เป็นการทำงานในช่วง O-Net ของ MTCNN ซึ่งเป็นขั้นตอนการทำงานที่มีการนำ Fully Convolutional Neural Network มาใช้งาน ซึ่งแตกต่างกับขั้นตอนในช่วง P-Net และ R-Net ที่นำมาใช้งานเฉพาะ Fully Connected Layer

ในการสกัดคุณลักษณะนี้ ภายในโมเดลจะมีการทำงานโดยการหาพื้นที่ 5 ตำแหน่งของใบหน้า บริเวณดวงตาทั้งสองข้าง, จมูก, และขอบริมฝีปากทั้งด้านซ้ายและขวา

3) การจำแนกประเภท (Classification) คือ การจำแนกประเภทข้อมูลภาพที่รับเข้ามา และผ่านการสกัดคุณลักษณะแล้ว ว่าเป็นภาพที่ถูกรับเข้ามาเป็นภาพใบหน้าหรือไม่

5) การรู้จำอารมณ์ (Emotion Recognition): จากผลการเคลื่อนไหวของใบหน้าที่เกี่ยวข้องกับอารมณ์ของผู้เรียน ระบบจะสามารถวิเคราะห์ความสัมพันธ์ระหว่างอารมณ์ของผู้เรียนกับการจำแนกประเภทอารมณ์ของผู้เรียน ระบบจะจำแนกอารมณ์บนใบหน้าเป็น 7 กลุ่มอารมณ์ และจะขึ้น

ข้อความบริเวณมุมซ้ายบนของกรอบสี่เหลี่ยมที่ตรวจจับใบหน้าตามอารมณ์ 7 อารมณ์ที่ตรวจจับได้ หลังจากนั้นจะจำแนกข้อมูลออกเป็น 2 กลุ่ม คือกลุ่มที่สนใจบทเรียน ซึ่งประกอบไปด้วย “ปกติ (Neutral)” “มีความสุข (Happy)” และ “ประหลาดใจ (Surprised)” และกลุ่มที่ไม่สนใจบทเรียน ประกอบด้วย “โกรธ (Angry)” “ขยะแขยง (Disgusted)” “เศร้า (Sad)” และ “กลัว (Fear)” เพื่อเก็บข้อมูลไปเปรียบเทียบกับภาพในวีณาที่ก่อนหน้านี้ และบันทึกลงในไฟล์ในขั้นตอนต่อไป

6) การบันทึกเวลา (Create Timestamp) : ระบบจะทำการบันทึกอารมณ์ที่เปลี่ยนแปลงไปในแต่ละช่วงลงในไฟล์ “.txt” ตามที่ได้รับมาจากขั้นตอนก่อนหน้านี้ และส่งอีเมลการบันทึกเวลา และวิดีโอที่บันทึกระหว่างการเรียนกลับไปยังผู้สอน

3.4 การพัฒนาโมเดลเพื่อตรวจจับอารมณ์ของผู้เรียน

ในงานวิจัยนี้ มีการใช้ pre-trained model เพื่อพัฒนาในการตรวจสอบอารมณ์ของผู้เรียน โดยในงานวิจัยนี้มีการใช้งาน ResNet-50 เพื่อเป็นต้นแบบในการสร้างโมเดล โดยมีการปรับค่าตัวแปรที่มีการใช้งานในโมเดล ดังนี้

- 1) ปรับขนาดภาพที่ใช้ในการฝึกฝนโมเดล ให้มีขนาด 224 * 224 pixels
- 2) ปรับโหมดสีของภาพเป็น RGB ซึ่งเป็นโหมดสีที่เหมาะสมกับการใช้งานบนคอมพิวเตอร์มากกว่า
- 3) สลับลำดับข้อมูลภายในชุดข้อมูล เพื่อให้ข้อมูลภาพทั้ง 7 อารมณ์กระจายตัวกันได้
- 4) สร้างฟังก์ชันเพื่อ create model โดยมีการปรับค่าตัวแปรต่างๆ ดังรูปที่ 3.3

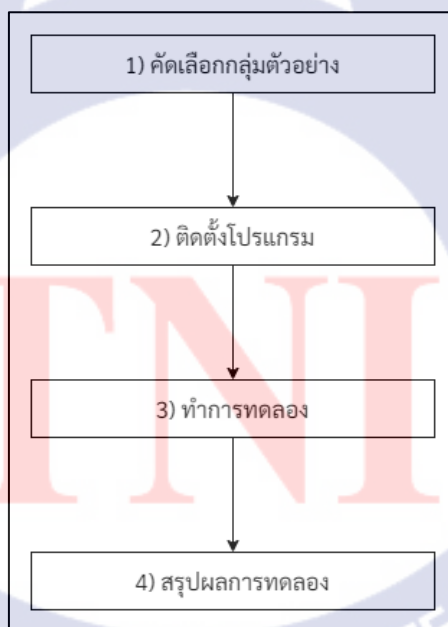
```
def Create_ResNet50V2_Model():
    model = Sequential([
        ResNet50V2,
        Dropout(.25),
        BatchNormalization(),
        Flatten(),
        Dense(64, activation='relu'),
        BatchNormalization(),
        Dropout(.5),
        Dense(7, activation='softmax')
    ])
    return model
```

รูปที่ 3.3 แสดงการปรับตัวแปรในการเรียนรู้ของโมเดล

- สร้างตัวแปร model เพื่อเก็บค่าโมเดลที่จะสร้าง โดยใช้ ResNet50V2 เป็น pre-trained model
- ปรับค่า dropout = 0.25 เพื่อลดการ overfit ของโมเดล
- เพิ่ม BatchNormalization เพื่อเพิ่มความเสถียรของโมเดล
- เพิ่ม Flatten layer เพื่อเปลี่ยนข้อมูลจากลักษณะของภาพ 2 มิติเป็นลักษณะแบบเวกเตอร์ 1 มิติ เพื่อเตรียมข้อมูลสำหรับการใช้ในเลเยอร์ Dense
- Dense layer เลือกใช้ activation function คือ ReLU
- หลังจากนั้นเพิ่ม BatchNormalization และ dropout อีกครั้ง
- เพิ่ม dense layer อีกหนึ่งชั้น และใช้ activation function คือ softmax

3.5 ขั้นตอนในการดำเนินงานวิจัย

ขั้นตอนในการดำเนินงานวิจัยนี้ จะแบ่งออกเป็น 4 ขั้นตอน ประกอบด้วย 1) การคัดเลือกกลุ่มตัวอย่าง 2) การติดตั้งโปรแกรม 3) การทำการทดลอง และ 4) การสรุปผลการทดลอง ดังรูปที่ 3.4



รูปที่ 3.4 แสดงขั้นตอนในการดำเนินงานวิจัยทั้งหมด

1) การคัดเลือกกลุ่มตัวอย่าง

ในการดำเนินงานวิจัยนี้ มีการคัดเลือกกลุ่มตัวอย่างโดยการสัมภาษณ์กลุ่มตัวอย่างก่อนทำการทดลอง ว่าเคยเรียนเนื้อหาวิชา Data Science หรือไม่ โดยจะคัดเลือกกลุ่มตัวอย่างจากนักศึกษา ระดับปริญญาตรี คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีไทย-ญี่ปุ่น โดยจะคัดเลือกเฉพาะผู้ที่ไม่เคยเรียนวิชา data science เท่านั้น

2) การติดตั้งโปรแกรมที่ใช้ในการทดสอบ

เนื่องจากโปรแกรมที่จัดทำขึ้น ถูกจัดทำด้วยภาษา python และตัวโปรแกรมจะต้องมีการปรับค่า parameter ให้เหมาะสมกับคอมพิวเตอร์แต่ละเครื่อง จึงจำเป็นที่จะต้องลง Python และ module ต่างๆ ที่เกี่ยวข้อง ตลอดจนถึงโปรแกรมที่ใช้ในการแก้ไขและพัฒนาโปรแกรม ดังนี้

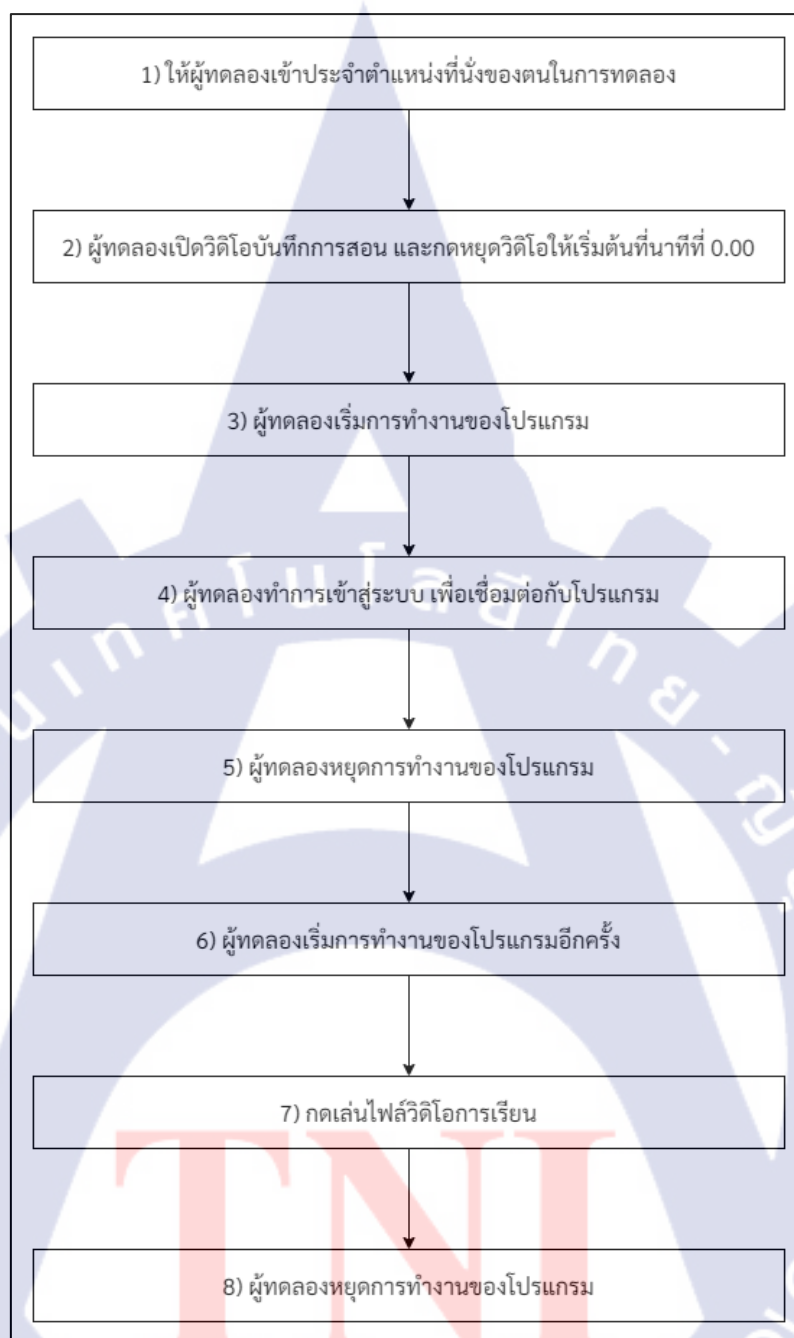
1) Install Python version 3 ขึ้นไป

2) Install Visual Studio code

3) Install python module ได้แก่ opencv-python , keras , tensorflow , pywin32 และ google-api-python-client

โดยค่า parameter ที่จำเป็นต้องปรับให้เหมาะสมกับคอมพิวเตอร์แต่ละเครื่อง คือจำนวนเฟรมภาพต่อวินาที เนื่องจากในการบันทึกวิดีโอ มีกระบวนการ emotion detection ซึ่งจะใช้เวลาไม่เท่ากันในแต่ละเครื่อง ขึ้นอยู่กับ RAM และ GPU ของเครื่องคอมพิวเตอร์ที่ใช้งาน

3) ขั้นตอนในการทดลอง จะประกอบด้วยขั้นตอนย่อยต่างๆ ดังรูปที่ 3.5



รูปที่ 3.5 แสดงขั้นตอนย่อยในการทำการทดลอง

1) ให้กลุ่มตัวอย่างที่คัดเลือกมา นั่งประจำที่เครื่องคอมพิวเตอร์ของตน โดยให้มุมกล้องจับภาพใบหน้าได้อย่างชัดเจน และอยู่ในพื้นที่ที่มีแสงพอเหมาะ ไม่มีดหรือสว่างเกินไป ดังรูปที่ 3.6



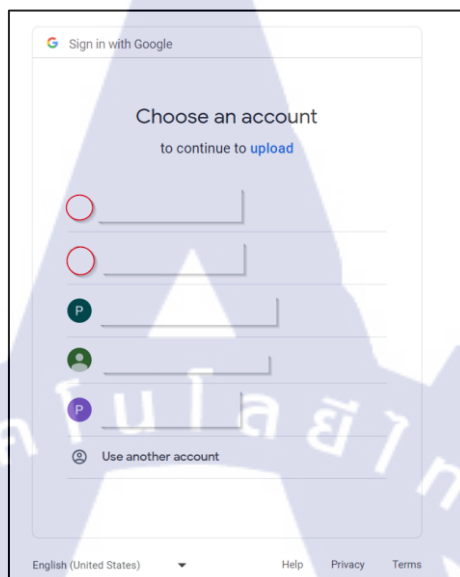
รูปที่ 3.6 แสดงการนั่งประจำจุดการทดลองของผู้ทำการทดลอง

2) เปิดไฟล์วิดีโอในการเรียนเตรียมไว้ แต่ยังไม่กดเล่นวิดีโอ และให้วิดีโอในการเรียนเริ่มต้นที่ช่วง 0.00 นาทีเท่านั้น ดังรูปที่ 3.7



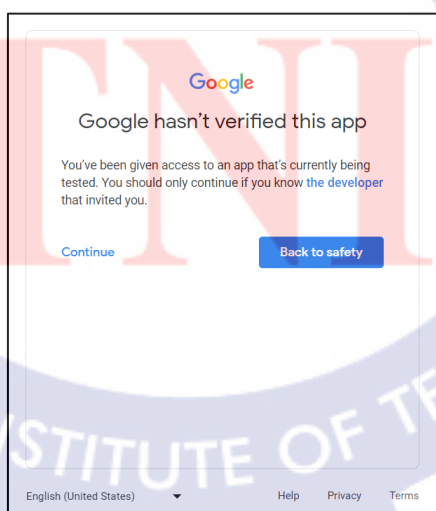
รูปที่ 3.7 แสดงการเปิดวิดีโอเพื่อเตรียมทำการทดลอง

- 3) จากนั้นเริ่มต้นรันโปรแกรม ล็อกอินเข้าสู่ระบบตามลำดับขั้นตอน ตามลำดับ ดังนี้
- เข้าสู่ระบบด้วยบัญชี gmail ดังรูปที่ 3.8



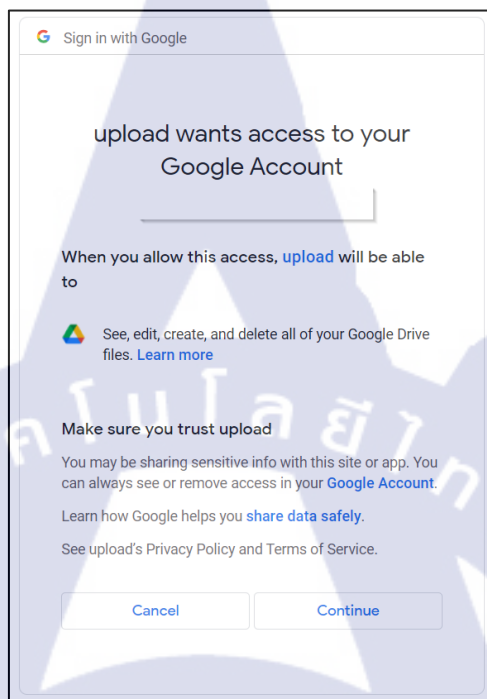
รูปที่ 3.8 แสดงการเข้าสู่ระบบ

- หน้าจอแจ้งเตือนว่าแอปพลิเคชันยังไม่ถูกรับรองด้วย Google โดยขั้นตอนนี้จะต้องกดปุ่ม Continue เพื่อไปยังขั้นตอนต่อไป ดังรูปที่ 3.9



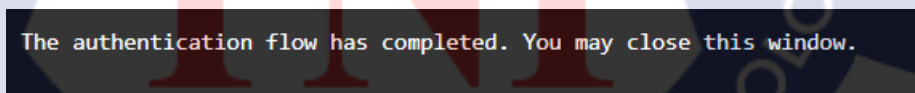
รูปที่ 3.9 หน้าจอแจ้งเตือนจาก Google

- หน้าจอการขอสิทธิ์การเข้าถึง เพื่ออัปโหลดไฟล์บน Google drive โดยในหน้าจอนี้จะต้องกดปุ่ม Continue เช่นกัน ดังรูปที่ 3.10



รูปที่ 3.10 แสดงการขอสิทธิ์ในการอัปโหลดไฟล์

- หลังจากทำตามขั้นตอนข้างต้นทั้งหมด จะขึ้นหน้าจอว่า “The authentication flow has completed. You may close this window.” แสดงถึงการเข้าสู่ระบบและยืนยันตัวตนเพื่อเข้าใช้งานแอปพลิเคชันได้เรียบร้อยแล้ว ดังรูปที่ 3.11



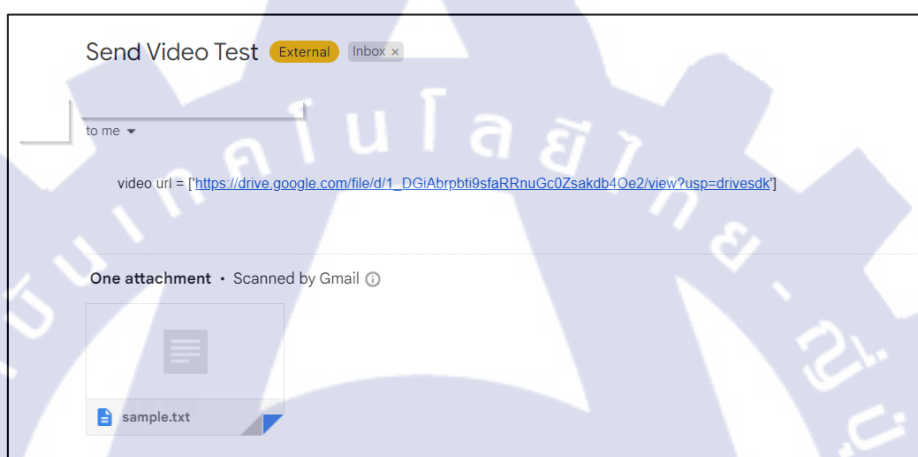
รูปที่ 3.11 หน้าจอแสดงการเข้าสู่ระบบสำเร็จ

- 4) หยุดการทำงานของโปรแกรม โดยการกดปุ่ม q บนแป้นพิมพ์ เพื่อไม่ให้ไฟล์วิดีโอผลลัพธ์ที่ได้มีการบันทึกเวลาที่ผู้ทดลองมองคอมพิวเตอร์ แต่ไม่ได้ทำการเรียน รวมไปถึงช่วงเวลาที่มีการเรียนจริง และตรวจสอบว่าการส่งอีเมลผลลัพธ์มีความถูกต้อง

5) รันโปรแกรมใหม่อีกครั้ง เนื่องจาก session ในการล็อกอินจะสามารถอยู่ได้ 4 ชม. จึงสามารถบันทึกวิดีโอได้ทันทีหลังจากการรันโปรแกรมโดยไม่ต้องล็อกอินใหม่อีกครั้ง

6) กดเล่นไฟล์วิดีโอในการเรียนที่เตรียมไว้ โดยเริ่มที่วินาทีที่ 0.00 เพื่อให้สามารถตรวจจับอารมณ์ของผู้เรียนได้ตั้งแต่ต้นจนถึงจบขั้นตอนการเรียน

7) เมื่อเรียนจบแล้ว กดปุ่ม q เพื่อออกจากโปรแกรม จากนั้นโปรแกรมจะบันทึกวิดีโอ และส่งอีเมล (ตามที่อยู่อีเมลที่ได้กำหนดไว้) โดยอัตโนมัติ โดยอีเมลที่ได้รับเป็นที่อยู่เว็บไซต์ในการจัดเก็บวิดีโอบน Google Drive และไฟล์นามสกุล “.txt” ที่บันทึก timestamp ไว้ ดังรูปที่ 3.12



รูปที่ 3.12 แสดงอีเมลที่ได้รับจากการรันโปรแกรม

4) สรุปผลการทดลอง

จากผลลัพธ์จากการทดลองที่ได้ออกมาเป็นไฟล์ video และไฟล์ timestamp ที่เป็น textfile โดยจะมีชื่อ file เป็น timestamp ในเวลาที่เริ่มต้นรันโปรแกรม โดยไฟล์ผลลัพธ์ที่ได้เป็นไฟล์ timestamp โดยจะระบุเวลา ซึ่งนับตั้งแต่โปรแกรมเริ่มการทำงาน ระบุว่าผู้เรียนมีความสนใจเรียนหรือไม่ และมีวงเล็บอารมณ์ที่สามารถตรวจจับได้อยู่ด้านหลัง หากไม่สามารถตรวจจับภาพใบหน้าได้ จะบันทึกข้อความว่า “Cannot detect the face” แทนข้อความระบุความตั้งใจเรียนและอารมณ์ โดยในการบันทึกข้อความเหล่านี้อยู่ในไฟล์ จะบันทึกเมื่ออารมณ์มีการเปลี่ยนแปลงเท่านั้น เพื่อให้ง่ายในการอ่านและสรุปผลการทดลอง ว่าในแต่ละช่วงเวลาผู้เรียนมีอารมณ์เช่นไรบ้าง และตั้งใจเรียนหรือไม่

บทที่ 4

ผลการวิจัย

- 4.1 ผลการตรวจสอบความถูกต้องของโมเดล
- 4.2 ผลการพัฒนาซอฟต์แวร์ในการตรวจจับใบหน้าและอารมณ์ของผู้เรียน
- 4.3 ผลการประเมินประสิทธิภาพซอฟต์แวร์ในการตรวจจับใบหน้าและอารมณ์ของผู้เรียน

4.1 ผลการตรวจสอบความถูกต้องของโมเดล

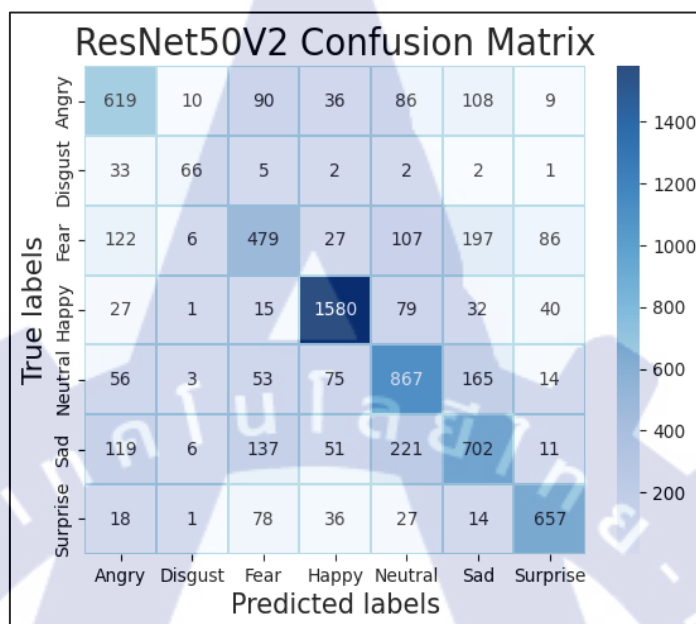
ในการวิเคราะห์อารมณ์ของผู้เรียนในระหว่างการเรียนออนไลน์ เราได้ใช้โมเดลที่ถูกเทรนจาก ResNet50 ซึ่งเป็น pre-trained model ที่ถูกสร้างขึ้นจาก Convolutional Neural Network (CNN) โดยใช้ชุดข้อมูล FER-2013 ซึ่งแบ่งอารมณ์เป็น 7 กลุ่มอารมณ์ ได้แก่ อารมณ์โกรธ (Angry), อารมณ์ขยะแขยง (Disgust), อารมณ์กลัว (Fear), มีความสุข (Happy), อารมณ์เป็นกลาง (Neutral), ความเศร้า (Sad), ประหลาดใจ (Surprise) ซึ่งเป็นการทดสอบการจำแนกอารมณ์บนใบหน้าจากโมเดลที่สร้างขึ้น โดยใช้กลุ่มตัวอย่างเป็นชุดข้อมูล FER-2013 ส่วนที่แบ่งไว้ทำการทดสอบจำนวน 30% และได้ผลลัพธ์ดังรูปที่ 4.1



รูปที่ 4.1 ตัวอย่างการจำแนกกลุ่มอารมณ์ด้วยโมเดล

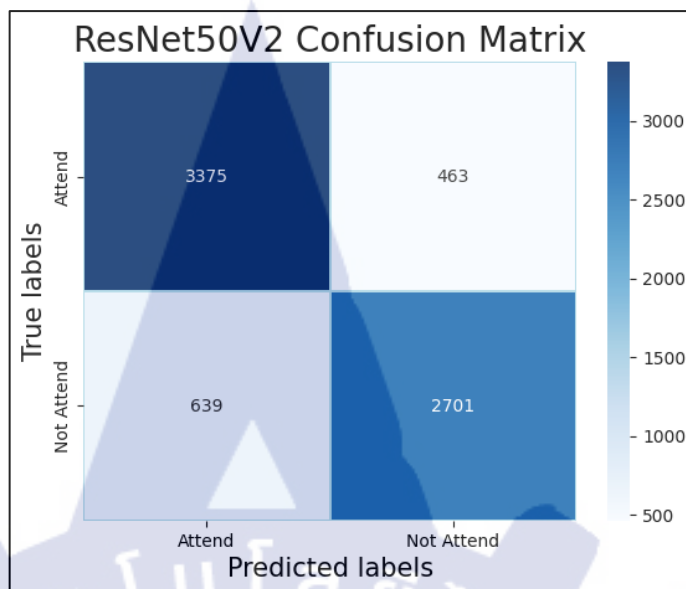
การตรวจสอบความถูกต้องของโมเดลนี้ ถูกวัดผลโดยการนำข้อมูลจากชุดข้อมูล FER-2013 จำนวน 30% ที่ถูกแบ่งไว้โดยการสุ่มจากชุดข้อมูล เพื่อการทดสอบประสิทธิภาพของโมเดล โดยโมเดล

มีค่าความถูกต้องดังนี้ Accuracy = 0.692 , Precision = 0.690 , Recall = 0.692 , F1 score = 0.690 และมี Confusion Matrix ดังรูปที่ 4.2



รูปที่ 4.2 Confusion Matrix ของโมเดลแบ่ง 7 กลุ่มอารมณ์

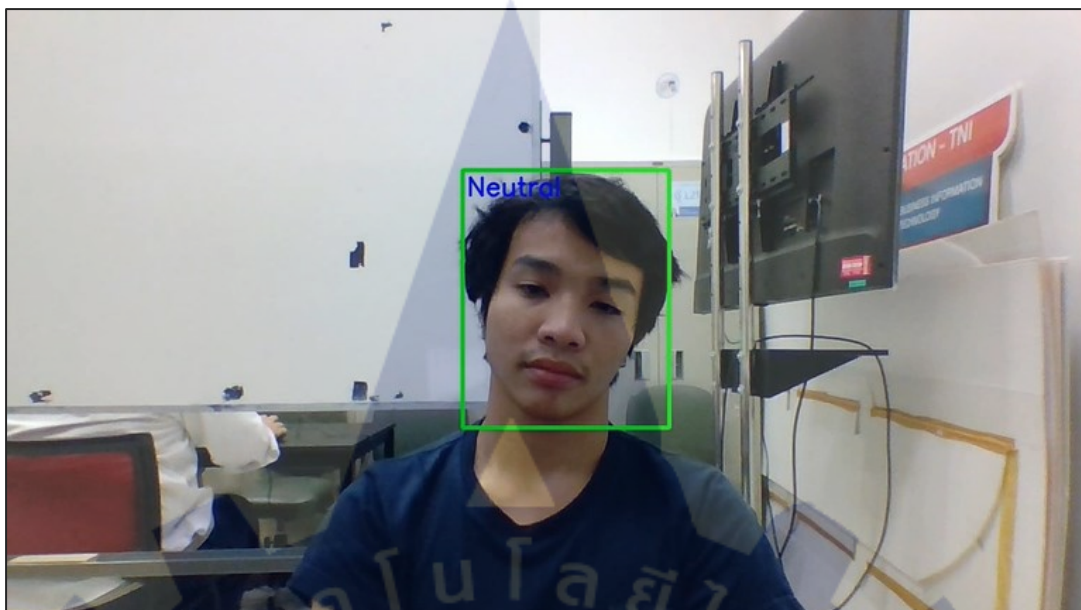
จาก 7 อารมณ์ที่กล่าวมาข้างต้น สามารถจำแนกอารมณ์ในการเรียนออกเป็น 2 กลุ่ม คือ 1) กลุ่มที่ตั้งใจเรียน (Attend) และ 2) กลุ่มที่ไม่ตั้งใจเรียน (Not Attend) โดยกลุ่มที่ตั้งใจเรียน (Attend) ประกอบไปด้วย 1) มีความสุข (Happy), อารมณ์เป็นกลาง (Neutral), ประหลาดใจ (Surprise) และกลุ่มที่ไม่ตั้งใจเรียน (Not Attend) ประกอบด้วยอารมณ์โกรธ (Angry), อารมณ์ขยะแขยง (Disgust), อารมณ์กลัว (Fear), ความเศร้า (Sad) โดยมีค่าความถูกต้องดังนี้ Accuracy: 0.846, Precision: 0.879, Recall: 0.841, F1 score: 0.860 และได้ confusion matrix ของโมเดลสองกลุ่มอารมณ์ดังรูปที่ 4.3



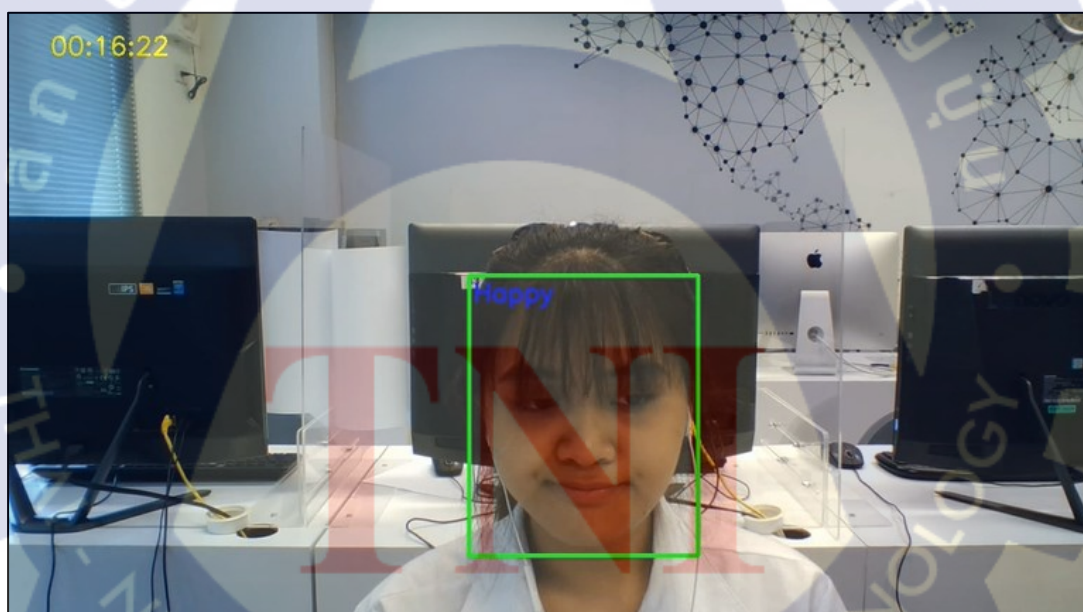
รูปที่ 4.3 Confusion Matrix ของโมเดลแบ่ง 2 กลุ่มอารมณ์

4.2 ผลการพัฒนาซอฟต์แวร์ในการตรวจจับใบหน้าและอารมณ์ของผู้เรียน

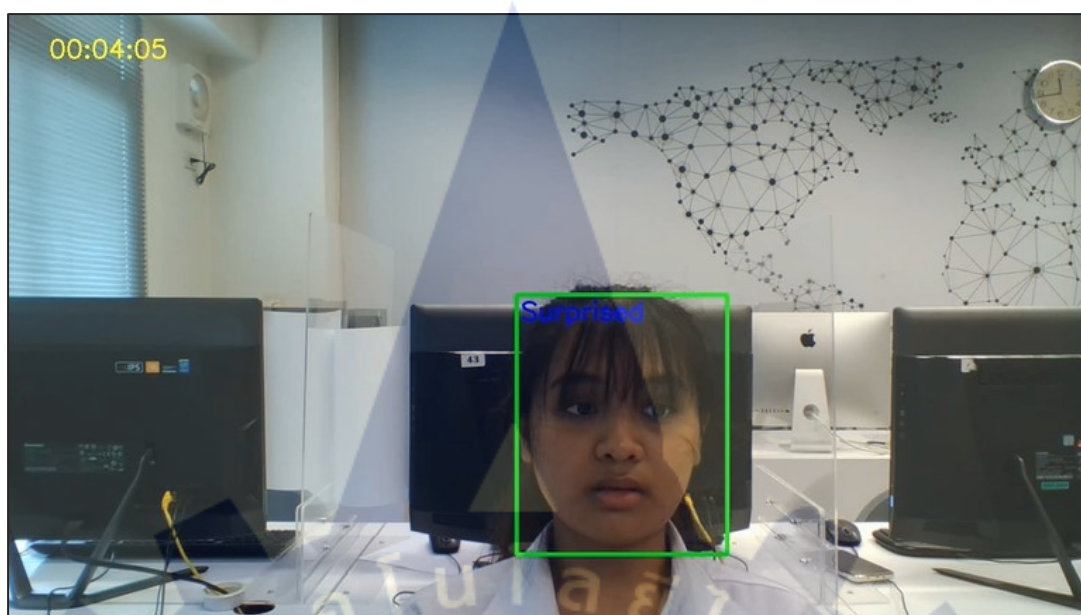
ผลการพัฒนาซอฟต์แวร์ในการตรวจจับใบหน้าและอารมณ์ของผู้เรียน ได้ทำการทดลองโดยใช้เนื้อหาของวิชา data science พื้นฐาน ซึ่งมีความยาว 20.00 นาที และแบ่งเนื้อหาเป็น 6 ส่วน ดังที่กล่าวมาข้างต้นแล้ว โดยในการตรวจสอบอารมณ์ในระหว่างการเรียนรู้ จะได้ผลลัพธ์เป็นวิดีโอที่แสดงใบหน้าของผู้เรียน โดยมีสีเหลี่ยมติกรอบล้อมรอบใบหน้าของผู้เรียน พร้อมแสดงอารมณ์ที่สามารถตรวจจับได้ที่ด้านมุมบนซ้ายของกรอบล้อมรอบใบหน้า ดังรูปที่ 4.4 – 4.10



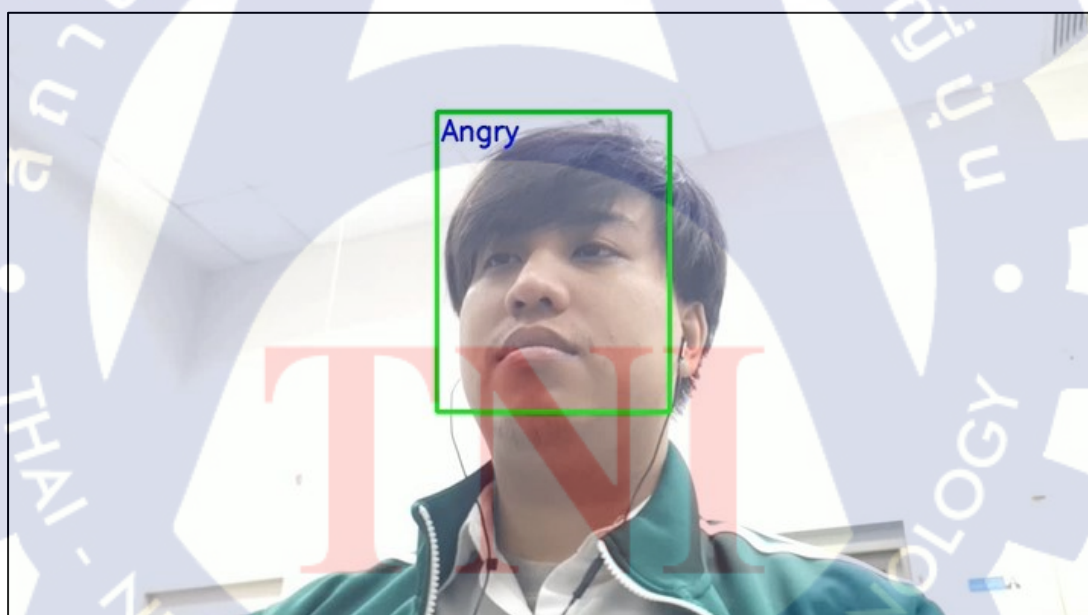
รูปที่ 4.4 แสดงภาพการตรวจจับใบหน้าในอารมณ์เป็นกลาง (Neutral)



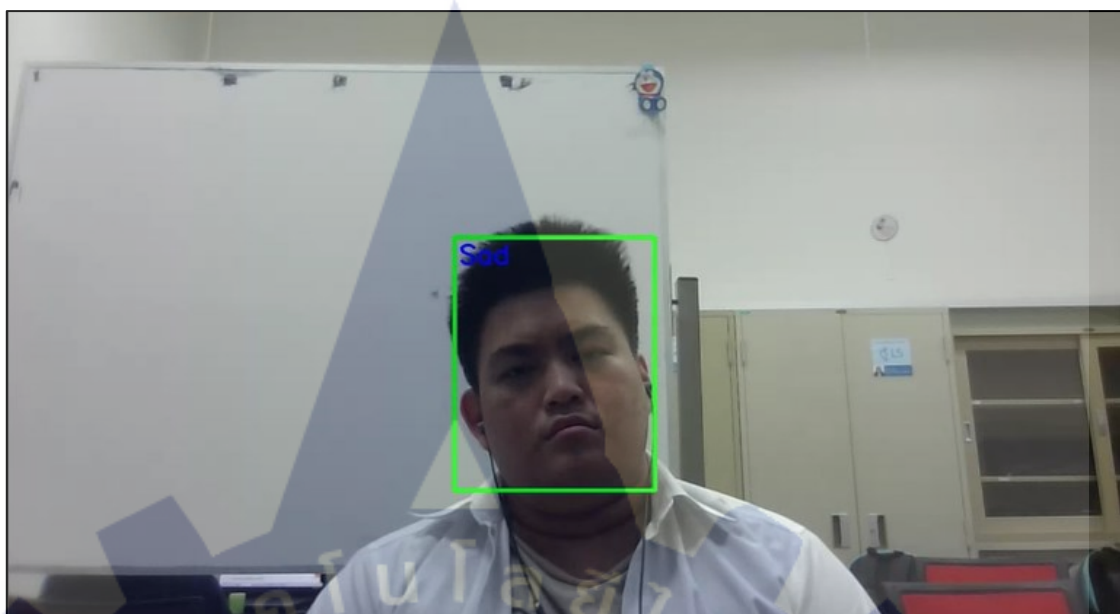
รูปที่ 4.5 แสดงภาพการตรวจจับใบหน้าในอารมณ์มีความสุข (Happy)



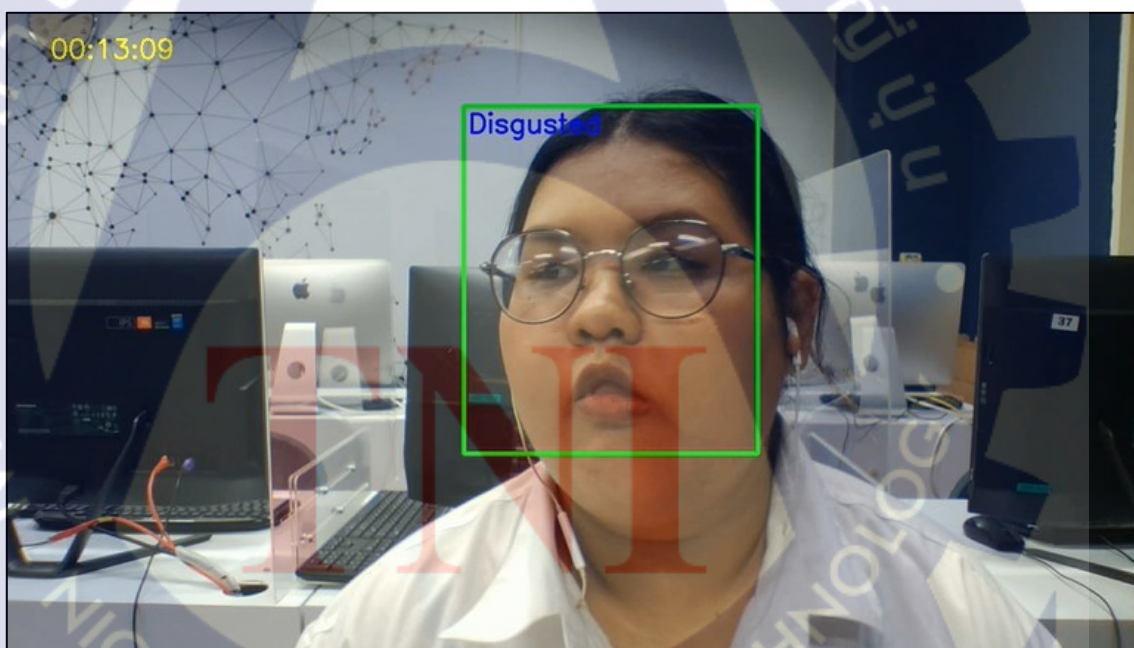
รูปที่ 4.6 แสดงภาพการตรวจจับใบหน้าในอารมณ์ประหลาดใจ (Surprised)



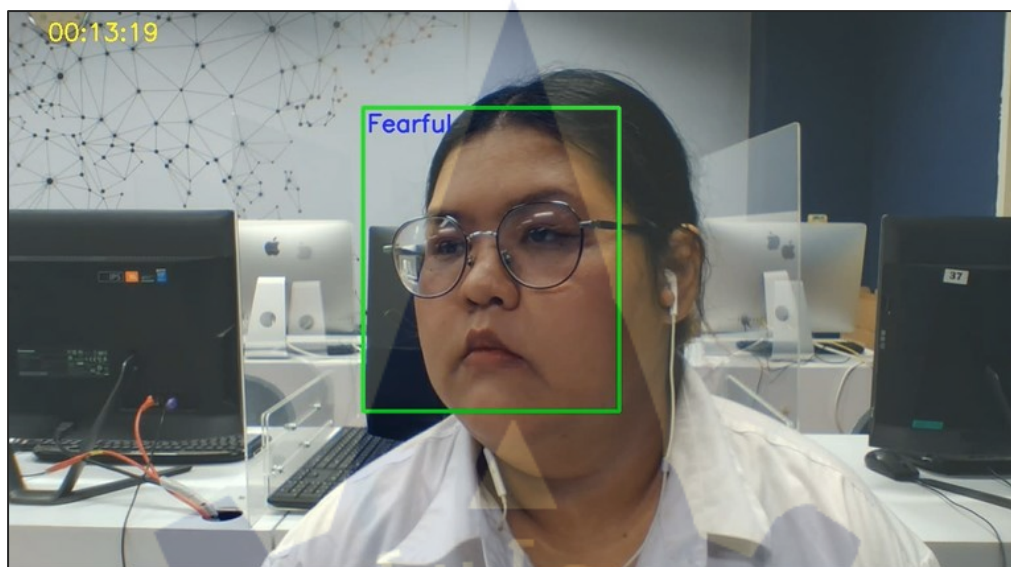
รูปที่ 4.7 แสดงภาพการตรวจจับใบหน้าในอารมณ์โกรธ (Angry)



รูปที่ 4.8 แสดงภาพการตรวจจับใบหน้าในอารมณ์เศร้า (Sad)

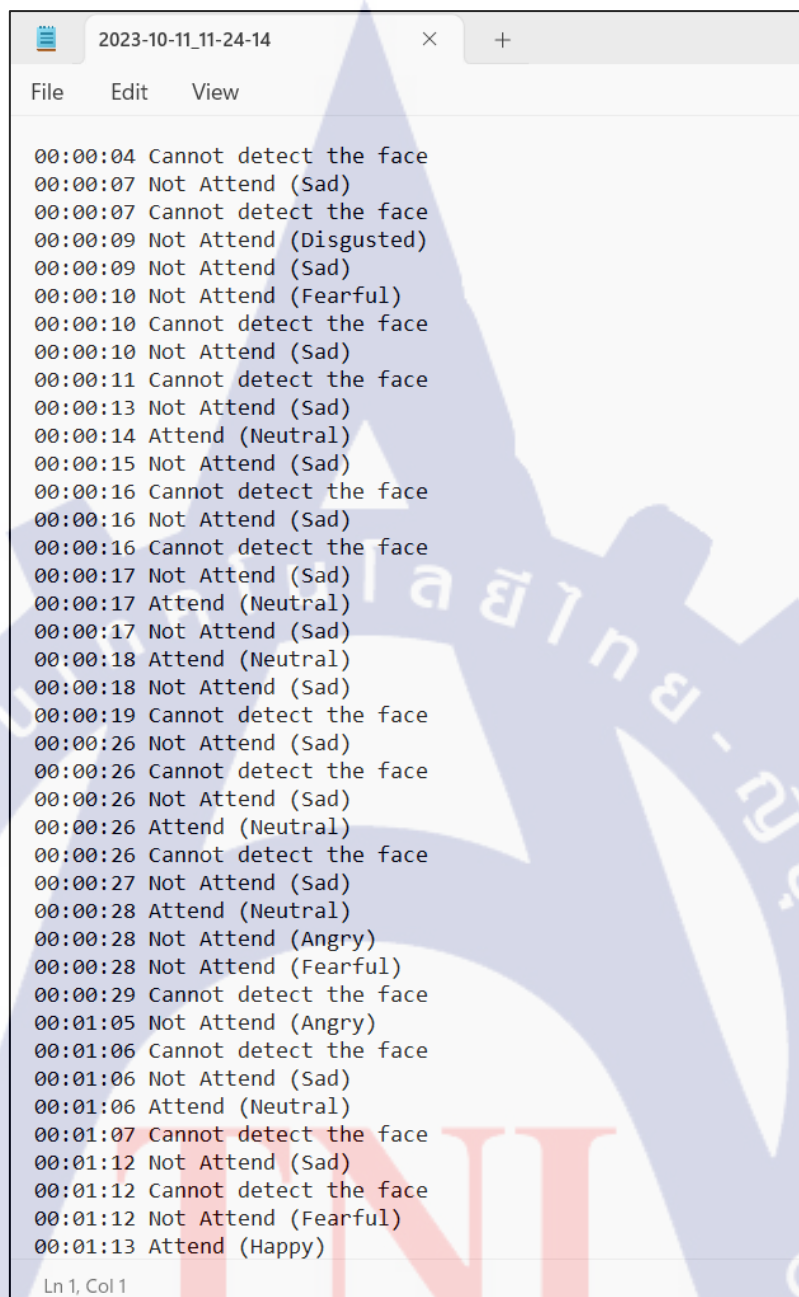


รูปที่ 4.9 แสดงภาพการตรวจจับใบหน้าในอารมณ์เศร้า (Sad)



รูปที่ 4.10 แสดงภาพการตรวจจับใบหน้าในอารมณ์กลัว (Fear)

และในขณะเดียวกันกับที่มีการบันทึกวิดีโอ ระบบจะสร้างไฟล์ text ขึ้นมา และบันทึกช่วงเวลาที่คุณเรียนมีการเปลี่ยนแปลงอารมณ์ลงในไฟล์ด้วย ดังรูปที่ 4.11



```

2023-10-11_11-24-14
File Edit View

00:00:04 Cannot detect the face
00:00:07 Not Attend (Sad)
00:00:07 Cannot detect the face
00:00:09 Not Attend (Disgusted)
00:00:09 Not Attend (Sad)
00:00:10 Not Attend (Fearful)
00:00:10 Cannot detect the face
00:00:10 Not Attend (Sad)
00:00:11 Cannot detect the face
00:00:13 Not Attend (Sad)
00:00:14 Attend (Neutral)
00:00:15 Not Attend (Sad)
00:00:16 Cannot detect the face
00:00:16 Not Attend (Sad)
00:00:16 Cannot detect the face
00:00:17 Not Attend (Sad)
00:00:17 Attend (Neutral)
00:00:17 Not Attend (Sad)
00:00:18 Attend (Neutral)
00:00:18 Not Attend (Sad)
00:00:19 Cannot detect the face
00:00:26 Not Attend (Sad)
00:00:26 Cannot detect the face
00:00:26 Not Attend (Sad)
00:00:26 Attend (Neutral)
00:00:26 Cannot detect the face
00:00:27 Not Attend (Sad)
00:00:28 Attend (Neutral)
00:00:28 Not Attend (Angry)
00:00:28 Not Attend (Fearful)
00:00:29 Cannot detect the face
00:01:05 Not Attend (Angry)
00:01:06 Cannot detect the face
00:01:06 Not Attend (Sad)
00:01:06 Attend (Neutral)
00:01:07 Cannot detect the face
00:01:12 Not Attend (Sad)
00:01:12 Cannot detect the face
00:01:12 Not Attend (Fearful)
00:01:13 Attend (Happy)

Ln 1, Col 1

```

รูปที่ 4.11 แสดงไฟล์ text ที่บันทึกการเปลี่ยนแปลงของอารมณ์

4.3 ผลการประเมินประสิทธิภาพซอฟต์แวร์ในการตรวจจับใบหน้าและอารมณ์ของผู้เรียน

การประเมินประสิทธิภาพซอฟต์แวร์ในการตรวจจับใบหน้าและอารมณ์ของผู้เรียน จะมีการประเมินจากผลลัพธ์ของวิดีโอและไฟล์ text ที่เป็นผลลัพธ์ในการทดสอบการทำงานของซอฟต์แวร์ในข้างต้น จึงทำการสรุปผลการประเมินประสิทธิภาพของซอฟต์แวร์แยกออกเป็น 3 หัวข้อ คือ 1) ผลการตรวจจับใบหน้าที่แสดงจำนวนวินาทีที่สามารถตรวจจับใบหน้าได้ และตารางที่ 4.2

แสดงจำนวนเปอร์เซ็นต์ที่สามารถตรวจจับใบหน้าได้ 2) ผลการตรวจสอบอารมณ์ของผู้เรียนที่ตั้งใจเรียน ในตารางที่ 4.3 แสดงจำนวนวินาทีที่ตรวจจับใบหน้าที่ตั้งใจเรียน และตารางที่ 4.4 แสดงจำนวนเปอร์เซ็นต์ที่ตรวจจับใบหน้าที่ตั้งใจเรียน 3) ผลการตรวจสอบอารมณ์ของผู้เรียนที่ไม่ตั้งใจเรียน ในตารางที่ 4.3 แสดงจำนวนวินาทีที่ตรวจจับใบหน้าที่ไม่ตั้งใจเรียน และตารางที่ 4.4 แสดงจำนวนเปอร์เซ็นต์ที่ตรวจจับใบหน้าที่ไม่ตั้งใจเรียน

ตารางที่ 4.1 จำนวนวินาทีที่สามารถตรวจจับใบหน้าของผู้เรียนได้

ผู้เรียน คนที่	เนื้อหา ส่วนที่ 1	เนื้อหา ส่วนที่ 2	เนื้อหา ส่วนที่ 3	เนื้อหา ส่วนที่ 4	เนื้อหา ส่วนที่ 5	เนื้อหา ส่วนที่ 6
1	0	0	0	0	11	3
2	180	344	164	262	209	38
3	149	337	156	257	204	38
4	177	104	44	232	176	34
5	144	279	117	254	207	38
6	180	344	164	261	208	38
7	23	153	31	150	22	13
8	97	59.5	102	112	54	24
9	13	149	162	256	205	37
10	180	343	163	262	208	38
11	155	262	146	165	168	27
12	153	152	101	156	82	7
13	124	32	34	121	39	7
14	33	40	7	15	37	5
15	135	299	156	257	206	38
16	127	92	150	67	85	27
ค่าเฉลี่ย	116.88	186.78	106.06	176.69	132.56	25.75

ตารางที่ 4.2 เปอร์เซ็นต์ที่สามารถตรวจจับใบหน้าของผู้เรียนได้

ผู้เรียน คนที่	เนื้อหา ส่วนที่ 1	เนื้อหา ส่วนที่ 2	เนื้อหา ส่วนที่ 3	เนื้อหา ส่วนที่ 4	เนื้อหา ส่วนที่ 5	เนื้อหา ส่วนที่ 6
1	0%	0%	0%	0%	5.26%	8.57%
2	100%	100%	100%	100%	100%	100%
3	82.78%	97.96%	95.12%	98.09%	97.60%	100%
4	98.33%	30.23%	26.83%	88.54%	84.21%	89.47%
5	80.00%	81.10%	71.34%	96.94%	89.47%	86.84%
6	100%	100%	100%	99.07%	99.04%	100%
7	12.66%	44.47%	18.90%	5.72%	10.52%	34.21%
8	53.89%	41.27%	66.46%	42.74%	25.83%	63.15%
9	7.22%	43.31%	98.78%	97.71%	98.09%	97.37%
10	100.00%	99.42%	99.39%	100.00%	99.52%	100.00%
11	86.11%	76.16%	89.02%	62.98%	80.38%	71.05%
12	85.00%	44.19%	61.59%	59.54%	39.23%	18.42%
13	68.89%	9.30%	20.73%	46.18%	18.66%	18.42%
14	18.33%	11.63%	4.27%	5.73%	17.70%	13.16%
15	75.00%	86.92%	95.12%	98.09%	98.56%	100.00%
16	70.56%	26.74%	91.46%	25.57%	40.67%	71.05%
ค่าเฉลี่ย	64.92%	55.79%	64.94%	64.18%	62.80%	66.98%

จากตารางที่ 4.1 ซึ่งแสดงจำนวนวินาทีที่สามารถตรวจจับใบหน้าได้ และตารางที่ 4.2 ซึ่งแสดงการคำนวณเปอร์เซ็นต์ของเวลาที่ระบบสามารถตรวจจับใบหน้าของผู้เรียน ที่สามารถคำนวณได้ โดยนำ จำนวนวินาทีที่ระบบตรวจจับใบหน้าได้สำเร็จหารด้วยจำนวนวินาทีทั้งหมดในส่วนนั้นๆ โดยจะได้ผลลัพธ์เป็นเปอร์เซ็นต์ของเวลาที่ระบบสามารถตรวจจับใบหน้าของผู้เรียนได้ ดังสมการที่ 4.1

เปอร์เซ็นต์ที่สามารถตรวจจับใบหน้าของผู้เรียน

$$= \frac{\text{เวลาสามารถตรวจจับใบหน้าได้ (วินาที)}}{\text{ความยาวของวิดีโอในช่วงนั้นๆ (วินาที)}} \times 100 \quad (4.1)$$

จากผลการทดลอง จะเห็นได้ว่าผู้เรียนคนที่ 1 สามารถตรวจจับใบหน้าได้น้อยที่สุด ซึ่งในการเรียนเนื้อหาส่วนที่ 1-4 จะไม่สามารถตรวจจับใบหน้าได้เลย และตรวจจับใบหน้าในเนื้อหาส่วนที่ 5 ได้เพียง 11 วินาทีจากระยะเวลาทั้งหมด 3.27 นาที เมื่อพิจารณาจากตารางที่ 4.1 ของเวลาที่ระบบสามารถจับใบหน้าได้ ซึ่งคำนวณได้เท่ากับ 5.26% ซึ่งแสดงในตารางที่ 4.2 และในเนื้อหาส่วนที่ 6 ระบบสามารถตรวจจับใบหน้าได้เพียง 3 วินาทีจากทั้งหมด 38 วินาที ดังตารางที่ 4.1 หรือ 8.57% ซึ่งแสดงในตารางที่ 4.2 โดยเมื่อเปรียบเทียบกับค่าเฉลี่ยแล้ว ผู้เรียนคนที่ 1 มีอัตราการตรวจจับใบหน้าได้ต่ำกว่าค่าเฉลี่ยมาก

ผู้เรียนคนที่ 2 สามารถตรวจจับใบหน้าได้ทั้งหมด 100% ตลอดช่วงเวลาในการเรียน ซึ่งเป็นผู้เรียนเพียงคนเดียวที่สามารถตรวจจับใบหน้าได้ครบ 100%

ผู้เรียนคนที่ 3 สามารถตรวจจับใบหน้าได้มาก โดยในเนื้อหาส่วนที่ 1 มีอัตราการตรวจจับใบหน้าได้ต่ำที่สุด คือ 82.78% ซึ่งเป็นอัตราการตรวจจับใบหน้าที่ยังอยู่ในระดับสูง เมื่อเปรียบเทียบกับค่าเฉลี่ยของการตรวจจับใบหน้าในเนื้อหาการเรียนช่วงที่ 1 ซึ่งมีค่าเฉลี่ยเท่ากับ 64.92% โดยในเนื้อหาส่วนที่ 2 เป็นต้นไป สามารถตรวจจับใบหน้าของผู้เรียนได้มากกว่า 90% ทั้งหมด

ผู้เรียนคนที่ 4 สามารถตรวจจับใบหน้าได้มากในเนื้อหาส่วนที่ 1, 4, 5 และ 6 โดยในช่วงการเรียนเนื้อหาทั้ง 4 ส่วนดังกล่าว สามารถตรวจจับใบหน้าได้มากกว่า 80% แต่ในช่วงเนื้อหาส่วนที่ 2 และ 3 กลับตรวจจับใบหน้าได้น้อยมาก โดยส่วนที่ 2 สามารถตรวจจับใบหน้าได้เพียง 30.23% และส่วนที่ 3 ตรวจจับได้เพียง 26.83% เท่านั้น

ผู้เรียนคนที่ 5 สามารถตรวจจับภาพใบหน้าได้มาก โดยผู้เรียนคนที่ 5 มีอัตราการตรวจจับใบหน้าคล้ายคลึงกับผู้เรียนคนที่ 3 คือ สามารถตรวจจับภาพใบหน้าได้มาก และมีอัตราการตรวจจับใบหน้าได้สูง โดยมีจุดที่แตกต่างกันคือ ผู้เรียนคนที่ 5 ตรวจจับใบหน้าได้น้อยที่สุดในเนื้อหาส่วนที่ 3 โดยตรวจจับได้ 71.34% แต่ในเนื้อหาส่วนอื่นๆ สามารถตรวจจับใบหน้าได้มากกว่า 80%

ผู้เรียนคนที่ 6 คือผู้เรียนที่มีอัตราการตรวจจับใบหน้าสูงเป็นอันดับที่ 2 โดยมีเพียงส่วนที่ 4 และส่วนที่ 5 เท่านั้น ที่ตรวจจับภาพใบหน้าได้ไม่ครบ 100% แต่ยังมีอัตราการตรวจจับใบหน้าได้สูงมาก คือ 99.07% และ 99.04% ตามลำดับ

ผู้เรียนคนที่ 7 มีอัตราการตรวจจับใบหน้าได้ต่ำกว่า 20% ในเนื้อหาส่วนที่ 1, 3, 4, และ 5 โดยเฉลี่ย อัตราการตรวจจับใบหน้าของระบบในเนื้อหาส่วนต่าง ๆ คือ: เนื้อหาส่วนที่ 1 ตรวจจับได้ 12.66% หรือ 23 วินาที , เนื้อหาส่วนที่ 2 ตรวจจับได้ 44.47% หรือ 153 วินาที , เนื้อหาส่วนที่ 3 ตรวจจับได้ 18.90% หรือ 31 วินาที , เนื้อหาส่วนที่ 4 ตรวจจับได้ 5.72% หรือ 150 วินาที , เนื้อหาส่วนที่ 5 ตรวจจับได้ 10.52% หรือ 22 วินาที และเนื้อหาส่วนที่ 6 ตรวจจับได้ 34.21% หรือ 13 วินาที

ผู้เรียนคนที่ 8 มีอัตราการตรวจจับใบหน้าในระดับค่อนข้างต่ำ โดยเมื่อเปรียบเทียบกับค่าเฉลี่ยแล้ว ผู้เรียนมีอัตราการตรวจจับใบหน้าได้สูงกว่าค่าเฉลี่ยเพียงช่วงเดียว คือ ช่วงที่ 3 สามารถ

ตรวจจับใบหน้าได้ 66.46% เมื่อเปรียบเทียบกับค่าเฉลี่ยในช่วงที่ 3 ที่มีค่าเฉลี่ยในการตรวจจับใบหน้าเท่ากับ 64.94%

ผู้เรียนคนที่ 9 มีอัตราการตรวจจับใบหน้าในช่วงที่ 1 ต่ำมาก คือ 7.22% หรือเพียง 13 วินาทีเท่านั้น และค่อยๆ เพิ่มขึ้นในช่วงที่ 2 คือ 43.31% ซึ่งเท่ากับ 149 วินาที และหลังจากนั้นจึงค่อยๆ เพิ่มขึ้นจนมากกว่า 90% ในช่วงที่ 3, 4, 5, และ 6

ผู้เรียนคนที่ 10 เป็นผู้เรียนที่สามารถตรวจจับใบหน้าได้มากเป็นอันดับที่ 3 โดยสามารถตรวจจับใบหน้าได้ 100% ในเนื้อหาส่วนที่ 1, 4 และ 6 โดยเนื้อหาอีก 3 ส่วน ที่น้อยกว่า 100% นั้น ได้แก่เนื้อหาส่วนที่ 2 สามารถตรวจจับได้ 99.42% เนื้อหาส่วนที่ 3 สามารถตรวจจับได้ 99.39% และเนื้อหาส่วนที่ 5 ที่สามารถตรวจจับใบหน้าได้ 99.52%

ผู้เรียนคนที่ 11 มีอัตราการตรวจจับใบหน้าได้ค่อนข้างสูง ยกเว้นในเนื้อหาส่วนที่ 4 ซึ่งเป็นช่วงเดียวที่มีอัตราตรวจจับใบหน้าได้ต่ำกว่าค่าเฉลี่ย คือ 62.98% ในขณะที่ค่าเฉลี่ยในการตรวจจับใบหน้าของช่วงที่ 4 คือ 64.18%

ผู้เรียนคนที่ 12 ในช่วงแรกมีอัตราการตรวจจับใบหน้าได้มาก คือ 85.00% ในช่วงที่ 1 และลดลงในช่วงที่ 2 เหลือเพียง 44.19% หลังจากนั้นเพิ่มขึ้นอีกครั้งในช่วงที่ 3 เป็น 61.59% และหลังจากนั้นลดลงอย่างต่อเนื่อง จนเหลือ 59.54% ในช่วงที่ 4 , 38.23% ในช่วงที่ 5 และ 18.42% ในช่วงที่ 6

เช่นเดียวกับผู้เรียนคนที่ 13 ซึ่งมีแนวโน้มเช่นเดียวกัน คือตรวจจับใบหน้าในช่วงที่ 1 ได้มาก และลดลงในช่วงที่ 2 เพิ่มขึ้นอีกครั้งในช่วงที่ 3 แต่มีความแตกต่างคือในช่วงที่ 4 อัตราการตรวจจับใบหน้ายังเพิ่มขึ้นอยู่ และลดลงในช่วงที่ 5 และ 6

ผู้เรียนคนที่ 14 มีแนวโน้มการตรวจจับใบหน้าได้ต่ำตลอดทั้งการเรียน โดยอัตราการตรวจจับใบหน้าทั้งหมดไม่เกิน 20% โดยตรวจจับใบหน้าได้มากที่สุดที่เนื้อหาช่วงที่ 1 คือ 18.33% และลดลงในช่วงที่ 2 เหลือเพียง 11.63% และลดลงอีกเป็นอย่างมากในช่วงที่ 3 จนเหลือเพียง 4.27% และเพิ่มขึ้นเล็กน้อยเป็น 5.73% ในเนื้อหาส่วนที่ 4 จากนั้นเพิ่มขึ้นอีกครั้งจนถึง 17.70% ในช่วงที่ 5 และลดลงจนเหลือ 13.16% ในเนื้อหาส่วนที่ 6

ผู้เรียนคนที่ 15 มีอัตราการตรวจจับใบหน้าสูง โดยสามารถตรวจจับใบหน้าได้น้อยที่สุดในเนื้อหาส่วนที่ 1 โดยสามารถตรวจจับได้ 75.00% และค่อยๆ เพิ่มมากขึ้นเป็นลำดับ โดยสามารถตรวจจับใบหน้าได้ 86.92% ในเนื้อหาส่วนที่ 2 , 95.12% ในเนื้อหาส่วนที่ 3 , 98.09% ในเนื้อหาส่วนที่ 4 , 98.56% ในเนื้อหาส่วนที่ 5 และสามารถตรวจจับใบหน้าได้ 100% ในเนื้อหาส่วนที่ 6

ผู้เรียนคนที่ 16 มีอัตราการตรวจจับใบหน้าที่ไม่สม่ำเสมอ โดยมีทั้งช่วงที่สามารถตรวจจับใบหน้าได้ดี ซึ่งสามารถตรวจจับใบหน้าได้มากกว่าค่าเฉลี่ยอยู่ 3 ช่วง คือช่วงที่ 1 ที่สามารถตรวจจับใบหน้าได้ 70.56% , ช่วงที่ 3 ที่สามารถตรวจจับใบหน้าได้ 91.46% , ช่วงที่ 6 ที่สามารถตรวจจับใบหน้าได้ 71.05%

จากตารางและบทสรุปข้างต้น จะเห็นได้ว่าผู้เรียนจะมีอัตราการตรวจจับใบหน้าที่สม่ำเสมอและไม่สม่ำเสมอในจำนวนใกล้เคียงกัน โดยผู้เรียนที่มีอัตราการตรวจจับใบหน้าที่ได้สูงกว่าค่าเฉลี่ยอย่างสม่ำเสมอ ได้แก่ ผู้เรียนคนที่ 2, 3, 5, 6, 10, 11, 15 ผู้เรียนที่มีอัตราการตรวจจับใบหน้าที่ต่ำกว่าค่าเฉลี่ยอย่างสม่ำเสมอ ได้แก่ ผู้เรียนคนที่ 1, 7 และผู้เรียนที่มีอัตราการตรวจจับใบหน้าที่ไม่สม่ำเสมอ ได้แก่ ผู้เรียนคนที่ 4, 8, 9, 12, 13, 14, 16

ตารางที่ 4.3 จำนวนวินาทีที่สามารถตรวจจับใบหน้าที่โมเดลตรวจสอบว่าตั้งใจเรียนของผู้เรียน

ผู้เรียน คนที่	เนื้อหา ส่วนที่ 1	เนื้อหา ส่วนที่ 2	เนื้อหา ส่วนที่ 3	เนื้อหา ส่วนที่ 4	เนื้อหา ส่วนที่ 5	เนื้อหา ส่วนที่ 6
1	-	-	-	-	10.5	2.5
2	180	344	164	262	209	38
3	148	335	149	251	197	38
4	167	52	18	114.5	34.5	24
5	135	218	67	125	114	27
6	180	344	164	261	208	38
7	3	5	20	82.3	14	8
8	87	42.7	73	78	5	16
9	6	124	162	251	202	35
10	180	342	163	262	208	38
11	155	260	146	164	164	23
12	152	151	101	155	82	7
13	117	30	31	117	39	7
14	33	40	7	15	37	5
15	118	291	155	256	206	37
16	103	61	129	58	64	0
ค่าเฉลี่ย	117.6	176.05	103.27	163.46	112.13	21.47

ตารางที่ 4.4 เปอร์เซ็นต์ความตั้งใจเรียนตามทีโมเดลตรวจสอบได้ของผู้เรียน

ผู้เรียน คนที่	เนื้อหา ส่วนที่ 1	เนื้อหา ส่วนที่ 2	เนื้อหา ส่วนที่ 3	เนื้อหา ส่วนที่ 4	เนื้อหา ส่วนที่ 5	เนื้อหา ส่วนที่ 6
1	-	-	-	-	94.7%	78.9%
2	100%	100%	100%	100%	100%	100%
3	97.9%	95%	95.6%	97.67%	96.6%	100%
4	94.36%	49.86%	40.63%	39.24%	19.67%	89.9%
5	93.75%	78.14%	57.26%	49.32%	55.23%	71.23%
6	100%	100%	100%	100%	100%	100%
7	13.04%	3.37%	54.83%	73.33%	63.63%	61.53%
8	89.69%	71.83%	71.55%	69.64%	9.25%	66.67%
9	46.15%	83.22%	100.00%	98.05%	98.54%	94.59%
10	100.00%	99.71%	100.00%	100.00%	100.00%	100.00%
11	100.00%	99.24%	100.00%	99.39%	97.62%	85.19%
12	99.35%	99.34%	100.00%	99.36%	100.00%	100.00%
13	94.35%	93.75%	91.18%	96.69%	100.00%	100.00%
14	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
15	87.41%	97.32%	99.36%	99.61%	100.00%	97.37%
16	81.10%	66.30%	86.00%	86.57%	75.29%	0.00%
ค่าเฉลี่ย	86.47%	82.51%	86.43%	87.26%	81.91%	84.09%

จากตารางที่ 4.3 ที่แสดงจำนวนวินาทีที่ผู้เรียนตั้งใจเรียน และตารางที่ 4.4 ซึ่งแสดงการคำนวณเปอร์เซ็นต์ของเวลาที่ผู้เรียนให้ความสนใจในการเรียน ซึ่งคำนวณโดยการนำจำนวนวินาทีที่โมเดลสามารถตรวจจับได้ว่าตั้งใจเรียน ทหารด้วยวินาทีที่สามารถตรวจจับใบหน้าได้ทั้งหมด โดยจะคำนวณผลลัพธ์ออกมาเป็นจำนวนเปอร์เซ็นต์ของเวลาที่ผู้เรียนมีส่วนร่วมในการเรียน ดังสมการที่ 4.2

$$\text{เปอร์เซ็นต์ที่ผู้เรียนตั้งใจเรียน} = \frac{\text{เวลาที่ผู้เรียนตั้งใจเรียน (วินาที)}}{\text{เวลาที่สามารถตรวจจับใบหน้าได้ (วินาที)}} \times 100 \quad (4.2)$$

เนื่องจากผู้เรียนคนที่ 1 เกิดปัญหาในการเก็บข้อมูล เนื่องจากการจัดแสงและมุมมองในการทดลองและจากผลการทดลอง จึงทำให้ในการเรียนช่วงที่ 1-4 ไม่สามารถสรุปผลได้ เนื่องจากไม่สามารถตรวจจับใบหน้าได้เลย จึงไม่สามารถเก็บข้อมูลที่ตั้งใจเรียนได้ ในเนื้อหาส่วนที่ 5 และ 6 ถึงแม้ว่าจะตรวจจับใบหน้าได้ไม่มาก แต่มีอัตราส่วนที่ตั้งใจเรียนค่อนข้างสูง โดยเนื้อหาส่วนที่ 5 ตรวจจับได้ว่าผู้เรียนตั้งใจเรียนถึง 94.7% และเนื้อหาส่วนที่ 6 ตรวจจับว่าผู้เรียนตั้งใจเรียนได้ 78.9% ซึ่งถึงแม้ว่าอัตราความตั้งใจเรียน 78% เมื่อเทียบกับ 100% จะเป็นอัตราส่วนที่ค่อนข้างมาก แต่ยังต่ำกว่าค่าเฉลี่ย ที่เนื้อหาส่วนที่ 6 มีค่าเฉลี่ยอัตราความตั้งใจเรียนถึง 84.09%

จะเห็นว่าผู้เรียนคนที่ 2 , 6 และ 14 มีความตั้งใจเรียนมากกว่าผู้เรียนคนอื่นๆ คือมีความตั้งใจเรียนทั้ง 100% ตลอดการเรียนเนื้อหาทั้ง 6 ช่วง

ผู้เรียนคนที่ 3 มีอัตราส่วนในความตั้งใจเรียนสูงมาก โดยเนื้อหาทั้ง 6 ส่วน มีความตั้งใจเรียนสูงกว่า 95% ทั้งหมด โดยส่วนที่มีความตั้งใจเรียนต่ำที่สุดคือช่วงที่ 2 ซึ่งมีความตั้งใจเรียน 95% และความตั้งใจเรียนสูงที่สุดคือส่วนที่ 6 ซึ่งมีความตั้งใจเรียนถึง 100%

ผู้เรียนคนที่ 4 มีความตั้งใจเรียนสูงในเนื้อหาส่วนที่ 1 และส่วนที่ 6 โดยเนื้อหาส่วนที่ 1 มีความตั้งใจเรียน 94.36% และส่วนที่ 6 มีความตั้งใจเรียน 89.9 แต่ในเนื้อหาช่วงที่ 2 , 3 , 4 และ 5 มีความตั้งใจเรียนต่ำกว่าค่าเฉลี่ย โดยมีความตั้งใจเรียน 49.86% , 40.63% , 39.24% และ 19.67% ตามลำดับ

ผู้เรียนคนที่ 5 มีความตั้งใจเรียนสูงกว่าค่าเฉลี่ยในเนื้อหาส่วนที่ 1 เท่านั้น โดยมีความตั้งใจเรียน 93.75% และค่อยๆ มีความตั้งใจเรียนลดลงในเนื้อหาส่วนที่ 2 , 3 และ 4 โดยมีความตั้งใจเรียนเท่ากับ 78.14% , 57.26% , 49.32% ตามลำดับ หลังจากนั้นในช่วงที่ 5 และ 6 มีความตั้งใจเรียนเพิ่มขึ้น เท่ากับ 55.23% , 71.23% ซึ่งถึงแม้ว่าจะมีความตั้งใจเรียนเพิ่มขึ้น แต่ยังมีค่าตั้งใจเรียนน้อยกว่าค่าเฉลี่ย

ข้อมูลที่ตรวจจับได้จากผู้เรียนคนที่ 7 ในเนื้อหาส่วนที่ 1 เท่ากับ 13.04% และส่วนที่ 2 เท่ากับ 3.37% ซึ่งมีอัตราความตั้งใจเรียนต่ำมาก โดยตั้งแต่เนื้อหาส่วนที่ 3 เป็นต้นไป มีอัตราความตั้งใจเรียนเพิ่มขึ้นจนมากกว่า 50% โดยมีความตั้งใจเรียนเท่ากับ 54.83% , 73.33% , 63.63% , 61.53% ตามลำดับ แต่ความตั้งใจเรียนทั้งหมดของผู้เรียนคนนี้อยู่ในเกณฑ์ที่ต่ำกว่าค่าเฉลี่ยทั้งหมด

ผู้เรียนคนที่ 8 มีความตั้งใจในเนื้อหาส่วนที่ 1 และ 2 มากกว่าค่าเฉลี่ย โดยมีความตั้งใจเท่ากับ 89.69% , 71.83% ตามลำดับ แต่เนื้อหาส่วนที่ 3 เป็นต้นไป มีความตั้งใจเรียนลดลง โดยในส่วนที่ 3 , 4 และ 6 จะมีความตั้งใจเท่ากับ 71.55% , 69.64% , 66.67% และเนื้อหาส่วนที่ 5 มีความตั้งใจเรียนต่ำที่สุด คือ 9.25%

ผู้เรียนคนที่ 9 มีความตั้งใจเรียนน้อยที่สุดในเนื้อหาส่วนที่ 1 และเป็นช่วงเดียวที่มีความตั้งใจเรียนน้อยกว่าค่าเฉลี่ย คือ 46.15% ในเนื้อหาส่วนอื่นๆ มีความตั้งใจเรียนมากกว่าค่าเฉลี่ย และในเนื้อหาการเรียนช่วงที่ 3 มีความตั้งใจเรียนสูงถึง 100%

ผู้เรียนคนที่ 10, 12, 13 มีลักษณะของความตั้งใจเรียนคล้ายคลึงกัน คือ มีความตั้งใจเรียนสูงมาก โดยในเนื้อหาการเรียนทุกช่วงจะมีความตั้งใจเรียนมากกว่า 90% ทั้งหมด โดยผู้เรียนทั้ง 3 คนนี้ ไม่มีผู้เรียนคนใดเลยที่มีความตั้งใจเรียนในช่วงที่ 2 สูงถึง 100% เนื่องจากเป็นเนื้อหาช่วงที่มีความยาวมากที่สุด โดยในผู้เรียนทั้ง 3 คนนี้ ผู้เรียนคนที่ 10 มีความตั้งใจเรียนสูงสุด คือ มีเพียงช่วงเดียวเท่านั้นที่อัตราความตั้งใจเรียนน้อยกว่า 100% คือช่วงที่ 2 ซึ่งมีความตั้งใจเรียน 99.71% และผู้เรียนที่มีความตั้งใจเป็นอันดับสอง จากผู้เรียนทั้ง 3 คนนี้ คือ ผู้เรียนคนที่ 12 ซึ่งมีความตั้งใจเรียนช่วงที่น้อยกว่า 100% อยู่ 3 ช่วง คือ 99.35% ในช่วงที่ 1 , 99.34% ในช่วงที่ 2 และ 99.36% ในช่วงที่ 4 และผู้เรียนคนที่ 13 ที่มีความตั้งใจเรียนถึง 100% เฉพาะช่วงที่ 5 และ 6 เท่านั้น โดยในเนื้อหาการเรียนช่วงที่ 1 จะมีความตั้งใจเรียนเท่ากับ 94.35% เนื้อหาช่วงที่ 2 มีความตั้งใจเรียนเท่ากับ 93.75% เนื้อหาช่วงที่ 3 ซึ่งเป็นเนื้อหาส่วนที่ผู้เรียนคนนี้มีค่าความตั้งใจเรียนน้อยที่สุด มีความตั้งใจเรียนเท่ากับ 91.18% และเนื้อหาส่วนที่ 4 ที่มีความตั้งใจเพิ่มขึ้น มีความตั้งใจเรียนเท่ากับ 96.69%

ผู้เรียนคนที่ 11 มีความตั้งใจเรียน 100% ในช่วงที่ 1 และ 3 , ช่วงที่ 2 มีความตั้งใจเรียนเท่ากับ 99.24% , ช่วงที่ 4 มีความตั้งใจเรียน 99.39% , ช่วงที่ 5 มีความตั้งใจเรียน 97.62% และช่วงที่ 6 มีความตั้งใจเรียนลดลงจนเหลือ 85.19% ซึ่งถึงแม้ว่าในเนื้อหาการเรียนช่วงที่ 6 จะมีความตั้งใจเรียนลดลง แต่ในเนื้อหาทุกส่วนยังมีความตั้งใจมากกว่าค่าเฉลี่ยทั้งหมด

ผู้เรียนคนที่ 15 มีลักษณะความตั้งใจเรียนคล้ายกับผู้เรียนคนที่ 11 คือ ในขณะที่เรียนเนื้อหาในบางช่วง มีความตั้งใจเรียนสูงถึง 100% มีความตั้งใจเรียนในเนื้อหาส่วนมากอยู่ระหว่าง 90-100% และมีเพียงหนึ่งช่วง ที่มีความตั้งใจต่ำกว่า 90% แต่ยังมากกว่าค่าเฉลี่ย โดยเนื้อหาส่วนที่ผู้เรียนคนที่ 15 มีความตั้งใจเรียนต่ำที่สุด คือ เนื้อหาการเรียนในช่วงที่ 1 เท่ากับ 87.41% และในช่วงที่ 2, 3, 4, 5 และ 6 มีความตั้งใจเท่ากับ 97.32% , 99.36% , 99.61% , 100.00% , 97.37% ตามลำดับ

ผู้เรียนคนที่ 16 มีความตั้งใจเรียนต่ำกว่าค่าเฉลี่ยในทุกช่วง แต่ในเนื้อหาการเรียนส่วนที่ 1, 3, 4, 5 มีอัตราความตั้งใจเรียนใกล้เคียงกับค่าเฉลี่ย คือ ในช่วงที่ 1 มีความตั้งใจเรียนเท่ากับ 81.10% ซึ่งใกล้เคียงกับค่าเฉลี่ยความตั้งใจเรียนในช่วงที่ 1 ที่มีค่าเท่ากับ 86.47% , ในช่วงที่ 2 มีค่าความตั้งใจเรียนเท่ากับ 66.30% ซึ่งแตกต่างกับค่าเฉลี่ยค่อนข้างมาก เมื่อเปรียบเทียบผลต่างของค่าเฉลี่ยของอัตราความตั้งใจเรียนของผู้เรียนคนนี้ในเนื้อหาการเรียนช่วงอื่นๆ , ช่วงที่ 3 มีอัตราความตั้งใจเรียนเท่ากับ 86.00% และค่าเฉลี่ยของความตั้งใจเรียนในช่วงที่ 3 เท่ากับ 86.43% , ช่วงที่ 4 มีอัตราความตั้งใจเรียนเท่ากับ 86.57% และค่าเฉลี่ยของความตั้งใจเรียนในช่วงที่ 4 เท่ากับ 87.26% , ช่วงที่ 5 มีอัตราความตั้งใจเรียนเท่ากับ 75.29% และค่าเฉลี่ยของความตั้งใจเรียนในช่วงที่ 5 เท่ากับ 81.91% แต่

ในช่วงที่ 6 ไม่สามารถตรวจจับใบหน้าที่ตั้งใจเรียนของผู้เรียนคนนี้ได้เลย จึงมีอัตราความตั้งใจเรียนเท่ากับ 0.00%

จากตารางที่ 4.1 , 4.2 , 4.3 และ 4.4 จะเห็นได้ว่า ข้อมูลระหว่างจำนวนวินาที และ เปอร์เซ็นต์ของจำนวนวินาทีที่ตรวจจับใบหน้าที่ เป็นไปในทิศทางเดียวกับจำนวนวินาที และ เปอร์เซ็นต์ของความตั้งใจเรียน เมื่อพิจารณาข้อมูลทั้ง 4 ตารางร่วมกันแล้ว จะได้ผลลัพธ์ว่า ผู้เรียนคนที่ 2 , 6, และ 14 มีความตั้งใจเรียนสูงมาก (100%) ตามที่แสดงในตารางที่ 4.4 และภายในตารางที่ 4.2 แสดงให้เห็นว่า ระบบสามารถตรวจจับใบหน้าของผู้เรียนในกลุ่มนี้ได้อย่างต่อเนื่อง (99-100%) อย่างไรก็ตาม ข้อสรุปนี้ไม่ตรงกับผู้เรียนลำดับที่ 7 ซึ่งมีอัตราการตรวจจับใบหน้าในช่วงแรกได้ต่ำ (เนื้อหาส่วนที่ 1-2) และเพิ่มขึ้นในภายหลัง (เนื้อหาส่วนที่ 4-6) อย่างไรก็ตาม เปอร์เซ็นต์ในการตรวจจับใบหน้าในตารางที่ 4.2 ไม่สอดคล้องกับผลลัพธ์ที่ได้ว่า ผลลัพธ์ในตารางที่ 4.2 และ 4.4 เป็นไปในทิศทางเดียวกัน โดยเฉพาะอย่างยิ่งในเนื้อหาส่วนที่ 2 ระบบตรวจจับใบหน้าได้ถึง 44.47% ตามที่แสดงในตารางที่ 4.4 แต่ตรวจจับใบหน้าที่แสดงความสนใจในการเรียนได้เพียง 3.37% ตามที่แสดงในตารางที่ 4.2

จากตารางและบทสรุปข้างต้น จะเห็นได้ว่าผู้เรียนที่มีอัตราความตั้งใจเรียนสูงกว่าค่าเฉลี่ยอย่างสม่ำเสมอ ได้แก่ ผู้เรียนคนที่ 2, 3, 6, 10, 11, 12, 13, 14, 15 ผู้เรียนที่มีอัตราความตั้งใจเรียนต่ำกว่าค่าเฉลี่ยอย่างสม่ำเสมอ ได้แก่ ผู้เรียนคนที่ 7 และผู้เรียนที่มีอัตราความตั้งใจเรียนไม่สม่ำเสมอ ได้แก่ ผู้เรียนคนที่ 1, 4, 5, 7, 8, 9, 16 โดยจากผู้เรียนที่มีความตั้งใจเรียนไม่สม่ำเสมอ ผู้เรียนที่มีอัตราความตั้งใจเรียนสูงกว่าค่าเฉลี่ยมากกว่า ได้แก่ผู้เรียนคนที่ 9 และผู้เรียนที่มีความตั้งใจเรียนต่ำกว่าค่าเฉลี่ยมากกว่ามีอัตราความตั้งใจเรียนสูงกว่าค่าเฉลี่ย ได้แก่ผู้เรียนคนที่ 1, 4, 5, 7, 8, 16

ตารางที่ 4.5 จำนวนวินาทีที่สามารถตรวจจับใบหน้า ที่โมเดลตรวจสอบว่าผู้เรียนไม่ตั้งใจเรียน

ผู้เรียน คนที่	เนื้อหา ส่วนที่ 1	เนื้อหา ส่วนที่ 2	เนื้อหา ส่วนที่ 3	เนื้อหา ส่วนที่ 4	เนื้อหา ส่วนที่ 5	เนื้อหา ส่วนที่ 6
1	-	-	-	-	0.5	0.5
2	0	0	0	0	0	0
3	1	2	7	6	7	0
4	10	52	26	117.5	141.5	10
5	9	61	50	129	93	11
6	0	0	0	0	0	0
7	20	148	11	67.7	8	5
8	10	16.8	29	34	49	8
9	7	25	0	5	3	2
10	0	1	0	0	0	0
11	0	2	0	1	4	4
12	1	1	0	1	0	0
13	7	2	3	4	0	0
14	0	0	0	0	0	0
15	17	8	1	1	0	1
16	24	31	21	9	21	27
ค่าเฉลี่ย	7.07	23.32	9.87	25.01	20.44	4.29

ตารางที่ 4.6 เปอร์เซ็นต์ที่โมเดลตรวจสอบได้ว่าผู้เรียนไม่ตั้งใจเรียน

ผู้เรียน คนที่	เนื้อหา ส่วนที่ 1	เนื้อหา ส่วนที่ 2	เนื้อหา ส่วนที่ 3	เนื้อหา ส่วนที่ 4	เนื้อหา ส่วนที่ 5	เนื้อหา ส่วนที่ 6
1	-	-	-	-	4.55%	16.67%
2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
3	0.67%	0.59%	4.49%	2.33%	3.43%	0.00%
4	5.65%	50.00%	59.09%	50.65%	80.40%	29.41%
5	6.25%	21.86%	42.74%	50.79%	44.93%	28.95%
6	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
7	86.96%	96.73%	35.48%	45.13%	36.36%	38.46%
8	10.31%	28.24%	28.43%	30.36%	90.74%	33.33%
9	53.85%	16.78%	0.00%	1.95%	1.46%	5.41%
10	0.00%	0.29%	0.00%	0.00%	0.00%	0.00%
11	0.00%	0.76%	0.00%	0.61%	2.38%	14.81%
12	0.65%	0.66%	0.00%	0.64%	0.00%	0.00%
13	5.65%	6.25%	8.82%	3.31%	0.00%	0.00%
14	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
15	12.59%	2.68%	0.64%	0.39%	0.00%	2.63%
16	18.90%	33.70%	14.00%	13.43%	24.71%	100.00%
ค่าเฉลี่ย	13.43%	17.24%	12.91%	13.31%	18.06%	16.85%

ผลสรุปในตารางที่ 4.5 และ 4.6 จะเป็นไปในทางตรงข้ามกับตารางที่ 4.3 และ 4.4 เนื่องจาก ตารางที่ 4.5 และ 4.6 เป็นตารางที่เก็บข้อมูลความไม่ตั้งใจเรียน แต่ตารางที่ 4.3 และ 4.4 จะแสดง ความตั้งใจเรียนของผู้เรียน โดยผู้เรียนแต่ละคน มีความไม่ตั้งใจเรียนดังนี้

ผู้เรียนคนที่ 1 ที่เกิดปัญหาในการเก็บข้อมูล จึงทำให้ในการเรียนช่วงที่ 1-4 ไม่สามารถสรุป ผลได้ เนื่องจากไม่สามารถตรวจจับใบหน้าได้เลย จึงไม่สามารถเก็บข้อมูลทั้งความตั้งใจเรียนและไม่ ตั้งใจเรียนได้ ในเนื้อหาส่วนที่ 5 และ 6 เมื่อเทียบกับอัตราการตรวจจับใบหน้าในตารางที่ 4.2 แล้ว จะ เห็นได้ว่ามีอัตราการตรวจจับใบหน้าต่ำ ซึ่งผู้เรียนมีอัตราความไม่ตั้งใจเรียนต่ำด้วย โดยเนื้อหาส่วนที่ 5 สามารถตรวจจับได้ว่าผู้เรียนไม่ตั้งใจเรียน 4.55% และเนื้อหาส่วนที่ 6 ตรวจจับว่าผู้เรียนไม่ตั้งใจเรียน 16.67% ซึ่งถือว่าเป็นผู้เรียนที่มีค่าความไม่ตั้งใจเรียนสูงกว่าค่าเฉลี่ย

ผู้เรียนคนที่ 2 , 6 และ 14 มีความตั้งใจเรียนมากกว่าผู้เรียนคนอื่น ๆ คือตลอดการเรียนเนื้อหาทั้ง 6 ช่วง ไม่มีช่วงที่สามารถตรวจจับความไม่ตั้งใจเรียนได้เลย

ผู้เรียนคนที่ 3 มีอัตราส่วนในความไม่ตั้งใจเรียนต่ำมาก โดยเนื้อหาทั้ง 6 ส่วน มีความไม่ตั้งใจเรียนต่ำกว่า 5% ทั้งหมด โดยส่วนที่มีความไม่ตั้งใจเรียนสูงที่สุดคือช่วงที่ 2 ซึ่งมีไม่ความตั้งใจเรียน 4.49% และความไม่ตั้งใจเรียนต่ำที่สุดคือส่วนที่ 3 คือ 0%

ผู้เรียนคนที่ 4 มีความไม่ตั้งใจเรียนสูงกว่าค่าเฉลี่ยในเนื้อหาการเรียนถึง 5 ช่วง โดยมีเพียงช่วงที่ 1 เท่านั้น ที่มีความไม่ตั้งใจเรียนต่ำกว่าค่าเฉลี่ย โดยมีค่าความไม่ตั้งใจเรียนเท่ากับ 5.65% เมื่อเปรียบเทียบกับค่าเฉลี่ยที่ไม่ตั้งใจเรียนถึง 13.43% แต่ในช่วงอื่นๆ มีอัตราความไม่ตั้งใจเรียนสูง โดยช่วงที่ 2, 3, และ 4 มีอัตราความไม่ตั้งใจเรียนอยู่ที่ช่วง 50-60% แต่ในเนื้อหาส่วนที่ 5 มีความไม่ตั้งใจเรียนสูงถึง 80.40% และลดลงในเนื้อหาการเรียนช่วงที่ 6 ที่ไม่ตั้งใจเรียน 29.41% ซึ่งถึงแม้ว่าจะลดลงเป็นอย่างมาก แต่ยังคงมีความไม่ตั้งใจเรียนมากกว่าค่าเฉลี่ยอยู่

ผู้เรียนคนที่ 5 มีแนวโน้มความไม่ตั้งใจเรียนคล้ายคลึงกับผู้เรียนคนที่ 4 คือ มีความไม่ตั้งใจเรียนต่ำกว่าค่าเฉลี่ยเพียงเนื้อหาส่วนที่ 1 เท่านั้น คือ มีความไม่ตั้งใจเรียน 6.25% ในช่วงที่ 2 และ 6 จะมีความไม่ตั้งใจเรียนอยู่ในช่วง 20-30% โดยช่วงที่ 2 มีความไม่ตั้งใจเรียนเท่ากับ 21.86% และช่วงที่ 6 มีความไม่ตั้งใจเรียนเท่ากับ 28.95% ช่วงที่ 3-5 มีค่าความไม่ตั้งใจเรียนอยู่ในช่วง 40-50% โดยมีค่าความตั้งใจเรียนในช่วงที่ 3, 4, 5 เท่ากับ 42.74% 50.79% 44.93% ตามลำดับ

ข้อมูลที่ตรวจจับได้จากผู้เรียนคนที่ 7 ในเนื้อหาส่วนที่ 1 เท่ากับ 86.96% และส่วนที่ 2 เท่ากับ 96.73% ซึ่งมีอัตราความไม่ตั้งใจเรียนสูงมาก โดยตั้งแต่เนื้อหาส่วนที่ 3 เป็นต้นไป มีอัตราความไม่ตั้งใจเรียนเพิ่มขึ้นลดลงจนต่ำกว่า 50% โดยมีความตั้งใจเรียนเท่ากับ 35.48% , 45.13% , 36.36% และ 38.46% ตามลำดับ แต่ความไม่ตั้งใจเรียนทั้งหมดของผู้เรียนคนนี้อยู่ในเกณฑ์ที่สูงกว่าค่าเฉลี่ยทั้งหมด

ผู้เรียนคนที่ 8 มีความไม่ตั้งใจเรียนในเนื้อหาส่วนที่ 1 ต่ำกว่าค่าเฉลี่ย โดยมีความไม่ตั้งใจเท่ากับ 10.31% แต่เนื้อหาส่วนที่ 3 เป็นต้นไป มีความตั้งใจเรียนลดลง โดยในส่วนที่ 2, 3, 4 และ 6 จะมีความไม่ตั้งใจเท่ากับ 28.24% 28.43% 30.36% 33.33% ตามลำดับ และเนื้อหาส่วนที่ 5 มีความไม่ตั้งใจเรียนสูงที่สุด คือ 90.74% ซึ่งสูงเป็นอันดับสอง เมื่อนับจากอัตราความไม่ตั้งใจเรียนทั้งหมดของผู้เรียนทั้ง 16 คน

ผู้เรียนคนที่ 9 มีความไม่ตั้งใจเรียนสูงที่สุดในเนื้อหาส่วนที่ 1 และเป็นช่วงเดียวที่มีความไม่ตั้งใจเรียนสูงกว่าค่าเฉลี่ย คือ 53.85% หลังจากนั้นในช่วงที่ 2 มีอัตราความไม่ตั้งใจเรียนลดลงเหลือ 16.78% และในช่วงที่ 3 อัตราความไม่ตั้งใจเรียนลดลงจนเหลือ 0% หลังจากนั้นเพิ่มขึ้นเล็กน้อย เป็น 1.95% 1.46% และ 5.41% ในส่วนที่ 4, 5 และ 6 ตามลำดับ

ผู้เรียนคนที่ 10, 12, 13 มีลักษณะของความไม่ตั้งใจเรียนคล้ายคลึงกัน คือ มีความไม่ตั้งใจเรียนต่ำมาก โดยในเนื้อหาการเรียนทุกช่วงจะมีความไม่ตั้งใจเรียนต่ำกว่า 10% ทั้งหมด โดยผู้เรียนทั้ง 3 คนนี้ ไม่มีผู้เรียนคนใดเลยที่มีความไม่ตั้งใจเรียนในช่วงที่ 2 เป็น 0% เนื่องจากเป็นเนื้อหาช่วงที่มีความยาวมากที่สุด โดยในผู้เรียนทั้ง 3 คนนี้ ผู้เรียนคนที่ 10 มีความไม่ตั้งใจเรียนต่ำสุด คือ มีเพียงช่วงเดียวเท่านั้นที่อัตราความไม่ตั้งใจเรียนมากกว่า 0% คือช่วงที่ 2 ซึ่งมีความไม่ตั้งใจเรียน 0.29% และผู้เรียนที่มีความไม่ตั้งใจต่ำเป็นอันดับสอง จากผู้เรียนทั้ง 3 คนนี้ คือ ผู้เรียนคนที่ 12 ซึ่งมีความไม่ตั้งใจเรียนช่วงที่มากกว่า 0% อยู่ 3 ช่วง คือ 0.65% ในช่วงที่ 1 , 0.66% ในช่วงที่ 2 และ 0.64% ในช่วงที่ 4 และผู้เรียนคนที่ 13 ที่มีความไม่ตั้งใจเรียนถึง 0% เฉพาะช่วงที่ 5 และ 6 เท่านั้น โดยในเนื้อหาการเรียนช่วงที่ 1 จะมีความไม่ตั้งใจเรียนเท่ากับ 5.65% เนื้อหาช่วงที่ 2 มีความไม่ตั้งใจเรียนเท่ากับ 6.25% เนื้อหาช่วงที่ 3 ซึ่งเป็นเนื้อหาส่วนที่ผู้เรียนคนนี้มี ความไม่ตั้งใจเรียนสูงที่สุด มีความไม่ตั้งใจเรียนเท่ากับ 8.82% และเนื้อหาส่วนที่ 4 ที่มีความไม่ตั้งใจเรียนลดลง มีความไม่ตั้งใจเรียนเท่ากับ 3.31%

ผู้เรียนคนที่ 11 มีความไม่ตั้งใจเรียน 0% ในช่วงที่ 1 และ 3 , ช่วงที่ 2 มีความไม่ตั้งใจเรียนเท่ากับ 0.76% , ช่วงที่ 4 มีความไม่ตั้งใจเรียน 0.61%, ช่วงที่ 5 มีความไม่ตั้งใจเรียน 2.38% และช่วงที่ 6 มีความไม่ตั้งใจเรียนเพิ่มขึ้นจนถึง 14.81% ซึ่งถึงแม้ว่าในเนื้อหาการเรียนช่วงที่ 6 จะมีความไม่ตั้งใจเรียนเพิ่มขึ้น แต่ในเนื้อหาทุกส่วนยังมีความไม่ตั้งใจต่ำกว่าค่าเฉลี่ยทั้งหมด

ผู้เรียนคนที่ 15 มีลักษณะความตั้งใจเรียนคล้ายกับผู้เรียนคนที่ 11 คือ ในขณะที่เรียนเนื้อหาในบางช่วง มีความไม่ตั้งใจเรียนเป็น 0% มีความไม่ตั้งใจเรียนในเนื้อหาส่วนมากอยู่ระหว่าง 0-5% และมีเพียงหนึ่งช่วง ที่มีความไม่ตั้งใจสูงกว่า 10% แต่ยังมีค่าความไม่ตั้งใจเรียนต่ำกว่าค่าเฉลี่ย โดยเนื้อหาส่วนที่ผู้เรียนคนที่ 15 มีความไม่ตั้งใจเรียนสูงที่สุด คือ เนื้อหาการเรียนในช่วงที่ 1 เท่ากับ 12.59% และในช่วงที่ 2, 3, 4, 5 และ 6 มีความตั้งใจเท่ากับ 2.68%, 0.64%, 0.39%, 0.00%, 2.63% ตามลำดับ

ผู้เรียนคนที่ 16 มีความไม่ตั้งใจเรียนต่ำกว่าค่าเฉลี่ยในทุกช่วง แต่ในเนื้อหาการเรียนส่วนที่ 1, 3, 4, 5 มีอัตราความตั้งใจเรียนใกล้เคียงกับค่าเฉลี่ย คือ ในช่วงที่ 1 มีความไม่ตั้งใจเรียนเท่ากับ 18.90% ซึ่งใกล้เคียงกับค่าเฉลี่ยความไม่ตั้งใจเรียนในช่วงที่ 1 ที่มีค่าเท่ากับ 13.43% , ในช่วงที่ 2 มีค่าความไม่ตั้งใจเรียนเท่ากับ 33.70% ซึ่งแตกต่างกับค่าเฉลี่ยค่อนข้างมาก เมื่อเปรียบเทียบผลต่างของค่าเฉลี่ยของอัตราความตั้งใจเรียนของผู้เรียนคนนี้ในเนื้อหาการเรียนช่วงอื่นๆ , ช่วงที่ 3 มีอัตราความตั้งใจเรียนเท่ากับ 14.00% และค่าเฉลี่ยของความตั้งใจเรียนในช่วงที่ 3 เท่ากับ 12.91%, ช่วงที่ 4 มีอัตราความตั้งใจเรียนเท่ากับ 13.43% และค่าเฉลี่ยของความตั้งใจเรียนในช่วงที่ 4 เท่ากับ 13.31% , ช่วงที่ 5 มีอัตราความตั้งใจเรียนเท่ากับ 24.71% และค่าเฉลี่ยของความตั้งใจเรียนในช่วงที่ 5 เท่ากับ 18.06% แต่ในช่วงที่ 6 สามารถตรวจจับได้เฉพาะใบหน้าที่ไม่ตั้งใจเรียนของผู้เรียนคนนี้ได้ จึงมีอัตราความไม่ตั้งใจเรียนเท่ากับ 100.00%

ผลรวมในตารางที่ 4.4 และ 4.6 จะรวมกันได้ 100% โดยตารางที่ 4.4 จะแสดงเปอร์เซ็นต์ที่ผู้เรียนตั้งใจเรียน และตารางที่ 4.6 จะแสดงเปอร์เซ็นต์ที่ผู้เรียนไม่ตั้งใจเรียน และจำนวนวินาทีในตารางที่ 4.3 และ 4.5 จะรวมกันได้เท่ากับเวลาภายในตารางที่ 4.1 เนื่องจากในการคำนวณเปอร์เซ็นต์ จะคิดจากจำนวนวินาทีที่สามารถตรวจสอบใบหน้าได้เท่านั้น

จากตารางและบทสรุปข้างต้น จะเห็นได้ว่าผู้เรียนที่มีอัตราความไม่ตั้งใจเรียนต่ำกว่าค่าเฉลี่ยอย่างสม่ำเสมอ ได้แก่ ผู้เรียนคนที่ 2, 3, 6, 10, 11, 12, 13, 14, 15 ผู้เรียนที่มีอัตราความไม่ตั้งใจเรียนต่ำกว่าค่าเฉลี่ยอย่างสม่ำเสมอ ได้แก่ ผู้เรียนคนที่ 7 และผู้เรียนที่มีอัตราความไม่ตั้งใจเรียนไม่สม่ำเสมอ ได้แก่ ผู้เรียนคนที่ 1, 4, 5, 7, 8, 9, 16 โดยจากผู้เรียนที่มีความตั้งใจเรียนไม่สม่ำเสมอ ผู้เรียนที่มีอัตราความไม่ตั้งใจเรียนต่ำกว่าค่าเฉลี่ยมากกว่า ได้แก่ผู้เรียนคนที่ 9 และผู้เรียนที่มีความไม่ตั้งใจเรียนสูงกว่าค่าเฉลี่ยมากกว่ามีอัตราความไม่ตั้งใจเรียนต่ำกว่าค่าเฉลี่ย ได้แก่ผู้เรียนคนที่ 1, 4, 5, 7, 8, 16



บทที่ 5

บทสรุป อภิปราย และข้อเสนอแนะ

งานวิจัยนี้ได้เสนอแนวทางในการพัฒนาระบบการเรียนรู้ออนไลน์ ที่มีประสิทธิภาพมากยิ่งขึ้น โดยการใช้ระบบตรวจสอบอารมณ์เพื่อช่วยในกระบวนการการสอนและการเรียนรู้ ซึ่งงานวิจัยนี้ได้ช่วยให้ผู้สอนได้รับข้อมูลเกี่ยวกับอารมณ์ของผู้เรียนในขณะที่กำลังเรียนรู้ ซึ่งสามารถนำมาปรับปรุงบทเรียนในอนาคตได้ ในงานวิจัยนี้ได้นำหลักการของ Convolutional Neural Network (CNN) มาใช้ โดยนำ ResNet-50 ซึ่งเป็น pre-trained model มาใช้งานในการตรวจจับอารมณ์ และนำชุดข้อมูล FER-2013 ซึ่งเป็นชุดข้อมูลของภาพใบหน้าประกอบไปด้วย 7 อารมณ์ ซึ่งมีจำนวนทั้งหมด 35,887 ภาพ และปรับใช้ผลจากงานวิจัยและทฤษฎีต่าง ๆ เพื่อจำแนกอารมณ์ของผู้เรียนออกเป็น 7 กลุ่มอารมณ์ และ 2 กลุ่มความตั้งใจเรียน โดยใช้เทคโนโลยีการตรวจจับใบหน้า (face detection) เพื่อตรวจสอบใบหน้าของผู้เรียน จากนั้นจึงใช้โมเดลที่ทำการปรับปรุงขึ้นเพื่อตรวจจับอารมณ์ของผู้เรียนในขณะนั้น และบันทึกช่วงเวลาที่มีการเปลี่ยนแปลงอารมณ์ของผู้เรียนลงในไฟล์ข้อความ (text file) รวมถึงการบันทึกวิดีโอของผู้เรียนในระหว่างกระบวนการเรียนรู้ด้วย เมื่อกระบวนการเรียนสิ้นสุดลง โปรแกรมจะทำการส่งวิดีโอที่บันทึกและไฟล์ข้อความที่บันทึก timestamp ไปยังที่อยู่อีเมลของอาจารย์ผู้สอนตามที่ได้กำหนดไว้ ซึ่งมีส่วนช่วยให้ผู้สอนสามารถตรวจสอบอารมณ์ของผู้เรียนอย่างละเอียดหรือสรุปผลตามความต้องการของผู้สอนต่อไปได้อย่างมีประสิทธิภาพและสะดวกสบาย

โดยในแง่ประสิทธิภาพของระบบ มีผู้เรียนที่มีความตั้งใจเรียนของทั้งคลิปวิดีโอเป็น 100% อยู่ทั้งหมด 3 คน ได้แก่ผู้เรียนในลำดับที่ 2 , 6 และ 14 เนื่องจากการแบ่งกลุ่มอารมณ์ที่แบ่งให้กลุ่มตั้งใจเรียนมีอารมณ์มีความสุข (Happy) , ประหลาดใจ (Surprised) และเป็นกลาง (Neutral) โดยที่ผู้เรียนมีความตั้งใจเรียนสูง เนื่องจากสาเหตุที่การจัดกลุ่มให้อารมณ์เป็นกลางอยู่ในกลุ่มตั้งใจเรียนเนื่องจากโดยทั่วไปแล้ว ใบหน้าของมนุษย์จะมีอารมณ์เป็นกลาง

จากผลการทดลอง ผู้เรียนบางส่วนจะมีความตั้งใจลดลงต่ำในเนื้อหาการเรียนช่วงที่ 3 และ 4 เป็นหลัก เนื่องจากเป็นเนื้อหาในช่วงกลางของวิดีโอ และมีเวลาในการเรียนผ่านไปแล้วประมาณ 5-10 นาทีเป็นต้นไป ซึ่งทำให้ผู้เรียนมีความสนใจน้อยลง และเริ่มกลับมาสนใจเนื้อหาในการเรียนอีกครั้งในเนื้อหาส่วนที่ 5 และ 6 เนื่องจากเป็นส่วนที่ใกล้จบเนื้อหาของวิดีโอแล้ว ทำให้ผู้เรียนมีความจดจ่อกับบทเรียนมากขึ้น

ในขณะเดียวกัน ผู้เรียนบางส่วนกลับมีความตั้งใจเรียนต่ำที่สุดในเนื้อหาการเรียนส่วนที่ 6 เนื่องจากเป็นช่วงท้ายของวิดีโอแล้ว และมีแนวโน้มเหมือนเนื้อหาสาระสำคัญกำลังจะจบลง จึงไม่ใส่ใจในบทเรียนหลังจากผ่านช่วงสำคัญของเนื้อหาการเรียน

และมีผู้เรียนอีกกลุ่มหนึ่ง ที่มีความตั้งใจเรียนต่ำที่สุดในเนื้อหาการเรียนส่วนที่ 1 ซึ่งอาจเป็นเพราะเป็นเนื้อหาในส่วนเริ่มต้น ผู้เรียนอาจมีความเข้าใจเบื้องต้นมาก่อนแล้วถึงแม้ว่าจะไม่เคยเรียนเนื้อหาในส่วนนี้ก็ตาม เนื่องจากเป็นเนื้อหาที่พบได้ตามบทความทางวิชาการทั่วไป ซึ่งผู้เรียนอาจเคยพบมาก่อน หรืออาจมีสาเหตุเนื่องมาจากเป็นเนื้อหาในส่วนแรก ทำให้ผู้เรียนยังไม่สามารถปรับอารมณ์ของตัวเองให้เหมาะสมกับการเรียนได้อย่างเต็มที่

จากผลสรุปข้างต้นนี้ ความไม่ตั้งใจเรียนในเนื้อหาส่วนที่ 1 ปัจจัยสำคัญจะมาจากตัวผู้เรียนเอง และเนื้อหาในส่วนของบทนำก็จำเป็นต้องมีสำหรับผู้เรียนที่ไม่ได้มีพื้นความรู้ที่เกี่ยวข้องมาก่อนเลย จึงอาจเป็นเนื้อหาส่วนที่มีการปรับปรุงได้น้อยที่สุด แต่สามารถปรับปรุงเนื้อหาในส่วนที่ 2 และ 3 ได้ เช่น เมื่อเนื้อหาผ่านไปประมาณ 10 นาทีแล้ว อาจสลับเป็นการพูดคุยในเนื้อหาอื่น หรือตรวจสอบให้ละเอียดขึ้นเป็นช่วงเวลาที่ผู้เรียนแต่ละคนไม่สนใจเนื้อหา และปรับปรุงการเรียนการสอนในส่วนนั้น

ในแง่ของผู้เรียน เมื่อผู้เรียนทราบแล้วว่าตัวเองไม่สนใจในเนื้อหาช่วงใด หรืออาจเกิดเหตุการณ์ที่ทำให้เสียสมาธิในช่วงใด อาจหาวิธีเพื่อแก้ไขปัญหของตน เพื่อให้สามารถทำให้ตนเองสนใจเนื้อหาการเรียนเพิ่มขึ้นได้

5.1 ข้อจำกัดในงานวิจัย

เนื่องจากสภาพในการทดลองที่ใช้เป็นห้องเรียน ทำให้แสงในการทดลองบริเวณคอมพิวเตอร์แต่ละเครื่องมีความไม่สม่ำเสมอ จึงส่งผลให้ในผู้ทดลองบางคนไม่สามารถตรวจจับใบหน้าได้ดี นอกจากนี้ ทรงผมของผู้เข้าร่วมการทดลองก็มีผลอีกด้วย ในผู้ทดลองที่มีเส้นผมบดบังใบหน้ามาก และผู้ทดลองที่ใส่แว่น และขยับแว่นบ่อย หรือมีการปิดผมบริเวณใบหน้าบ่อยครั้ง จะทำให้การตรวจจับใบหน้าได้ลดลง

5.2 ปัญหาที่พบในงานวิจัย

ภายในงานวิจัยนี้ มีผู้เข้าร่วมการทดลองทั้งหมด 21 คน แต่สามารถทำการทดลองได้เสร็จสมบูรณ์ทั้งหมดเพียง 16 คน โดยมีผู้ทดลอง 5 คน ที่เมื่อการทดลองเสร็จสิ้นแล้ว ไม่สามารถส่งวิดีโอมายังอีเมลของผู้สอนได้ แต่ยังสามารถบันทึกวิดีโอลงบนเครื่องคอมพิวเตอร์ที่ทดลองได้ ในการสรุปผลจึงนำข้อมูลมาสรุปผล เฉพาะผู้เข้าร่วมการทดลองที่สามารถทำงานได้จนจบกระบวนการสมบูรณ์เท่านั้น และควรป้องกันปัญหาที่จะเกิดขึ้นนี้ โดยให้การส่งอีเมลเพื่อแจ้งผลมายังผู้สอนมีการทำงานซ้ำอีก 1-2 ครั้ง เพื่อป้องกันปัญหาที่เกิดขึ้นเนื่องจากสัญญาณอินเทอร์เน็ต

5.3 ข้อเสนอแนะเพิ่มเติม

1. หากนำระบบไปใช้งานจริง ควรปรับปรุงให้สามารถเปิดโปรแกรม และบันทึกวิดีโอได้สะดวกกว่านี้
2. โมเดลในส่วนการตรวจจับอารมณ์สามารถทำงานได้แล้ว แต่ปัญหาที่ไม่สามารถ detect ใบหน้าของผู้ทำการทดลองได้ เกิดจากการเรียกใช้งาน module ในการ detect ใบหน้า อาจจะมีการปรับเปลี่ยนให้เหมาะสมยิ่งขึ้น
3. สามารถพัฒนาระบบนี้ต่อไป เพื่อนำไปใช้ในการตรวจสอบการทุจริตในการสอบได้
4. สามารถบันทึกเสียงเพิ่มเติมได้ เพื่อเป็นการตรวจสอบสภาพแวดล้อมของผู้เรียนด้วย
5. สามารถพัฒนาระบบต่อไป เพื่อให้สรุปผลในการเรียนของผู้เรียนแต่ละคนได้อัตโนมัติ



บรรณานุกรม

- [1] ภาควิชาจุลชีววิทยา คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล, “ไวรัสโคโรนาสายพันธุ์ใหม่ (Novel Coronavirus 2019, 2019-nCoV),” [Online]. Available: https://www.si.mahidol.ac.th/sidoctor/sirirajonline2021/article_files/1410_1.pdf [Accessed : 10 ธันวาคม 2564].
- [2] จิราภรณ์ ศรีแจ่ม, “วันที่ไทยรู้จัก COVID-19,” [Online]. Available: <https://news.thaipbs.or.th/content/290347>. [Accessed : 10 ธันวาคม 2564].
- [3] พิรัช วัชรสินธุ, “วิพากษ์ระบบการศึกษาไทย ลู่อัดเสนอ 6 มาตรการ (6 Re's) เพื่อหาทางออกร่วมกันในช่วงโควิด,” [Online]. Available : <https://workpointtoday.com/covid-policy-lab-education/> [Accessed : 10 กันยายน 2564].
- [4] PISA Thailand, “การแถลงข่าวผลการประเมิน PISA 2022,” [Online]. Available : <https://pisathailand.ipst.ac.th/news-21/> [Accessed : 30 พฤศจิกายน 2566].
- [5] C. I. Martha and C. Prisca, “Emotional presence, learning, and the online learning environment,” *The International Review of Research Open and Distance Learning*. vol 13, no.4, pp.269-292, October 2012.
- [6] Reinhard Pekrun, Emotions and Learning. *Educational Practices Series-24*, Belley: Gonnet Imprimeur, 2014.
- [7] jazz.ordinaryday, “รวมหลักสูตรป.โทออนไลน์ในไทย อยู่บ้านก็คว้าใบปริญญาได้,” [Online]. Available : <https://www.mangozero.com/thai-online-master-degree/> [Accessed : 4 ตุลาคม 2565].
- [8] ฐานเศรษฐกิจดิจิทัล, “Blended Learning จุดเปลี่ยนการศึกษาไทยหลังโควิด,” [Online]. Available : <https://www.thansettakij.com/business/marketing/547535> [Accessed : 4 ตุลาคม 2565].
- [9] สุวรรณ จัยทอง, “ผลการจัดการเรียนรู้แบบผสมผสาน ในระบบออนไลน์ยุคปกติใหม่ โควิด-19 ที่มีต่อความสามารถในการเรียนรู้ของนักศึกษาครู,” *วารสารวิจัยรำไพพรรณี*, ปีที่ 16, ฉบับที่ 2, หน้า 126-134, พฤษภาคม – สิงหาคม 2565.

- [10] พิริยะ วรธนไทย และคณะ, “ผลของการใช้บทเรียนออนไลน์ต่อผลสัมฤทธิ์ทางการเรียนและความก้าวหน้าทางการเรียน ของผู้เรียนโรงเรียนมัธยมศึกษาขนาดเล็กในจังหวัดเพชรบุรี,” *คุรุสภาวិทยาจารย์*, ปีที่ 4, ฉบับที่ 1, หน้า 76-86, มกราคม - เมษายน 2566.
- [11] J. McCarthy, *Defending AI Research: A Collection of Essays and Reviews*, Chicago, USA: Center for the Study of Language and Information Distributed by Cambridge University Press, 1996.
- [12] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, California: Pearson, 2009.
- [13] จารุณี ดวงสุวรรณ, “เอกสารประกอบการสอน ปัญญาประดิษฐ์ 1,” หาดใหญ่ : คณะวิทยาศาสตร์ มหาวิทยาลัยสงขลานครินทร์, 2559.
- [14] E. Rich and K. Knight, *Artificial Intelligence*, New Dehli: McGraw Hill Companies, 2006.
- [15] D. L. Poole and A. K. Mackworth, *Artificial Intelligence: Foundations of Computational Agents*, (2nd ed.), Cambridge:Cambridge University Press, 2017.
- [16] Daniel Saez Trigueros, Li Meng, “Face Recognition: From Traditional to Deep Learning Methods,” [Online]. Available : <https://arxiv.org/pdf/1811.00116.pdf> [Accessed :October 12 ,2022].
- [17] R. Lienhart and J. Maydt, “An extended set of haar-like features for rapid object detection,” *Proceedings International Conference on Image Processing*, Rochester, NY, USA, September 22-25, 2002, pp.900-903.
- [18] J. Redmon and A. Farhadi, “YOLO9000: Better, faster, stronger,” *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR2017*, Honolulu, HI, USA, July 21-26, 2017, pp. 6517-6525.
- [19] Y. Tian et al., “Recognizing action units for facial expression analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 97-115, February 2001.
- [20] P. Ekman and W. V. Friesen, *Facial Action Coding System*, Oxford: Oxford University Press, 1978.

- [21] W. Zhao et al., "Multifeature extreme ordinal ranking machine for facial age estimation, mathematical problems in engineering," *Mathematical Problems in Engineering*, vol. 2015, no. 3, pp.1-9, September 2015.
- [22] I. M. Revina, W.R. Sam Emmanuel, "A survey on human face expression recognition techniques," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, issue 6, pp. 619-628, June 2021.
- [23] P. Ekman, "An argument for basic emotions," *Cognition & Emotion*, vol.6 issue3-4, pp.169-200, May 1992.
- [24] M. Lyons et al., "Coding facial expressions with Gabor wavelets," *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, Japan, April 14-16, 1998, pp. 200-205.
- [25] I. Goodfellow et al., *Deep Learning (Vol. 1)*, Cambridge: MIT press, 2016.
- [26] S. M. Mavadati et al., "DISFA: A spontaneous facial action intensity database," *IEEE Transactions on Affective Computing*, vol. 4, no. 2, pp. 151-160, April-June 2013.
- [27] W. Mellouk and W. Handouzi, "Facial emotion recognition using deep learning: Review and insights," *The 2nd International Workshop on the Future of Internet of Everything, FioE*, Leuven, Belgium, August 9-12, 2020, pp. 689-694.
- [28] Y. L. Cun et al., "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol.86 no.11, pp. 2278-2324, November 1998.
- [29] Y. L. Cun et al., "Handwritten zip code recognition with multilayer networks," *10th International Conference on Pattern Recognition, PreMI 1990*, Atlantic City, NJ, USA, June 16-21, 1990, pp. 35-40.
- [30] N. Raksaard and O. Surinta, "Comparative study between local descriptors and deep learning for silk pattern image retrieval," *Journal of Science and Technology Mahasarakham University*, vol.37 no.6, pp.736-746, November-December, 2018.
- [31] J. Sanuksan and O. Surinta, "Deep convolutional neural networks for plant recognition in natural environment," *Journal of Science and Technology Mahasarakham University*, vol.38 no.2, pp. 113-124, March-April 2019.

- [32] ชิตพงษ์ กิตตินราทร, “Convolutional Neural Network,” [Online]. Available: <https://guopai.github.io/ml-blog19.html>. [Accessed: 12 ตุลาคม 2565].
- [33] P. J. Flynn et al., “3D Face Recognition,” [Online]. Available: https://link.springer.com/chapter/10.1007/978-0-387-71041-9_11 [Accessed: October 16, 2022].
- [34] H. Li et al., “A convolutional neural network cascade for face detection,” *2015 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015*, Boston, MA, USA, June 7-12, 2015, pp. 5325-5334.
- [35] K. He et al., “Deep residual learning for image recognition,” *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, Las Vegas, June 27-30, 2016, pp. 770-778.
- [36] K. Zhang et al., “Joint face detection and alignment using multi-task cascaded convolutional networks,” *IEEE Signal Processing Letters*, vol.23, pp. 1499-1503, December 2016.
- [37] J. A. Russell, “A circumplex model of affect,” *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, December 1980.
- [38] Cleveland Clinic, “Facial Muscles,” [Online]. Available : <https://my.clevelandclinic.org/health/body/21672-facial-muscles>. [Accessed: October 18, 2022].
- [39] B. Farnsworth, “Facial Action Coding System (FACS) – A Visual Guidebook,” [Online]. Available: <https://imotions.com/blog/facial-action-coding-system/> [Accessed: October 17, 2022].
- [40] The Nature of Things, “The Seven Universal Emotions We Wear On Our Face,” [Online]. Available: <https://www.cbc.ca/natureofthings/features/the-seven-universal-emotions-we-wear-on-our-face> [Accessed: October 19, 2022].
- [41] L. Zahara et al., “The facial emotion recognition (FER-2013) dataset for prediction system of micro-expressions face using the convolutional neural network (CNN) algorithm based raspberry pi,” *Conference: 2020 Fifth International Conference on Informatics and Computing, ICIC 2020*, Gorontalo, Indonesia, November 3-4, 2020, pp. 1-9.

- [42] Manas Sambare, “FER-2013,” [Online]. Available : <https://www.kaggle.com/datasets/msambare/fer2013>. [Accessed : February 19, 2023].
- [43] I. J. Goodfellow et al., “Challenges in representation learning: A report on three machine learning contests,” *Neural Networks*, vol. 64, pp. 59-63, April 2015.
- [44] S. Raschka, “*Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning*,” [Online]. Available: <https://arxiv.org/pdf/1811.12808.pdf>. [Accessed : December 16, 2023].
- [45] J. Mohajon, “*Confusion Matrix for Your Multi-Class Machine Learning Model*,” [Online]. Available: <https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826>. [Accessed: October 16, 2022].
- [46] Microsoft, “*Visual Studio Code*,” [Online]. Available: <https://code.visualstudio.com/docs>. [Accessed: October 16, 2022].
- [47] Google, “*Google Docs API*,” [Online]. Available: <https://developers.google.com/docs/api/reference/rest>. [Accessed: October 16, 2022].
- [48] G. V. Rossum, “*The History of Python*,” [Online]. Available: <http://python-history.blogspot.com/2009/01/personal-history-part-1-cwi.html>. [Accessed: October 16, 2023].
- [49] Python, “*General Python FAQ*,” [Online]. Available: <https://docs.python.org/3/faq/general.html#why-is-it-called-python>. [Accessed: December 16, 2023].
- [50] M. Jaworski and T. Ziade, *Expert Python Programming*, 2nd ed., Birmingham: Packt, 2016.
- [51] B. Kernighan and Dennis Ritchie, *The C Programming Language*, New Jersey: Prentice Hall, 1978.
- [52] AWS, “*SMTP คืออะไร*,” [Online]. Available: <https://aws.amazon.com/th/what-is/smtp/>. [Accessed : 8 ธันวาคม 2566].
- [53] J. Rhoton, *Programmer's Guide to Internet Mail: SMTP, POP, IMAP, and LDAP*, Boston: Digital Press, 1999.

- [54] Google, “Google Drive,” [Online]. Available: <https://www.google.com/drive/> [Accessed : December 9, 2022].
- [55] พิษณุสินี เสถียรธราดล, “ผลกระทบของการเรียนออนไลน์ภายใต้สถานการณ์การแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019: กรณีศึกษานิสิตหลักสูตรการศึกษาด้านจิตและศิลปศาสตรบัณฑิต สาขาวิชาภาษาอังกฤษ มหาวิทยาลัยพะเยา” *วารสารบัณฑิตแสงโคมคำ*, ปีที่ 6, ฉบับที่ 3, หน้า 423 – 439, กันยายน – ธันวาคม, 2564.
- [56] ตริรัตน์ เสริมทรัพย์ และคณะ, “วิจัยเชิงสำรวจการรู้จำใบหน้า” *วิทยานิพนธ์ วท.ม. (วิทยาการคอมพิวเตอร์)*, มหาวิทยาลัยขอนแก่น, ขอนแก่น, 2557.
- [57] ศศิชา บุญเก่า และวิฑิต ฉัตรรัตนกุลชัย, “การจับความรู้สึกของคนจากใบหน้าด้วยเทคนิคปัญญาประดิษฐ์,” *การประชุมวิชาการเครือข่ายวิศวกรรมเครื่องกลแห่งประเทศไทยครั้งที่ 33, NETT33*, จังหวัดอุดรธานี, 2-5 กรกฎาคม 2562, หน้า 570.
- [58] วชิรชัย คงศิริวัฒนา และคณะ, “ระบบตรวจสอบการเข้าชั้นเรียนและประเมินความสนใจผ่านลักษณะอารมณ์ทางใบหน้าด้วยกล้องเว็บแคม,” *วารสารวิทยาศาสตร์ลาดกระบัง*, ปีที่ 30, ฉบับที่ 2, หน้า 42-57, กรกฎาคม-ธันวาคม 2564.
- [59] จูติพงษ์ รักสาริกรณ์ และธนภัทร ชังคะจิตร, “การจำแนกอารมณ์จากใบหน้าแบบ Real-Time บนอุปกรณ์ฝังตัว ด้วยแบบจำลองการเรียนรู้เชิงลึก,” *วิทยานิพนธ์ วศ.ม. (วิศวกรรมข้อมูลขนาดใหญ่)*, มหาวิทยาลัยธุรกิจบัณฑิต, กรุงเทพฯ, 2563.
- [60] ภูบดี ศิวาวงศ์ และคณะ, “ระบบตรวจสอบเพื่อยืนยันตัวบุคคลด้วยใบหน้าบนแอนดรอยด์,” *ปริญญานิพนธ์ วศ.บ. (คอมพิวเตอร์)*, มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ, 2558.
- [61] P. Claudio et al., “Face and eye tracking algorithm based on digital image processing,” *2001 IEEE International Conference on Systems, Man and Cybernetics, e-Systems and e-Man for Cybernetics in Cyberspace, Cat.No.01CH37236*, Tucson, AZ, USA, October 7-10, 2001, pp. 1178-1183.

ภาคผนวก

TNI

ภาคผนวก ก

โค้ดโปรแกรม (System Code)



โค้ด Python สำหรับการสร้าง Model ด้วยอัลกอริทึม ResNet-50

```
#Reference : https://www.kaggle.com/code/abduulrahmankhalid/emotion-based-music-recommender-resnet50v2
```

```
from tensorflow.keras.applications import ResNet50V2
from tensorflow.keras import layers, models, optimizers
from tensorflow.keras.layers import *
from tensorflow.keras.models import Sequential, Model
from tensorflow.keras.utils import plot_model
from tensorflow.keras.callbacks import EarlyStopping, ModelCheckpoint,
ReduceLROnPlateau
from tensorflow.keras.preprocessing.image import ImageDataGenerator,
load_img, img_to_array
from sklearn.model_selection import train_test_split
import cv2
import keras
import tensorflow as tf
import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import confusion_matrix, accuracy_score, precision_score,
recall_score, f1_score
from datetime import datetime

train_dir = 'emotion_detection/data/train/'
test_dir = 'emotion_detection/data/test/'
```

```
# Visualizing Classes
def Classes_Count(path, name):
    Classes_Dict = {}

    for Class in os.listdir(path):

        Full_Path = path + Class
        Classes_Dict[Class] = len(os.listdir(Full_Path))

    df = pd.DataFrame(Classes_Dict, index=[name])

    return df

Train_Count = Classes_Count(train_dir, 'Train').transpose(
).sort_values(by="Train", ascending=False)
Test_Count = Classes_Count(test_dir, 'Test').transpose(
).sort_values(by="Test", ascending=False)

pd.concat([Train_Count, Test_Count], axis=1)

Train_Count.plot(kind='barh')

Test_Count.plot(kind='barh')

plt.style.use('default')
plt.figure(figsize=(25, 8))
image_count = 1
BASE_URL = train_dir
```

```

for directory in os.listdir(BASE_URL):
    if directory[0] != '.':
        for i, file in enumerate(os.listdir(BASE_URL + directory)):
            if i == 1:
                break
            else:
                fig = plt.subplot(1, 7, image_count)
                image_count += 1
                image = cv2.imread(BASE_URL + directory + '/' + file)
                plt.imshow(image)
                plt.title(directory, fontsize=20)

# Data Preprocessing
img_shape = 48
batch_size = 64
train_data_path = train_dir
test_data_path = test_dir

train_preprocessor = ImageDataGenerator(
    rescale=1 / 255.,
    # Data Augmentation
    rotation_range=10,
    zoom_range=0.2,
    width_shift_range=0.1,
    height_shift_range=0.1,
    horizontal_flip=True,
    fill_mode='nearest',
)

test_preprocessor = ImageDataGenerator(
    rescale=1 / 255.,

```

```
)

train_data = train_preprocessor.flow_from_directory(
    train_data_path,
    class_mode="categorical",
    target_size=(img_shape, img_shape),
    color_mode='rgb',
    shuffle=True,
    batch_size=batch_size,
    subset='training',
)

test_data = test_preprocessor.flow_from_directory(
    test_data_path,
    class_mode="categorical",
    target_size=(img_shape, img_shape),
    color_mode="rgb",
    shuffle=False,
    batch_size=batch_size,
)

# Building ResNet50V2 Model
# specifying new image shape for resnet
img_shape = 224
batch_size = 64
train_data_path = train_dir
test_data_path = test_dir

train_preprocessor = ImageDataGenerator(
    rescale=1 / 255.,
    rotation_range=10,
    zoom_range=0.2,
```



```
width_shift_range=0.1,
height_shift_range=0.1,
horizontal_flip=True,
fill_mode='nearest',
)

test_preprocessor = ImageDataGenerator(
    rescale=1 / 255.,
)

train_data = train_preprocessor.flow_from_directory(
    train_data_path,
    class_mode="categorical",
    target_size=(img_shape, img_shape),
    color_mode='rgb',
    shuffle=True,
    batch_size=batch_size,
    subset='training',
)

test_data = test_preprocessor.flow_from_directory(
    test_data_path,
    class_mode="categorical",
    target_size=(img_shape, img_shape),
    color_mode="rgb",
    shuffle=False,
    batch_size=batch_size,
)

# ### Fine-Tuning ResNet50V2
```

```

ResNet50V2 = tf.keras.applications.ResNet50V2(input_shape=(224, 224, 3),
                                              include_top=False,
                                              weights='imagenet'
                                              )

# Freezing all layers except last 50

ResNet50V2.trainable = True

for layer in ResNet50V2.layers[:-50]:
    layer.trainable = False

def Create_ResNet50V2_Model():

    model = Sequential([
        ResNet50V2,
        Dropout(.25),
        BatchNormalization(),
        Flatten(),
        Dense(64, activation='relu'),
        BatchNormalization(),
        Dropout(.5),
        Dense(7, activation='softmax')
    ])
    return model

ResNet50V2_Model = Create_ResNet50V2_Model()
ResNet50V2_Model.summary()
ResNet50V2_Model.compile(
    optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])

```

```

# Create Callback Checkpoint
checkpoint_path = "ResNet50V2_Model_Checkpoint"

Checkpoint = ModelCheckpoint(
    checkpoint_path, monitor="val_accuracy", save_best_only=True)

# Create Early Stopping Callback to monitor the accuracy
Early_Stopping = EarlyStopping(
    monitor='val_accuracy', patience=7, restore_best_weights=True, verbose=1)

# Create ReduceLROnPlateau Callback to reduce overfitting by decreasing
learning
Reducing_LR = tf.keras.callbacks.ReduceLROnPlateau(monitor='val_loss',
                                                    factor=0.2,
                                                    patience=2,
                                                    verbose=1)

callbacks = [Early_Stopping, Reducing_LR]

steps_per_epoch = train_data.n // train_data.batch_size
validation_steps = test_data.n // test_data.batch_size

now = datetime.now()
dt_string = now.strftime("%Y%m%d_%H%M%S")
print(dt_string)

ResNet50V2_history = ResNet50V2_Model.fit(train_data,
validation_data=test_data, epochs=30, batch_size=batch_size,
                                       callbacks=callbacks,
                                       steps_per_epoch=steps_per_epoch, validation_steps=validation_steps)
ResNet50V2_Model.save(f"resnet_30epoch_{dt_string}.h5")

```

```
# Evaluating ResNet50V2

ResNet50V2_Model = tf.keras.models.load_model(
    f"Model/resnet_30epoch_20230807_094237.h5")

ResNet50V2_Score = ResNet50V2_Model.evaluate(test_data)

print("    Test Loss: {:.5f}".format(ResNet50V2_Score[0]))
print("Test Accuracy: {:.2f}%".format(ResNet50V2_Score[1] * 100))

def plot_curves(history):

    loss = history.history["loss"]
    val_loss = history.history["val_loss"]

    accuracy = history.history["accuracy"]
    val_accuracy = history.history["val_accuracy"]

    epochs = range(len(history.history["loss"]))

    plt.figure(figsize=(15, 5))

    # plot loss
    plt.subplot(1, 2, 1)
    plt.plot(epochs, loss, label="training_loss")
    plt.plot(epochs, val_loss, label="val_loss")
    plt.title("Loss")
    plt.xlabel("epochs")
    plt.legend()

    # plot accuracy
```



```

plt.subplot(1, 2, 2)
plt.plot(epochs, accuracy, label="training_accuracy")
plt.plot(epochs, val_accuracy, label="val_accuracy")
plt.title("Accuracy")
plt.xlabel("epochs")
plt.legend()

ResNet50V2_Predictions = ResNet50V2_Model.predict(test_data)
# Choosing highest probalbilty class in every prediction
ResNet50V2_Predictions = np.argmax(ResNet50V2_Predictions, axis=1)

fig, ax = plt.subplots()
# fig, ax = plt.subplots(figsize=(15, 10))
cm = confusion_matrix(test_data.labels, ResNet50V2_Predictions)
sns.heatmap(cm, annot=True, fmt='g', ax=ax, cmap=plt.cm.Blues,
            linewidths=0.1, linecolor='lightblue')

ax.set_xlabel('Predicted labels', fontsize=15)
ax.set_ylabel('True labels', fontsize=15)
ax.set_title('ResNet50V2 Confusion Matrix', fontsize=20)
# Visualizing Predictions
Emotion_Classes = ['Angry',
                    'Disgust',
                    'Fear',
                    'Happy',
                    'Neutral',
                    'Sad',
                    'Surprise']

# Shuffling Test Data to show diffrent classes
test_preprocessor = ImageDataGenerator(
    rescale=1 / 255.,

```

```

)

test_generator = test_preprocessor.flow_from_directory(
    test_data_path,
    class_mode="categorical",
    target_size=(img_shape, img_shape),
    color_mode="rgb",
    shuffle=True,
    batch_size=batch_size,
)

Random_batch = np.random.randint(0, len(test_generator) - 1)

Random_Img_Index = np.random.randint(0, batch_size - 1, 10)

fig, axes = plt.subplots(nrows=2, ncols=5, figsize=(25, 10),
                        subplot_kw={'xticks': [], 'yticks': []})

for i, ax in enumerate(axes.flat):

    Random_Img = test_generator[Random_batch][0][Random_Img_Index[i]]

    Random_Img_Label = np.argmax(
        test_generator[Random_batch][1][Random_Img_Index[i]])

    Model_Prediction = np.argmax(ResNet50V2_Model.predict(
        tf.expand_dims(Random_Img, axis=0), verbose=0))

    ax.imshow(Random_Img)

# Display 10 random pictures from the dataset with their labels

```

```

Random_batch = np.random.randint(0, len(test_generator) - 1)

Random_Img_Index = np.random.randint(0, batch_size - 1, 10)

fig, axes = plt.subplots(nrows=2, ncols=5, figsize=(25, 10),
                        subplot_kw={'xticks': [], 'yticks': []})

for i, ax in enumerate(axes.flat):

    Random_Img = test_generator[Random_batch][0][Random_Img_Index[i]]
    Random_Img_Label = np.argmax(
        test_generator[Random_batch][1][Random_Img_Index[i]])

    Model_Prediction = np.argmax(ResNet50V2_Model.predict(
        tf.expand_dims(Random_Img, axis=0), verbose=0))

    ax.imshow(Random_Img)

    if Emotion_Classes[Random_Img_Label] ==
    Emotion_Classes[Model_Prediction]:
        color = "green"
    else:
        color = "red"
    ax.set_title(
        f"True: {Emotion_Classes[Random_Img_Label]}\nPredicted:
        {Emotion_Classes[Model_Prediction]}", color=color)
    plt.show()
    plt.tight_layout()

# accuracy: (tp + tn) / (p + n)
accuracy = accuracy_score(test_data.labels, ResNet50V2_Predictions)

```

```

print('Accuracy: %f' % accuracy)
# precision tp / (tp + fp)
precision = precision_score(
    test_data.labels, ResNet50V2_Predictions, average='weighted')
print('Precision: %f' % precision)
# recall: tp / (tp + fn)
recall = recall_score(
    test_data.labels, ResNet50V2_Predictions, average='weighted')

print('Recall: %f' % recall)
# f1: 2 tp / (2 tp + fp + fn)
f1 = f1_score(test_data.labels, ResNet50V2_Predictions, average='weighted')
print('F1 score: %f' % f1)

ytrain = []
ytest = []
for i in range(len(test_data.labels)):
    # print(test_data.labels[i])
    if test_data.labels[i] == 3 or test_data.labels[i] == 4 or test_data.labels[i] ==
6:
        ytrain.append('attend')
    elif test_data.labels[i] == 0 or test_data.labels[i] == 1 or test_data.labels[i]
== 2 or test_data.labels[i] == 5:
        ytrain.append('not_attend')
    else:
        ytrain.append('wow')
    if ResNet50V2_Predictions[i] == 3 or ResNet50V2_Predictions[i] == 4 or
ResNet50V2_Predictions[i] == 6:
        ytest.append('attend')
    elif ResNet50V2_Predictions[i] == 0 or ResNet50V2_Predictions[i] == 1 or
ResNet50V2_Predictions[i] == 2 or ResNet50V2_Predictions[i] == 5:

```

```

        ytest.append('not_attend')
    else:
        ytest.append('wow')

y_train = np.array(ytrain)
y_test = np.array(ytest)

for i in range(len(test_data.labels)):
    if y_train[i] != y_test[i]:
        print('Y train = '+str(y_train[i]))
        print('Y test = '+str(y_test[i]))
        print(i)
        print('-----')

fig, ax = plt.subplots()
cm = confusion_matrix(y_train, y_test)

sns.heatmap(cm, annot=True, fmt='g', ax=ax, cmap=plt.cm.Blues,
            linewidths=0.1, linecolor='lightblue')

ax.set_xlabel('Predicted labels', fontsize=15)
ax.set_ylabel('True labels', fontsize=15)
ax.set_title('ResNet50V2 Confusion Matrix', fontsize=20)

accuracy = accuracy_score(y_train, y_test)
print('Accuracy: %f % accuracy)
precision = precision_score(
    y_train, y_test, average='weighted')
print('Precision: %f % precision)
recall = recall_score(

```



```
y_train, y_test, average='weighted')  
print('Recall: %f' % recall)  
f1 = f1_score(y_train, y_test, average='weighted')  
print('F1 score: %f' % f1)
```



โค้ด Python สำหรับการเชื่อมต่อ Service Google

```
import pickle
import os
from google_auth_oauthlib.flow import Flow, InstalledAppFlow
from googleapiclient.discovery import build
from googleapiclient.http import MediaFileUpload, MediaIoBaseDownload
from google.auth.transport.requests import Request

def Create_Service(client_secret_file, api_name, api_version, *scopes):
    print(client_secret_file, api_name, api_version, scopes, sep='-')
    CLIENT_SECRET_FILE = client_secret_file
    API_SERVICE_NAME = api_name
    API_VERSION = api_version
    SCOPES = [scope for scope in scopes[0]]
    print(SCOPES)

    cred = None

    pickle_file = f'token_{API_SERVICE_NAME}_{API_VERSION}.pickle'
    # print(pickle_file)

    if os.path.exists(pickle_file):
        with open(pickle_file, 'rb') as token:
            cred = pickle.load(token)

    if not cred or not cred.valid:
        # if cred and cred.expired and cred.refresh_token:
        #     cred.refresh(Request())
        # else:
```

```
flow = InstalledAppFlow.from_client_secrets_file(CLIENT_SECRET_FILE,
SCOPE)

cred = flow.run_local_server()

with open(pickle_file, 'wb') as token:
    pickle.dump(cred, token)

try:
    service = build(API_SERVICE_NAME, API_VERSION, credentials=cred)
    print(API_SERVICE_NAME, 'service created successfully')
    return service
except Exception as e:
    print('Unable to connect.')
    print(e)
    return None

def convert_to_RFC_datetime(year=1900, month=1, day=1, hour=0,
minute=0):
    dt = datetime.datetime(year, month, day, hour, minute, 0).isoformat() + 'Z'
    return dt
```

โค้ด Python สำหรับการอัปโหลดไฟล์ลงใน Google Drive

```
from googleapiclient.http import MediaFileUpload
from Google import Create_Service

CLIENT_SECRET_FILE = 'Authentication_Files\client_secret.json'
API_NAME = 'drive'
API_VERSION = 'v3'
SCOPES = ['https://www.googleapis.com/auth/drive']

service = Create_Service(CLIENT_SECRET_FILE , API_NAME , API_VERSION ,
SCOPES)

def uploadvideo(file_names,mime_types) :
    folder_id = '1vV2GP4sTISGusqGUXV34o2ZV35LP5Ant'
    url = []

    for filename , mime_type in zip(file_names,mime_types) :
        file_metadata = {
            'name' : filename,
            'parents' :[folder_id]
        }
        media =
MediaFileUpload('Output/{0}'.format(filename),mimetype=mime_type)

        response = service.files().create(
            body=file_metadata,
            media_body=media,
            fields='*'
        ).execute()
```

```
file_id = response['id']
urlvid = response['webViewLink']
url.append(urlvid)

request_body = {
    'role' : 'reader',
    'type' : 'anyone'
}

response_permissions = service.permissions().create(
    fileId = file_id,
    body = request_body
).execute()

print(response_permissions)

return url
```


โค้ด Python สำหรับการใช้งานโมเดลและระบบทั้งหมด

```
import cv2
import numpy as np
from keras.models import model_from_json
import os
from other_python_file.timer import Timer
import smtplib
from email.mime.text import MIMEText
from email.mime.multipart import MIMEMultipart
from email.mime.base import MIMEBase
from email import encoders
from upload import uploadvideo
import tensorflow as tf

# Setup port number and server name
smtp_port = 587 # Standard secure SMTP port
smtp_server = "smtp.gmail.com" # Google SMTP Server

# Set up the email lists
email_from = "ch.phurichaya_st@tni.ac.th "
email_list = ["ch.phurichaya_st@tni.ac.th"]

# Define the password (better to reference externally)
pswd = "iynolewewrwlhfcj" # As shown in the video this password is now dead, left
in as example only

# name the email subject
subject = "Send Video Test"
```

```

# set resolution of device
def change_res(cap, width, height):
    cap.set(3, width)
    cap.set(4, height)

# grab resolution dimensions and set video capture to that resolution.
def get_dims(cap, res):
    # Function adapted from https://krr.co/0l6qmh
    # Standard Video Dimensions Sizes
    STD_DIMENSIONS = {
        "480p": (640, 480),
        "720p": (1280, 720),
        "1080p": (1920, 1080),
        "4k": (3840, 2160),
    }

    width, height = STD_DIMENSIONS[res]
    if res in STD_DIMENSIONS:
        width, height = STD_DIMENSIONS[res]
        # change the current caputre device
        # to the resulting resolution
        change_res(cap, width, height)
        return width, height

#set filetype to video (avi,mp4)
def get_video_type(filename):
    VIDEO_TYPE = {
        'avi': cv2.VideoWriter_fourcc(*'XVID'),
    }

    filename, ext = os.path.splitext(filename)

```

```

if ext in VIDEO_TYPE:
    return VIDEO_TYPE[ext]
return VIDEO_TYPE['avi']

# Define the email function (dont call it email!)
def send_emails(filename,link):

    for person in email_list:
        # Make the body of the email
        body = f"""
        video url = {link}
        """

        # make a MIME object to define parts of the email
        msg = MIMEMultipart()
        msg['From'] = email_from
        msg['To'] = person
        msg['Subject'] = subject

        # Attach the body of the message
        msg.attach(MIMEText(body, 'plain'))

    for file in filename:
        # Open the file in python as a binary
        attachment= open(file, 'rb') # r for read and b for binary

        # Encode as base 64
        attachment_package = MIMEBase('application', 'octet-stream')
        attachment_package.set_payload((attachment).read())
        encoders.encode_base64(attachment_package)

```

```

        attachment_package.add_header('Content-Disposition', "attachment;
filename= " + file)

        msg.attach(attachment_package)

    # Cast as string
    text = msg.as_string()

    # Connect with the server
    print("Connecting to server...")
    TIE_server = smtplib.SMTP(smtp_server, smtp_port)
    TIE_server.starttls()
    TIE_server.login(email_from, pswd)
    print("Succesfully connected to server")
    print()

    # Send emails to "person" as list is iterated
    print(f'Sending email to: {person}...')
    TIE_server.sendmail(email_from, person, text)
    print(f'Email sent to: {person}')
    print()

    # Close the port
    TIE_server.quit()

#set video properties
filename = 'Output/video.avi'
frames_per_second = 24.0
res = '720p'
t = Timer()

```

```

attend = ['Happy' , 'Neutral' , 'Surprised']
not_attend = ["Angry", "Disgusted","Fearful", 'Sad']

emotion_dict = {0: "Angry", 1: "Disgusted", 2: "Fearful",
                 3: "Happy", 4: "Neutral", 5: "Sad", 6: "Surprised"}

emotion_model
tf.keras.models.load_model("Model/resnet_30epoch_20230807_094237.h5")

print("Loaded model from disk")

# start the webcam feed
cap = cv2.VideoCapture(1)
out = cv2.VideoWriter(filename, get_video_type(
    filename), 5.93 , get_dims(cap, res))
#start timer
t.start()
file_object = open('Output/sample.txt', 'w')

#declare variable for save emotion to file
emotionNow = ''
emotionPrev = ''
while True:
    # Find haar cascade to draw bounding box around faceq
    ret, frame = cap.read()
    if not ret:
        break

    #call face detector
    face_detector = cv2.CascadeClassifier(
        'Model/haarcascade_frontalface_default.xml')

```



```

#change color of frame
gray_frame = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)

# detect faces available on camera
num_faces = face_detector.detectMultiScale(gray_frame, scaleFactor=1.3,
minNeighbors=5)

time = t.gettime()

# take each face available on the camera and Preprocess it
for (x, y, w, h) in num_faces:
    cv2.rectangle(frame, (x, y-50), (x+w, y+h+10), (0, 255, 0), 4)
    roi_gray_frame = gray_frame[y:y + h, x:x + w]
    cropped_img = cv2.resize(roi_gray_frame, (224, 224))
    cropped_img = cv2.merge((cropped_img, cropped_img, cropped_img))
    img = np.expand_dims(cropped_img, axis=0)
    new_img = img/255.0

# predict the emotions
#try :
emotion_prediction = emotion_model.predict(new_img)
#print(emotion_prediction)

#find max peercentage to show the detected emotion
maxindex = int(np.argmax(emotion_prediction))
cv2.putText(frame, emotion_dict[maxindex], (x+5, y-20),
cv2.FONT_HERSHEY_SIMPLEX, 1, (255, 0, 0), 2, cv2.LINE_AA)
emotionNow = emotion_dict[maxindex]

if emotionPrev != emotionNow :

```

```

print('Now : '+emotionNow)
print('Prev : '+emotionPrev)
emotionPrev = emotionNow

text_attend = "
if emotionNow in attend :
    text_attend = 'Attend (' + emotionNow + ')'
if emotionNow in not_attend :
    text_attend = 'Not Attend (' + emotionNow + ')'

time = t.gettime()
hr = int(time/3600)
min = int((time/3600) % 60)
sec = int(time % 60)
file_object.write(text_attend+" "+str(hr) + ':' + str(min) + ':' + str(sec)+'\n')
# except Exception as err:
#     print(err)

font = cv2.FONT_HERSHEY_SIMPLEX
datet = "
out.write(frame)

#show windows
cv2.imshow('Emotion Detection', frame)

#close windows
if cv2.waitKey(1) & 0xFF == ord('q'):
    file_object.close()
    break

t.stop()

```

```
file_object.close()

#close all windows
cv2.destroyAllWindows()
#wait for close windows
cv2.waitKey(1)

link = uploadvideo(['video.avi'],["video/x-msvideo"])
print(link)
send_emails(['Output\\sample.txt'],link)
cap.release()
out.release()
```



ภาคผนวก ข
การตีพิมพ์ผลงานวิจัย



Conference Paper

Detecting Student Emotions During Online Learning Using Artificial Intelligence Systems

Choksettakit Phurichaya
Information Technology
Thai-Nichi Institute of Technology
Bangkok, Thailand
ch.phurichaya_st@tni.ac.th

Charoenpit Saromporn
Information Technology
Thai-Nichi Institute of Technology
Bangkok, Thailand
saromporn@tni.ac.th

Tiwatthanont Kanda
Information Technology
Thai-Nichi Institute of Technology
Bangkok, Thailand
kanda@tni.ac.th

Boonsieng Pramuk
Information Technology
Thai-Nichi Institute of Technology
Bangkok, Thailand
b.pramuk@tni.ac.th

Abstract—This study focuses on the application of artificial intelligence in the realm of online education, specifically addressing the detection of emotional states among online learners. The primary objective is to develop a robust Convolutional Neural Network (CNN) based system for accurately assessing the emotional engagement of learners during online educational sessions. The learners are categorized into two distinct groups based on their level of engagement: those who actively participate and those who exhibit disinterest. To evaluate the effectiveness of this emotion detection model, pre- and post-learning assessments are conducted to gauge learners' comprehension and the impact of the emotion detection system on their learning outcomes. The research findings demonstrate a notable level of accuracy, with the two emotions classification 84.65%. This research contributes to the broader field of online education by providing a sophisticated tool for assessing and enhancing the emotional well-being and engagement of learners in digital learning environments.

Keywords—CNN, Online Education, Emotion Detection

I. INTRODUCTION

The education system is the main point for the development of human resources for driving the nation forward. It stands as a pivotal avenue for affording opportunities to all citizens, enabling them to acquire knowledge to be applied in their future professions. However, the state of education in Thailand is beset by challenges, encompassing both qualitative and quantitative aspects.

In terms of quality, there exists a predicament exemplified by scores derived from the PISA assessment, which are lower than those of countries within the ASEAN region. Simultaneously, from a quantitative standpoint, issues persist, such as educational disparities and the dearth of educational opportunities for youth in various regions [1].

The COVID-19 pandemic crisis has precipitated a disruption in the educational landscape, resulting in the suspension of conventional teaching and learning activities within educational institutions since the commencement of the 2020 academic year. This predicament has necessitated learners across the nation to primarily adopt alternative learning avenues, such as online platforms and satellite television broadcasts. However, stemming from this crisis, the educational challenges faced in Thailand have been further

exacerbated, particularly in terms of issues pertaining to the quality of education.

Due to the online learning environment, there is a lack of face-to-face communication, which differs from traditional classroom instruction. In the latter, it was possible to gauge learners' emotions and sentiments. In the context of online teaching, limitations arise from the environmental conditions, potentially leading to incomplete comprehension and understanding [2].

The aims of this paper include facilitating the assessment of learners' emotions during online learning and developing a program that utilizes facial recognition technology for the purpose of emotion detection.

In this work, we use a CNN model developed from ResNet50 and trained with the FER-2013 dataset, which classifies human emotions into seven categories.

II. RELATED WORK

A. Facial Emotional Expression Detection

Facial expression [3] is the manner in which emotions are conveyed through the features of facial muscles at a given moment. This involves changes in the facial muscles that indicate emotions. Hence, technology has been employed to detect and analyze these emotional cues on the face.

In their work, Wafa Mellouk and Wahida Handouzi [4] conducted a literature review titled "Facial emotion recognition using deep learning: review and insights." The review revealed that emotions are often categorized into six fundamental categories: sad, fear, disgust, surprise, angry, and happy. Research in this field typically involves the classification of these six basic emotions as well as the recognition of neutral facial expressions (faces without apparent emotions).

To achieve this recognition, convolutional neural network (CNN) has gained popularity as a technology of choice. CNN have demonstrated high accuracy levels, exceeding 90% in emotion recognition tasks. This underscores the effectiveness of CNN in accurately detecting and categorizing emotions based on facial expressions.

B. Convolution Neural Networks

Convolution Neural Network (CNN) is one of the characteristics of neural Network. It is often used to detect faces and emotions. Because it can do both feature extraction and classification, which can be regarded as the highlight of this network [5]-[9].

Since convolution layers are employed in the system, they play a crucial role in extracting distinctive features from images. Convolution layers simulate the human visual system, which perceives images as a collection of subregions and combines them, such as when detecting object edges. This allows the model to learn image characteristics efficiently and accurately. Convolution layers work in conjunction with other layers like Pooling Layers and Fully-Connected Layers, which differentiate Convolutional Neural Network (CNN) from other neural network architectures. These layers collectively enable the neural network to process and analyze images effectively, making CNN well-suited for tasks involving image recognition and feature extraction.

The CNN architecture, characterized by its multiple layers, plays a crucial role in image processing. At its core, the convolutional layer is pivotal, augmented by additional layers structured as follows:

- 1) Input Layer: At the forefront of the CNN, the input layer receives the raw pixel values of the image, serving as the foundational step in the neural network's analysis.
- 2) Convolutional Layers: CNN implements convolutional layers to process input data. These layers employ small, learnable filters known as kernels, which slide over the input data, facilitating feature detection. In the initial stages, these layers identify basic features like edges, corners, and textures. As the network deepens, these features amalgamate to detect intricate patterns and complex objects.
- 3) Activation Function: Following each convolution operation, an activation function (e.g., ReLU, Sigmoid) is applied element-wise. This introduces non-linearity into the system, empowering the network to grasp intricate patterns and relationships within the data.
- 4) Pooling Layers / Subsampling Layers / Down-sampling Layers: Pooling layers play a vital role in reducing the spatial dimensions of the input volume. Utilizing methods such as max-pooling and average-pooling, these layers curtail the computational load and the number of parameters within the network. Additionally, they act as a control mechanism against overfitting, enhancing the network's overall efficiency.
- 5) Fully Connected Layers / Dense Layers: After traversing multiple convolutional and pooling layers, the neural network engages in high-level reasoning through fully connected layers. Neurons within these layers establish connections with all activations in the preceding layer, allowing the network to discern intricate relationships within the data. This intricate connectivity enables the network to comprehend complex patterns, a fundamental aspect of its learning process.
- 6) Output Layer: The final layer of the CNN constitutes the output layer. In classification tasks, this layer often employs a softmax activation function. This function transforms the network's internal activations into class probabilities, providing a clear and interpretable output for

the given input, making it invaluable for various real-world applications, especially in the realm of computer vision and pattern recognition.

The utilization of Convolutional Neural Network (CNN) is particularly well-suited for Face Recognition due to the architecture of CNN, which includes multiple layers of filters. This architecture enhances the quality of facial detection, even when there are changes in image proportions. Additionally, CNN contributes to improving the efficiency and effectiveness of Face Recognition while reducing the number of false positives by learning and identifying facial features in a hierarchical manner.

CNN excels in feature extraction, which is vital for recognizing faces accurately. They can capture complex facial details and patterns, making them highly effective at differentiating between individuals. Moreover, CNN can adapt to variations in lighting, facial expressions, and pose, making them robust for real-world face recognition applications.

In summary, CNN are a suitable choice for Face Recognition due to their ability to extract relevant features, adapt to variations, and enhance recognition accuracy, ultimately contributing to the reduction of false identifications.

C. Emotion Classification

Facial expressions are a fundamental aspect of human communication [10], as outlined in reference. They represent a non-verbal form of communication that is equally significant. Human beings convey a wide range of emotions through their facial expressions, which can be categorized into six primary emotions. Each emotion is typically associated with distinct facial expressions, as follows:

- Angry: furrowed brows, upper and lower eyelids open wide, tightened lips, and a tense mouth.
- Fear: knitted brows, upper eyelids moving upwards, and lips slightly apart.
- Disgust: furrowed brows, a wrinkled nose, an upper lip lifted, and lips pressed together.
- Happy: Tightened muscles around the eyes, visible crow's feet around the eyes, raised cheeks, and corners of the mouth lifted upwards.
- Sadness: inner corners of the eyebrows raised, corners of the mouth pulled downward.
- Surprise: raised eyebrows, widened upper eyelids, an open mouth, and expanded eye sockets.

These facial expressions serve as a universal means of conveying emotions, allowing individuals to communicate their feelings without the need for words. They are a fundamental component of human interaction and understanding.

D. Facial Muscular Movement

On the human face, there are approximately twenty muscles, each with specific functions [11]. These facial muscles play two primary roles:

- Auriculars: Some individuals can use these muscles to move their ears.
- Corrugator Supercilia: Located around the eyebrows, these muscles enable brow furrowing.
- Depressor Anguli Oris: Situated around both sides of the chin, they assist in brow furrowing.
- Depressor Labii Inferioris: Found around the chin, these muscles aid in lowering the lower lip.
- Levator Labii Superioris Alaeque Nasi: These muscles contribute to raising the upper lip and the nostrils.
- Mentalis: Located in the center of the chin, they control the movement of the lower lip.
- Nasalis: These muscles help in opening the nostrils.
- Occipitofrontalis: Positioned above the eyebrows and extending to the crown of the head, they lift the eyebrows and allow the raising of the corners of the mouth.
- Orbicularis oculi: They control eyelid movement.
- Orbicularis oris: Muscles encircling the lips, controlling lip opening and closing.
- Procerus: Located between the eyebrows, they are involved in controlling the eyebrows and nose bridge.
- Risorius: Positioned on the sides of the mouth, they contribute to smiling.
- Zygomaticus Major and Minor: They play a role in smiling.

E. Resnet50

The term "ResNet" is derived from "Residual Network," and it represents a specific category of Convolutional Neural Network (CNN) as introduced in the paper titled "Deep Residual Learning for Image Recognition," authored by He Kaiming, Zhang Xiangyu, Ren Shaoqing, and Sun Jian in 2015. CNN are widely employed in various computer vision applications [13-15].

ResNet-50, a prominent instance, is a convolutional neural network comprising fifty layers, which encompasses forty-eight convolutional layers, one MaxPool layer, and one

average pool layer. Residual neural networks belong to the realm of Artificial Neural Network (ANNs), characterized by their construction through the stacking of residual blocks.

The original ResNet model, called ResNet-34, was upgraded for better performance. Nowadays, we have ResNet-50, ResNet-101, ResNet-152 are popular, but for this study, the smaller ResNet-50 is the best option due to our limited data. ResNet-50 works well for this research because our dataset is not very large. In comparison, ResNet-101 and ResNet-152 need more computer storage and processing power, which goes beyond what this paper needs.

F. Related Documents and Research

Wacharachai Kongsiriwatthanakul et al. [16] conducted a research study on a system that monitors classroom attendance and assesses learners' interest through facial expressions using a webcam. This study focused on monitoring attendance and capturing the emotions of learners during on-site learning. They categorized emotions into two groups: the first group included emotions related to interest in learning, such as neutral, happy, and surprised, while the second group included emotions indicating disinterest in learning, such as angry, sad, and fear.

III. RESEARCH METHODOLOGY

A. Limitation

- The sample group consists of undergraduate learners from the Thai-Nichi Institute of Technology.
- Learning a data science lesson in twenty minutes using video recordings from a data science class. The learners' comprehension is assessed through pre- and post-learning quizzes.
- Learners are required to use a desktop computer or laptop with the following specifications:
- Capable of internet connectivity.
- Equipped with a camera capable of capturing images of the learner.
- Storage space on the computer of more than one GB.

B. Working Process

Online learning system has been developed to address these issues. The system aims to incorporate a solution by leveraging technology. While learners are engaged in their virtual classrooms, the system will employ facial detection to

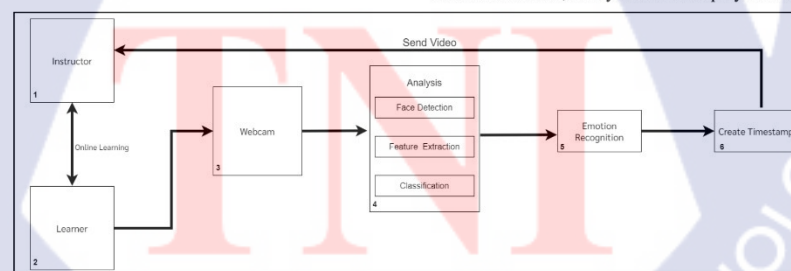


Fig. 1. Working process of emotion detection program.

analyze and detect learners' emotions. This feature will allow instructors to assess the level of satisfaction among learners.

The system is composed of six components, each with specific roles and functionalities, as described below and shown in Fig. 1.

1) "Learner": Individuals who have registered for online learning

2) "Instructor": A pivotal element of the system, responsible for creating and designing the curriculum, content, assessments, and evaluations for online learning.

3) "Webcam": Video footage is recorded using a webcam. The camera captures the specific perspectives of learners participating in the virtual class. The system tries to identify regions of interest (ROIs) such as facial expressions, postures, gestures, or gazes of the learners during the learning process.

4) "Analysis": This part consist of three subcomponents as "Face Detection", "Feature Extraction", and "Classification". This "Analysis" component is responsible for analyzing the captured data. It detects and interprets learners' facial expressions and emotions based on the recorded video.

In this research, we use CNN for "Feature Extraction" and "Classification" to recognize emotions. In ResNet-50, there are various layers combined, including:

- Input layer to receive image data with dimensions (224,224,3), meaning the image is 224 pixels wide, 224 pixels long, and uses the RGB color model. Convolution and Pooling Layers : We using Convolutional layers (Conv2D) apply filters to the input to create feature maps.
- Residual Blocks: The model has special blocks called residual blocks which are the key innovation of ResNet architectures. Each block has layers that help process the data: convolutional layers, batch normalization to adjust the data, and activation functions (ReLU) to add complexity.
- Batch Normalization and Activation: Batch Normalization layers normalize the activations of the previous layer at each batch, stabilizing and accelerating the training process. Activation functions (ReLU) introduce non-linearities to the model, allowing it to learn complex patterns in the data.
- Pooling Layer: After the last residual block, there is a MaxPooling2D layer that further reduces the spatial dimensions to (28, 28, 256).

For training the model, we use FER-2013 dataset [17]-[18]. This dataset collects human facial expression images. The data was gathered from the Kaggle website and is freely accessible. It comprises a total of 35,887 images, as indicated in Table I.

5) "Emotion Recognition": The system is designed to recognize emotions from learners' facial expressions, aiming to provide insights into their emotional states during learning.

6) "Create Timestamp" is shown Time Logging into ".txt" file: The system logs the instances of detected changes in

TABEL I. THE NUMBER OF IMAGES IN THE FER-2013 DATASET.

Emotion	Training data	Test data	Train and test data
Angry	3,995	958	4,953
Disgust	436	111	547
Fear	4,097	1,024	5,121
Happy	7,215	1,774	8,989
Neutral	4,965	1,233	6,198
Sad	4,830	1,247	6,077
Surprise	3,171	831	4,002
Total	28,709	7,178	35,887

emotions during the video. This information is stored in a text file, indicating the timestamps of emotional fluctuations.

Each component contributes to the functionality of the system in the following manner, encompassing their respective roles.

IV. EXPERIMENTAL RESULTS

A. Model Accuracy

The model was trained using the ResNet50 architecture, a pre-trained convolutional neural network (CNN) framework.

This training utilized the FER-2013 dataset, which categorizes facial expressions into seven distinct emotional states: 1) angry, 2) disgust, 3) fear, 4) happy, 5) neutral, 6) sad, and 7) surprise as shown in Fig.2

These results demonstrate the model's capability in accurately classifying emotional states based on the FER-2013 dataset and are suitable for inclusion in a research paper.



Fig 2. Seven emotions detection output.

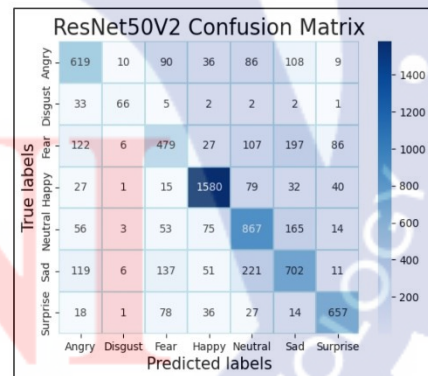


Fig 3. Seven emotions model's confusion matrix.

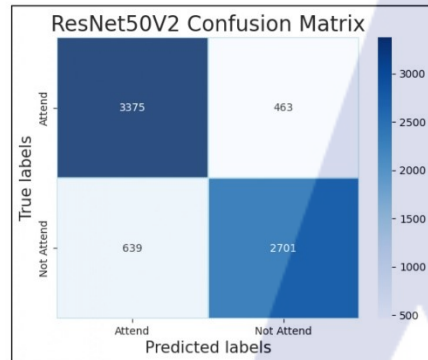


Fig 4. Two emotions model's confusion matrix.

This model has accuracy = 0.692, precision = 0.690, recall = 0.692, and F1 score = 0.690. This model's confusion matrix is in Fig. 3.

From the seven emotions detection model, we classify the emotion into two groups [16,19]. 1) "Attend" group has happy, neutral, and surprise. 2) "Not attend" group has angry, disgust, fear, and sad. In this model, the accuracy is 0.846, precision = 0.879, recall = 0.841 and F1 score = 0.860. The confusion matrix for this model is shown in Fig. 4.

B. Using program for detect

We use the basic of data science in this experiment. In this experiment we divided the video into six parts.

Part 1 (0.00 – 3.00 minutes): Introduction of data science course

Part 2 (3.01 – 8.45 minutes): Describe the meaning and the difference of data science and machine learning.

Part 3 (8.46 – 11.30 minutes): Describe the type of data that can be used in machine learning.

Part 4 (11.31 – 15.53 minutes): Describe what is AI company and show an example.

Part 5 (15.54 – 19.21 minutes): Describe machine learning and show the example.

Part 6 (19.22 – 20.00 minutes): Introduction to next chapter.

In the study, eight learners from Thai-Nichi Institute of Technology were part of the research. However, there was a problem with one learner's data. We could not use their information about how much they paid attention to different topics (shown in Table I). So, that data was not included in our final analysis. Table II shows the times when the face detection was successful.

Table II shows the calculation of the percentage of time learners were attentive to their learning. This was determined by dividing the seconds they were focused by the seconds the system could detect their faces. The results represent the percentage of time learners were engaged in learning. Based on the findings of the experiment, it is apparent that learners no. 2 and no. 6 showed a high level of attention. Abnormalities in the data were observed for learner no. 7 in parts 1 (13.04%) and 2 (3.37%), learner no. 4 in part 5

TABEL II. PERCENTAGE OF LEARNERS' ATTENTION (%)

LEARNER No.	PART 1	PART 2	PART 3	PART 4	PART 5	PART 6
1	-	-	-	-	94.7	78.9
2	100	100	100	100	100	100
3	97.90	95.00	95.60	97.67	96.60	100
4	94.36	49.86	40.63	39.24	19.67	89.90
5	93.75	78.14	57.26	49.32	55.23	71.23
6	100	100	100	100	100	100
7	13.04	3.37	54.83	73.33	63.63	61.53
8	89.69	71.83	71.55	69.64	9.25	66.67
AVERAGE	84.11	71.17	74.27	76.60	67.39	83.53

TABEL III. PERCENTAGE OF LEARNERS FACES DETECTION (%)

LEARNER No.	PART 1	PART 2	PART 3	PART 4	PART 5	PART 6
1	0	0	0	0	5.26	8.57
2	100	100	100	100	100	100
3	82.78	97.96	95.12	98.09	97.60	100
4	98.33	30.23	26.83	88.54	84.21	89.47
5	80.00	81.10	71.34	96.94	89.47	86.84
6	100	100	100	99.07	99.04	100
7	12.66	44.47	18.90	5.72	10.52	34.21
8	53.89	41.27	66.46	42.74	25.83	63.15
AVERAGE	65.96	61.88	59.83	66.39	63.99	72.78

(39.24%), and learner no. 8 in part 5 (9.25%), indicating "not attend". On average, learners' interest levels in different parts were as follows: part 1 (84.11%), part 2 (71.17%), part 3 (74.27%), part 4 (76.60%), part 5 (67.39%), and part 6 (83.53%).

Table III shows the calculation of the percentage of time when the system could detect learners' faces. This was determined by dividing the seconds the system successfully detected faces by the total seconds in that specific part. The results represent the percentage of time the system was able to detect learners' faces. Based on the results of the experiment, it becomes evident that learner no. 1 exhibited the lowest rate of detection. In parts 1-4, no faces were detected. In part 5, only 5.26% (11 seconds out of the total 3.27 minutes) of the time could the system capture faces. In part 6, the system detected faces for only 8.57% (3 seconds out of the total 38 seconds). Learner no. 7 had detection rates lower than 20% in parts 1, 3, 4, and 5. On average, the system's face detection rates in different parts were as follows: part 1 (65.96%), part 2 (61.88%), part 3 (59.83%), part 4 (66.39%), part 5 (63.99%), and part 6 (72.78%).

Learners no. 2 and no. 6 showed high interest in learning (100%), as shown in Table I. Similarly, Table II demonstrates their faces were detected consistently during learning sessions (99-100%). However, abnormal data was observed for learner no. 7, whose face detection started low in the initial parts (part 1-2) and increased in the later parts (part 4-6). However, the percentages of face detection in Table III did not align with this pattern. Particularly, in Part 2, the system detected faces up to 44.47% as shown in Table III, but it captured faces indicating interest in learning only at 3.37% as shown in Table II.

Regarding emotion detection, the model can function well. However, challenges in face detection due to the module's limitations. This was primarily due to low lighting conditions during the experiment and learners' hairstyles sometimes

obstructing their faces significantly. These issues can be improved upon in the future.

V. CONCLUSION AND FUTURE WORKS

We developed an emotion detection model capable of classifying emotions into seven emotions. These emotions were further grouped into "attend" and "not attend". Utilizing the FER-2013 dataset, which includes emotions as a 1) angry, 2) disgust, 3) fear, 4) happy, 5) neutral, 6) sad, and 7) surprise, we categorized them into two groups as "attend" and "not attend" emotions [16]-[19]. In an experiment involving eight children watching a 20-minutes video on data science, our program successfully detected facial expressions with an average accuracy exceeding 60%. Specifically, it accurately recognized focused facial expressions with an average accuracy of over 55%. After the class, the program automatically sends the video and its timestamp file via email to the professor. This system empowers instructors to effectively assess and adjust their teaching methods according to learner engagement levels.

In future developments, the system should be enhanced for easier program access and video recording in practical usage scenarios. Additionally, it can be further developed for detecting academic dishonesty during exams. There is also room to expand the system's capabilities by incorporating additional audio recording to assess the learning environment of the learners. Moreover, continuous development could enable the system to automatically summarize individual learners' learning performance.

REFERENCES

- [1] Parit Wacharasindhu, "The karma of the Thai education system has led to a proposal of 6 measures (6 Re's) to find a common solution during the COVID period." [Online]. Available : <https://workpointtoday.com/covid-policy-lab-education/> [10 September 2022]
- [2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [3] Martha C. L., and Prisca C. (2012). Emotional Presence, Learning, and the Online Learning Environment. The International Review of Research open and distance learning. Vol 13 (4).
- [4] I. Michael Revina, W.R. Sam Emmanuel, "A Survey on Human Face Expression Recognition Techniques", Journal of King Sand University - Computer and Information Sciences, Volume 33, Issue 6, 2021, Pages 619-628, ISSN 1319-1578
- [5] Wafa Mellouk, Wahida Handouzi, "Facial emotion recognition using deep learning: review and insights", Procedia Computer Science, Volume 175, 2020, Pages 689-694, ISSN 1877-0509M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
- [6] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-Based Learning Applied to Document Recognition. In: Proceedings of the IEEE. 1998, pp. 2278–2324.
- [7] Le Cun Y, Matan O, Boser B, Denker JS, Henderson D, Howard RE, et al. Handwritten Zip Code Recognition with Multilayer Networks. In: 10th International Conference on Pattern Recognition. 1990, pp. 35–40.
- [8] Raksaard N and Surinta O. Comparative Study Between Local Descriptors and Deep Learning for Silk Pattern Image Retrieval, Journal of Science and Technology Mahasarakham University, 37(6), 2018, pp. 736–746.
- [9] Sanuksan J and Surinta O. Deep Convolutional Neural Networks for Plant Recognition in Natural Environment, Journal of Science and Technology Mahasarakham University, 38(2), 2019, pp. 113–124.
- [10] Chitpong Kittinaradom, "Convolutional Neural Network" [Online]. Available : <https://guopai.github.io/ml-blog/19.html> [12 October 2022]
- [11] The Nature of Things "The Seven Universal Emotions We Wear On Our Face" [Online]. Available : <https://www.cbc.ca/natureofthings/features/the-seven-universal-emotions-we-wear-on-our-face> [19 October 2022]
- [12] Cleveland Clinic "Facial Muscles" [Online]. Available : <https://my.clevelandclinic.org/health/body/21672-facial-muscles> [18 October 2022]
- [13] Datagen "ResNet-50: The Basics and a Quick Tutorial" [Online]. Available : <https://datagen.tech/guides/computer-vision/resnet-50/#> [12 September 2023]
- [14] Aliyari Mostafa et. al, "UAV-Based Bridge Inspection via Transfer Learning." [Online]. Available : https://www.researchgate.net/publication/355249492_UAV-Based_Bridge_Inspection_via_Transfer_Learning [20 October 2023]
- [15] Cabezas, Mariano et al., "Detection of Invasive Species in Wetlands: Practical DL with Heavily Imbalanced Data. Remote Sensing" [Online]. Available : https://www.researchgate.net/publication/345393748_Detection_of_Invasive_Species_in_Wetlands_Practical_DL_with_Heavily_Imbalanced_Data [20 October 2023]
- [16] Watcharachai Kongsiriwattana et al., "Student Class Attendance and Interest Assessment System with Facial Expression Detection via Webcam" [Online]. Available : https://i01.tci-thaijo.org/index.php/science_kmitl/article/download/248748/172046/8872 [20 October 2023]
- [17] Manas Sambare "FER-2013" [Online]. Available : <https://www.kaggle.com/datasets/msambare/fer2013> [19 February 2023]
- [18] Ian J. Goodfellow et al. "Challenges in representation learning: A report on three machine learning contests", Neural Networks, Volume 64, 2015, Pages 59–63, ISSN 0893-6080
- [19] T. Raksarikorn and T. Kangkachit, "Facial Expression Classification using Deep Extreme Inception Networks," 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE), Nakhonpathom, Thailand, 2018, pp. 1–5, doi: 10.1109/JCSSE.2018.8457396.



ประวัติย่อผู้วิจัย



ชื่อ – สกุล

นางสาวภูริชญา โชคเศรษฐกิจ

วัน เดือน ปีเกิด

6 มิถุนายน 2543

ที่อยู่ปัจจุบัน

126/1199 ถนนติวานนท์-ปากเกร็ด 15
ตำบลปากเกร็ด อำเภอปากเกร็ด จังหวัดนนทบุรี

ประวัติการศึกษา

พ.ศ. 2566

วิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีไทย-ญี่ปุ่น

พ.ศ. 2563

วิทยาศาสตรบัณฑิต สาขาเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีไทย-ญี่ปุ่น

พ.ศ. 2563

บริหารธุรกิจบัณฑิต สาขาการเงินและการธนาคาร
มหาวิทยาลัยรามคำแหง

ประวัติการทำงาน

พ.ศ. 2563 – ปัจจุบัน

Data Engineer
MFEC Public Company Limited