

**N<sub>x</sub>**

**AI**

**สำหรับทุกคน**

**เริ่มต้นกับ**

**ปัญญาประดิษฐ์ได้**

**ภายใน 1 เล่ม**

**N<sub>x</sub>**

**AI**

**สำหรับทุกคน**

**เริ่มต้นกับ**

**ปัญญาประดิษฐ์ได้**

**ภายใน 1 เล่ม**

# AI สำหรับทุกคน

เริ่มต้นกับปัญญาประดิษฐ์ได้ภายใน 1 เล่ม

โดย

NEXTCODECAMP

“Next learning, next skill.”

สงวนสิทธิ์ตามพระราชบัญญัติลิขสิทธิ์ 2537

ห้ามคัดลอกเนื้อหาก่อนได้รับอนุญาต

ข้อมูลทางบรรณานุกรมหนังสือ

กฤดาภัตร์ สีหารี,

AI สำหรับทุกคน เริ่มต้นกับปัญญาประดิษฐ์ได้ภายใน 1 เล่ม, 2569.

305 หน้า.

ISBN (e-book): 978-616-630-614-9

1. ปัญญาประดิษฐ์ 2. AI 3. Artificial Intelligence

หนังสือ E-Book จัดจำหน่ายออนไลน์

จัดทำโดย

รองศาสตราจารย์ ดร. กฤดาภัตร์ สีหารี

19/76 ซอยแจ้งวัฒนะ-ปากเกร็ด 41

ตำบลคลองเกลือ อำเภอปากเกร็ด

จังหวัดนนทบุรี 11120

# คำนำ

ปัจจุบันทุกสาขาอาชีพต้องปรับตัวกับเทรนที่มาแรงในยุคนี้ นั่นก็คือ AI ภาพของการนำไปใช้ในทุก ๆ โดเมนเริ่มมีความจริงจัง และได้รับการยอมรับอย่างแพร่หลายมากขึ้นเรื่อย ๆ การสร้างความเข้าใจ และเรียนรู้ AI ในการประยุกต์ใช้ อย่างมีประสิทธิภาพเป็นสิ่งที่เราทุกคนควรบริหารจัดการ

ผู้เขียนพยายามเรียบเรียงและถ่ายทอดหลักการพื้นฐานโดยผู้อ่านทั่วไป หรือ นิสิตและนักศึกษาปริญญาตรีขึ้นไป รวมถึงผู้ที่อยากจะจริงจังกับ AI สามารถอ่านได้ ผู้เขียนได้รวบรวมโค้ดตัวอย่างซึ่งสร้าง/ปรับด้วย AI และทดสอบ/ปรับแล้วโดยมนุษย์ (ผู้เขียน) สามารถนำไปรันเพื่อการศึกษาได้ใน Google Colab บทที่ 1-4 เป็นเนื้อหาพื้นฐานสำหรับการเรียนรู้เบื้องต้นทั่วไป บทที่ 5-10 เข้มข้นขึ้นสำหรับผู้ที่จริงจังกับ AI และบทที่ 11 เป็นเรื่องเกี่ยวกับจริยธรรมปัญญาประดิษฐ์ที่ควรรู้

หนังสืออาจมีเนื้อหาค่อนข้างมาก หากมีข้อผิดพลาดประการใดต้องขออภัยมา ณ ที่นี้ และท่านสามารถติชมหรือแนะนำได้ที่ [nextcodecamp@gmail.com](mailto:nextcodecamp@gmail.com)

สุดท้ายนี้หวังเป็นอย่างยิ่งว่าใคร ๆ ก็อ่านได้สมกับชื่อเรื่องของหนังสือเล่มนี้

รศ.ดร. กฤดาภัทธ สีสารี

1 กุมภาพันธ์ 2569

## สารบัญ

คำนำ	i
บทที่ 1 แนะนำ AI	1
1.1 ความสำคัญของ AI ในโลกปัจจุบัน	1
1.2 ChatGPT	3
1.3 GEMINI	5
1.3.1 การวิเคราะห์ข้อมูลหลายรูปแบบ	5
1.3.2 การช่วยวิเคราะห์เอกสาร	7
1.3.3 ผู้ช่วยเขียนโค้ด	7
1.3.4 ผู้ช่วยด้านการเรียนรู้	8
1.3.5 ผู้ช่วยคอนเทนต์อัจฉริยะ	8
1.4 Roboflow	10
1.5 การใช้ AI ในธุรกิจ	12
1.5.1 การตลาด	12
1.5.2 การเงิน (Finance)	14
1.5.3 การบริการลูกค้า	15
1.5.4 การจัดการห่วงโซ่อุปทาน	16
1.6 การใช้ AI ในชีวิตประจำวัน	17
1.6.1 การเดินทางและการคมนาคม	18
1.6.2 สุขภาพและการออกกำลังกาย	19
1.6.3 ความบันเทิงและสื่อ	20
1.6.4 บ้านอัจฉริยะ	21
1.6.5 การใช้ AI ในการสื่อสารและภาษา	22
1.6.6 ข้อควรระวังในการใช้ AI ในชีวิตประจำวัน	23

บทที่ 2 เครื่องมือในการเขียนโปรแกรม	25
2.1 การเตรียมเครื่องมือและการเตรียมตัว	25
2.2 Google Colab	26
2.3 เข้าใช้งาน Google Colab	27
2.4 การอัปโหลดไฟล์ลงใน Google Drive	35
บทที่ 3 การเรียนรู้ของเครื่องจักร	39
3.1 การเรียนรู้ของเครื่องจักร	39
3.1.1 การเรียนรู้แบบมีผู้สอน	39
3.1.2 การเรียนรู้แบบไม่มีผู้สอน	40
3.1.3 การเรียนรู้แบบเสริมแรง	41
3.2 สถานการณ์กับการนำไปใช้	43
3.3 ปัญหาเกี่ยวกับการเลือกเทคนิค	45
3.3.1 ระบบแนะนำสินค้า	45
3.3.2 รถยนต์ไร้คนขับ	48
3.3.3 การเงินและการลงทุน	52
3.3.4 การแพทย์	57
3.3.5 การเกษตร	61
3.3.6 พลังงาน	66
3.3.7 เมือง	72
บทที่ 4 เทคนิคในการเรียนรู้ของเครื่องจักร	77
4.1 เทคนิคและโมเดล	77
4.2 เทคนิคการถดถอย	81
4.1.1 การถดถอยเชิงเส้น	81
4.1.2 การถดถอยด้วยป่าสุ่ม	86

4.3	เทคนิคการจำแนก	90
4.3.1	การจำแนกประเภทด้วยการถดถอยโลจิสติก	90
4.3.2	การจำแนกประเภทด้วยต้นไม้ตัดสินใจ	94
4.3.3	การจำแนกประเภทด้วยป่าสุ่ม	97
4.3.4	K-Nearest Neighbors	104
4.4	เทคนิคการจัดกลุ่ม	107
4.4.1	K-Means	107
4.4.2	DBSCAN	110
4.5	เทคนิคการลดมิติ	114
4.5.1	PCA	114
4.5.2	LDA	118
4.6	เทคนิคการเชื่อมโยงความสัมพันธ์	121
4.6.1	Apriori	121
4.6.2	FP-Growth	124
4.7	เทคนิคเอนเซมเบิล	126
บทที่ 5	การเรียนรู้เชิงลึก	131
5.1	Artificial Neural Network	131
5.2	Recurrent Neural Network	138
5.3	Long Short-Term Memory	142
5.4	โครงข่ายประสาทเทียมแบบคอนโวลูชัน	147
บทที่ 6	ข้อมูลและการประเมินโมเดล	153
6.1	ความสำคัญของข้อมูลกับโมเดล	153
6.2	แหล่งข้อมูลสำหรับการฝึกโมเดล AI	155
6.3	การติดป้ายกำกับ	156
6.4	วิธีให้น้ำหนักกับความเชื่อถือของข้อมูล	156

6.5 ตัวอย่างการให้น้ำหนักของข้อมูล	159
6.6 ข้อคำนึงถึงเรื่องคุณภาพข้อมูล	162
6.7 ตัวอย่างการวิเคราะห์คุณภาพข้อมูล	163
6.8 การปรับสมดุลข้อมูล	165
6.9 ข้อควรพิจารณาสำหรับการเรียนรู้แบบไม่มีผู้สอน	167
6.10 ข้อมูลในการเรียนรู้แบบเสริมแรง	171
6.11 การประเมินผล โมเดล	172
6.11.1 การประเมินผล โมเดลสำหรับการจำแนก	172
6.11.2 การประเมินผล โมเดลถดถอย	179
6.11.3 การประเมินผลการเรียนรู้แบบเสริมแรง	182
บทที่ 7 AI กับภาพ	185
7.1 ภาพดิจิทัลคืออะไร?	185
7.2 AI มองภาพอย่างไร?	188
7.3 การสกัดคุณลักษณะ	191
7.3.1 การสกัดคุณลักษณะด้วยการตรวจจับขอบ	193
7.3.2 การสกัดคุณลักษณะเชิงรูปร่าง	197
7.3.3 การสกัดคุณลักษณะด้วย Color Histograms	200
7.4 แนวทางการนำคุณลักษณะไปใช้งานกับ โมเดล AI	201
7.4.1 การนำคุณลักษณะที่สกัดมาใช้โดยตรง	201
7.4.2 การนำคุณลักษณะไปใช้แบบผสม	202
บทที่ 8 CNN กับ โมเดลเชิงกำเนิด	205
8.1 โครงข่ายประสาทเทียมแบบคอนโวลูชัน	205
8.1.1 Convolution Layer	206
8.1.2 Pooling Layer	207

8.1.3 Flatten Layer	208
8.1.4 Fully Connected Layer	208
8.2 การแก้ปัญหาด้วย CNN	209
8.2.1 การจำแนกภาพ	209
8.2.2 การตรวจจับวัตถุ	213
8.2.3 การแบ่งส่วนของภาพ	216
8.3.4 การสร้างภาพใหม่	219
8.2.5 การสกัดคุณลักษณะใน CNN	220
8.3 โมเดลเชิงกำเนิด	223
8.3.1 Autoencoder	223
8.3.2 Variational Autoencoder	224
8.3.3 Adversarial Autoencoder	228
8.3.4 GAN	229
8.3.5 Diffusion Model	232
8.4 การประมวลผลข้อมูลที่ไม่ใช่ภาพ	234
8.4.1 CNN กับข้อมูลเสียง	234
8.4.2 CNN กับข้อความ	236
8.5 แนวคิดของ CNN ที่แท้จริง	237
บทที่ 9 AI กับ Text	239
9.1 เท็กซ์	239
9.2 BERT	241
9.3 Hugging Face	244
9.4 Fine-tuning	245
9.4.1 ตัวอย่างการทำ Fine-tuning ในงาน NER	247
9.4.2 ปัญหาในการทำ Fine-tuning	251

9.5 ตัวอย่างโค้ดของงานที่ BERT สามารถทำงานได้	252
9.5.1 การวัดความคล้ายของคำ	252
9.5.2 การหาคำที่ถูกต้อง	254
9.5.3 คัดแยกอีเมลสแปม	255
9.5.4 วิเคราะห์ความขัดแย้งของประโยค	259
บทที่ 10 AI กับ Art	263
10.1 AI กับ Art	263
10.2 AI Style Transfer	265
10.3 AI กับการเคลื่อนไหว	266
10.3.1 การถ่ายโอนการเคลื่อนไหว	267
10.3.2 การถ่ายโอนสไตล์	270
10.4 การสร้างสรรค์งานศิลปะด้วยเครื่องมือ AI	271
10.4.1 การออกแบบลายผ้า	271
10.4.2 การออกแบบเครื่องครัวเซรามิก	276
10.4.3 การออกแบบฟอนต์	277
10.4.4 การออกแบบงานไม้แกะสลักลายไทย	278
10.5 จริยธรรม กับ AI Art	280
10.5.1 ประเด็นเรื่องลิขสิทธิ์ของภาพ	280
10.5.2 ประเด็นเรื่องภาพใหม่ที่ AI สร้างขึ้น	281
10.5.3 ความเสี่ยงเรื่องสิทธิภาพบุคคล (Personality Rights)	282
10.5.4 การฟ้องร้องสามารถเอาผิดกับใครได้บ้าง	283
10.6 แนวทางปฏิบัติในการใช้ภาพคนจริงฝึก AI	284
บทที่ 11 จริยธรรมปัญญาประดิษฐ์	287
11.1 จริยธรรมกับปัญญาประดิษฐ์	287

11.2 แนวคิดพื้นฐานในจริยธรรมที่ใช้ใน AI	288
11.2.1 แนวคิดพื้นฐานสากล	288
11.2.2 แนวคิดพื้นฐานประเทศไทย	291
11.3 ปัญหาทางจริยธรรมใน AI	292
11.3.1 ความลำเอียงและความยุติธรรม	292
11.3.2 ความเป็นส่วนตัวและการปกป้องข้อมูล	293
11.3.3 ความรับผิดชอบ	294
11.3.4 ความโปร่งใสและการอธิบายได้	295
11.3.5 ผลกระทบต่อแรงงาน	296
11.3.6 การปลอมแปลงข้อมูล	297
11.4 แนวปฏิบัติเพื่อส่งเสริมความสำคัญของจริยธรรม AI	299
11.4.1 การพัฒนา Sandbox	299
11.4.2 แนวทาง AI Audit Template	301
บรรณานุกรม	305

# บทที่ 1

## แนะนำ AI

### 1.1 ความสำคัญของ AI ในโลกปัจจุบัน

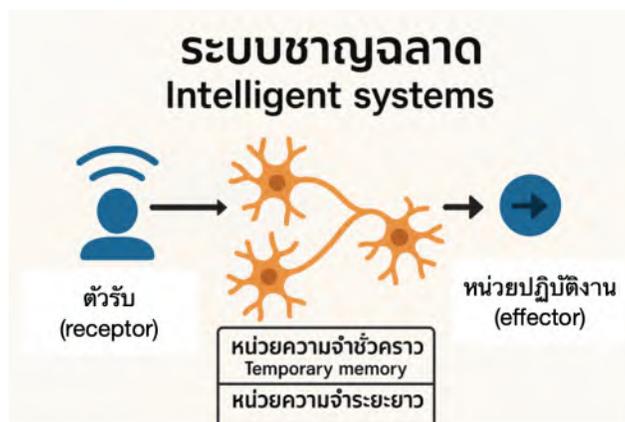
ปัญญาประดิษฐ์ (Artificial Intelligence) หากพูดให้เข้าใจง่าย ๆ ก็คือ การสร้างให้คอมพิวเตอร์คิดอย่างมนุษย์ อย่างไรก็ตาม เราอาจจะศึกษานิยามจากนักวิชาการเพื่อสร้างความเข้าใจให้ลึกซึ้งขึ้น ยกตัวอย่างเช่น

Turing (1950) ให้นิยามว่า AI คือระบบที่สามารถแสดงพฤติกรรมได้จนมนุษย์ไม่สามารถแยกออกได้ว่ากำลังสื่อสารกับมนุษย์หรือเครื่องจักร

John McCarthy และคณะ (1956) ให้นิยามปัญญาประดิษฐ์ว่าเป็นสาขาวิชาที่มุ่งศึกษาการทำให้เครื่องจักรสามารถแสดงพฤติกรรมที่ต้องอาศัยความฉลาดในระดับเดียวกับมนุษย์

Russell และ Norvig (1995) ให้นิยามว่าเป็นการศึกษาและออกแบบระบบอัจฉริยะที่สามารถรับรู้สิ่งรอบตัว และตัดสินใจเพื่อให้บรรลุเป้าหมายได้อย่างเหมาะสม เขาได้วางกรอบแนวคิดของ AI ไว้ 4 อย่าง คือ คิดเหมือนมนุษย์ แสดงพฤติกรรมเหมือนมนุษย์ คิดอย่างมีเหตุผล และกระทำอย่างเหมาะสม

McCulloch และ Walter (1943) ได้บรรยายปัญญาประดิษฐ์ว่าเป็นความรู้แขนงหนึ่งซึ่งศึกษาเกี่ยวกับโครงสร้าง และฟังก์ชันการทำงานของระบบชาวนฉลาด (Intelligent systems) มีการทำงานด้วยระบบประสาทซึ่งมีตัวรับ (receptor) และหน่วยปฏิบัติการ (effector) เชื่อมโยงเป็นเครือข่าย มีหน่วยความจำชั่วคราว และหน่วยความจำระยะยาว (ภาพที่ 1.1)



ภาพที่ 1.1 ระบบชาญฉลาด

การศึกษาต่าง ๆ ในปัญญาประดิษฐ์จึงคล้ายกับการศึกษาการทำงานของสมอง และหน่วยรับรู้โดยศึกษาไปที่การสะท้อนตอบต่อสิ่งที่มีเงื่อนไข และไม่มีเงื่อนไข การอนุมานหาข้อสรุป และการแสดงพฤติกรรมของระบบเมื่อเรียนรู้หรือได้รับการฝึกฝน อาจกล่าวกว้าง ๆ ว่าปัญญาประดิษฐ์เป็นศาสตร์ด้านหนึ่งในวิทยาการคอมพิวเตอร์ที่สอนให้เครื่องจักรหรือคอมพิวเตอร์สามารถทำงานได้เสมือนเป็นมนุษย์ โดยมีสมองกลเป็นองค์ประกอบหลักในการทำงาน

งานที่สามารถประยุกต์ปัญญาประดิษฐ์ได้ก็คือ งานที่มุ่งเน้นให้คอมพิวเตอร์สามารถทำงานได้แทนมนุษย์ เช่น การทำงานที่มีการทำซ้ำ ๆ และงานที่สามารถตัดสินใจได้โดยใช้เงื่อนไข

การศึกษปัญญาประดิษฐ์เกี่ยวข้องกับสาขาวิชาความรู้ในหลาย ๆ เรื่อง อาทิ

- ระบบผู้เชี่ยวชาญ (Expert System)
- การเรียนรู้ของเครื่องจักร (Machine Learning)
- การอนุมานด้วยตรรกะ (Logic reasoning) และ
- วิศวกรรมความรู้ (Knowledge engineering)

การเรียนรู้ของเครื่องจักร (Machine Learning) เป็นด้านที่มีการประยุกต์ใช้กับระบบงานต่าง ๆ อย่างกว้างขวาง ไม่ว่าจะเป็นด้านธุรกิจ การแพทย์ การประมวลผลภาพ และการวิเคราะห์ข้อความ กระบวนการเรียนรู้ของเครื่องจักรต้องอาศัยโมเดลซึ่งมีให้เลือกใช้หลาย ๆ แบบ และโมเดลที่เป็นที่รู้จักกันอย่างกว้างขวางคือ โครงข่ายประสาทเทียมหรือนิวรัลเน็ตเวิร์ก (Neural Network) ซึ่งเป็นโมเดลพื้นฐานของการเรียนรู้เชิงลึก (Deep Learning)

ในปัจจุบันมีผู้ให้บริการ AI แพลตฟอร์มต่าง ๆ อย่างมากมาย บางแพลตฟอร์มอาจให้บริการเจาะจงในบางเรื่อง เช่น เรื่องภาพ เรื่องวิดีโอ และเรื่องงานวิจัย บางแพลตฟอร์มเป็นแพลตฟอร์มที่สามารถทำงานได้ทั่วไป ซึ่งก็มักจะมีความสามารถหลากหลายในการประมวลผลภาพ วิดีโอ เสียง และข้อความ อาทิ Gemini, GhatGPT และ Claude ในฐานะของผู้ใช้อาจต้องพิจารณาว่าจะสามารถนำแพลตฟอร์มเหล่านั้นมาใช้ประโยชน์ได้อย่างไร หรือจะนำมาพัฒนาหรือช่วยงานในด้านใด ในหัวข้อต่อไปนี้จะได้กล่าวถึงความสำคัญของ AI Platform ที่สามารถทำงานได้ทั่วไป อาทิ ChatGPT และ Gemini นอกจากนี้แล้วได้ยกตัวอย่าง Roboflow ซึ่งให้บริการในการสร้างโมเดล AI ที่เกี่ยวข้องกับคอมพิวเตอร์วิทัศน์ (Computer Vision) โดยเฉพาะ

## 1.2 ChatGPT

ปัจจุบันคงปฏิเสธไม่ได้ว่า ChatGPT นั้นเก่งเหลือเกิน ผู้เขียนเองเขียนโปรแกรมคอมพิวเตอร์ได้ก็ได้ลองทดสอบให้ ChatGPT ช่วยสร้างโค้ดให้ ผลที่ได้คือไม่ต้องเหนื่อย และเสียเวลามาก ChatGPT ไม่ใช่ AI ที่เก่งแต่เรื่องเขียนโปรแกรมในหลาย ๆ เรื่องหรือโดเมนอื่น ๆ ก็ทำได้ดี ChatGPT เน้นค้นหาข้อมูลที่มีความหลากหลาย และเอาข้อมูลนั้นมาผสมรวมกัน มีการเทรนหรือฝึกสอนให้ระบบเกิดการ

เรียนรู้ สามารถสร้างเนื้อหาเพิ่มเติมได้เอง มีการกลั่นกรองข้อมูลที่ดี ความแม่นยำที่ได้จากการค้นหาข้อมูลนั้นก็ค่อนข้างดีทีเดียว

นอกจากนี้แล้ว สิ่งที่น่าสนใจของ ChatGPT คือ การค้นหาในรูปแบบของการสนทนา มันทำให้ผู้ใช้งานรู้สึกว่ามีปฏิสัมพันธ์ได้ไม่ยาก การเลือกใช้คำในการตอบที่มีวิธีการใช้คำอันนุ่มนวลและสุภาพ ทำให้ผู้ใช้งานรับรู้ได้ถึงอารมณ์ อันนี้อาจจะเป็นเทคนิคที่ดึงดูดใจให้หลาย ๆ คนเปลี่ยนใจค้นกับ Google มาใช้บริการของ ChatGPT ซึ่งแน่นอนว่าบริษัทยักษ์ใหญ่ต่าง ๆ (โดยเฉพาะ Google ซึ่งเสียฐานผู้ใช้ไปแล้วบางส่วน) พยายามทำระบบ และปล่อยบริการต่าง ๆ ให้เราได้ทดลองเพื่อแย่งส่วนแบ่งการตลาดขั้นโต

ความเก่งของ ChatGPT คือการประมวลผลข้อความ (Text processing) ด้วยการประมวลผลภาษาธรรมชาติ (Natural Language Processing) ทำให้ผู้ใช้งานสื่อสารกับระบบด้วยการกำหนดคำค้นที่เป็นภาษาของมนุษย์ อีกทั้งความสามารถในการประมวลผลด้วยการตีความหมายของคำจึงทำให้การค้นหาข้อมูลมีความแม่นยำมาก ๆ

คำถามที่ว่า ChatGPT จะมาแย่งงานเราหรือจะมาช่วยงานเรา อาจมีคำตอบทั้งจริงและไม่จริง ในส่วนที่จริงนั้น ChatGPT สามารถทำงานแทนเราได้ ยกตัวอย่างง่าย ๆ สมมติว่าต้องการให้ ChatGPT แปลจดหมายจากภาษาไทยเป็นภาษาอังกฤษ หากให้มนุษย์ทำเราก็อาจจะต้องเชื่อมั่นว่าภาษาของเรานั้นดี บางคนที่ขาดความมั่นใจอาจจะให้ผู้ที่มีความรู้ช่วยกลั่นกรองให้ เมื่อคำนวณเวลารวม ๆ ในการจัดการก็อาจจะใช้เวลาหลายวัน ถ้าเราใช้ ChatGPT แปลภาษาให้ก็อาจจะเสร็จในเวลาไม่ถึงหนึ่งนาที การกลั่นกรองด้านภาษาให้ดูสะอาดสวยอาจจะต้องอาศัยคนที่เก่งภาษานั้น ๆ มาช่วยอ่าน อย่างไรก็ตาม เราคงไม่สามารถปฏิเสธได้ว่าการแปลของ ChatGPT ถึงแม้จะไม่สะอาดสวยมาก แต่ก็สามารถสื่อสารได้อย่างดีเลยทีเดียว

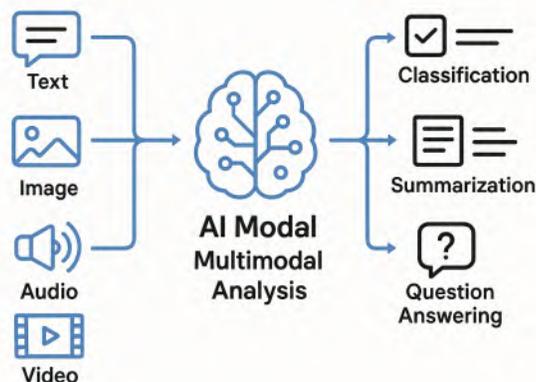
ผู้เขียนขอก้าวถึงความสามารถของ ChatGPT เพียงเท่านี้เพราะเข้าใจว่าหลาย ๆ คนน่าจะเคยใช้ และเห็นความสามารถอันหลากหลายมาแล้ว หรืออาจศึกษาในหนังสือ<sup>1</sup>ที่ผู้เขียนเคยถ่ายทอดมาแล้วเพื่อให้มองเห็นหลาย ๆ มิติในการใช้งาน

## 1.3 GEMINI

Gemini ของกูเกิลมีความสามารถที่หลากหลายไม่แพ้ ChatGPT ตัวอย่าง Use Case ที่ดีของ Gemini อาทิ

### 1.3.1 การวิเคราะห์ข้อมูลหลายรูปแบบ

การวิเคราะห์ข้อมูลหลายรูปแบบ (Multimodal Analysis) (ภาพที่ 1.2) สามารถประมวลผลข้อมูล เช่น ข้อความ ภาพ วิดีโอ และเสียง เข้าด้วยกัน



ภาพที่ 1.2 Multimodal Analysis

ความชาญฉลาดของ Gemini อาจนำไปใช้ในระบบ AI ที่พัฒนาขึ้นโดยเฉพาะ เช่น ระบบใช้ภาพจากกล้องวงจรปิดพร้อม ๆ กับการวิเคราะห์ข้อความรายงานเหตุการณ์ เพื่อสรุปว่าเกิดเหตุรถชนตรงไหน และมีใครเกี่ยวข้องบ้าง ผู้อ่านอาจลองจินตนาการการใช้ Google Map ในชีวิตประจำวัน ขณะขับรถบนถนนเส้นหนึ่ง อาจมี

<sup>1</sup> กฤดาภรณ์ สิริหารี, Generative AI, OpenAI API กับ Python สร้างแอปล้ำ ๆ จาก 0 ถึงโปร, 2568. 189 หน้า. ISBN (e-book): 978-616-623-347-6

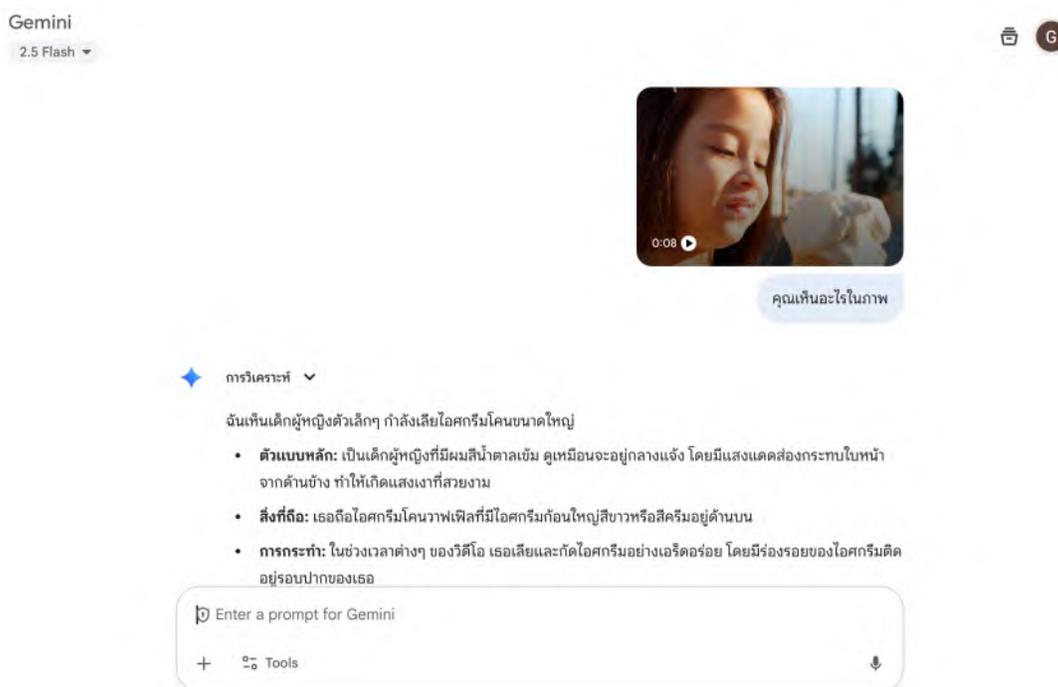
ผู้ใช้รถใช้ถนนส่งข้อความรายงานเหตุการณ์ไปยังเว็บไซต์ของหน่วยงานที่เกี่ยวข้อง ขณะเดียวกันอาจมีผู้บันทึกภาพเหตุการณ์และเผยแพร่ผ่านสื่อสังคมออนไลน์ หากระบบ AI อย่าง Gemini สามารถนำข้อมูลจากหลายแหล่งดังกล่าวมาประมวลผลร่วมกันได้ ก็ย่อมมีความเป็นไปได้ที่จะสรุปว่า ณ ช่วงเวลานั้นอาจเกิดเหตุการณ์บางอย่างขึ้นในตำแหน่งหนึ่งบนแผนที่ดิจิทัล ทั้งนี้ ข้อมูลการใช้งาน Google Map จากผู้ใช้จำนวนมากซึ่งเกิดขึ้นแบบเรียลไทม์ สามารถสะท้อนสภาพการจราจรหรือความผิดปกติในพื้นที่ได้ในระดับหนึ่ง (คล้ายระบบ Realtime Video Analytics<sup>2</sup> ในภาพที่ 1.3)



ภาพที่ 1.3 ระบบ Realtime Video Analytics

ผู้เขียนได้ทดลองใช้ Gemini 2.5 Flash เพื่อทำ Video Question Answering หรือการสอบถามระบบว่าในภาพหรือวิดีโอมีอะไรบ้าง ได้ลองอัปโหลดวิดีโอเด็กหญิงกำลังทานไอศกรีม (ภาพที่ 1.4) แล้วป้อนคำถามว่าเห็นอะไรบ้าง Gemini ก็สามารถวิเคราะห์ได้เป็นอย่างดี

<sup>2</sup> [https://cities-today.com/industry/real-time-video-analytics-and-the-road-to-vision-zero/?utm\\_source=chatgpt.com](https://cities-today.com/industry/real-time-video-analytics-and-the-road-to-vision-zero/?utm_source=chatgpt.com)



ภาพที่ 1.4 Video Question Answering

### 1.3.2 การช่วยวิเคราะห์เอกสาร

Gemini สามารถอ่านข้อความในเอกสาร, ตาราง, กราฟ หรือไฟล์ PDF แล้วสรุป/วิเคราะห์ข้อมูลสำคัญออกมา เช่น วิเคราะห์รายงานทางการเงิน สรุปบทเรียนหรืองานวิจัย แปลงเอกสารกระดาษเป็นข้อมูลดิจิทัล ความสามารถด้านนี้อาจเป็นที่ต้องการของแวดวงธุรกิจ นักเรียน และครูอาจารย์ รวมไปถึงนักวิเคราะห์การเงิน

### 1.3.3 ผู้ช่วยเขียนโค้ด

ผู้ช่วยเขียนโค้ด Gemini ช่วยเขียน แก้ไข และอธิบายโค้ดในหลายภาษา เช่น Python, JavaScript, C++ สามารถเชื่อมกับ Google Colab หรือ VS Code เพื่อช่วย debug<sup>3</sup> หรืออธิบายโค้ดให้เข้าใจง่าย

<sup>3</sup> การ ค้นหา วิเคราะห์ และแก้ไขข้อผิดพลาด (bug) ที่เกิดขึ้นใน โปรแกรมหรือระบบ

### 1.3.4 ผู้ช่วยด้านการเรียนรู้

Gemini เข้าใจทั้งข้อความและภาพ ทำให้ช่วยสอนได้หลายแบบ เช่น อธิบาย โจทย์คณิตศาสตร์โดยดูจากรูปถ่ายสมการ แปลงโน้ตในสมุดเป็นสรูปพร้อมภาพประกอบ สร้างแบบฝึกหัดอัตโนมัติจากเนื้อหาที่ผู้เรียนอัปโหลด นักเรียนสามารถ สอบถามเนื้อหาต่าง ๆ ที่ตนอาจจะยังไม่เข้าใจ ครูอาจใช้ให้ช่วยออกข้อสอบหรือ เรียบเรียงเนื้อหาการสอน

### 1.3.5 ผู้ช่วยคอนเทนต์อัจฉริยะ

Gemini เข้าใจบริบทของข้อความและภาพ ทำให้สร้างสื่อได้หลากหลาย เช่น เขียนบทความ และออกแบบภาพประกอบ สร้าง Storyboard จากคำอธิบาย ออกแบบโปสเตอร์ แผ่นพับ หรือโฆษณา

ที่กล่าวมาก็เป็นเพียงตัวอย่างบางส่วนที่สามารถทดลองใช้ได้โดยไม่เสียเงิน ทั้ง Gemini และ ChatGPT มีความสามารถคล้ายคลึงกัน ตารางที่ 1.1 แสดงการ เปรียบเทียบคุณสมบัติบางอย่างระหว่างสองแพลตฟอร์ม ในการเปรียบเทียบแสดงให้เห็นความแตกต่างของทั้งสองแพลตฟอร์มซึ่งทั้งคู่มีจุดเด่นที่แตกต่างกันไป แต่ไม่ ได้พิจารณาว่าใครเหนือกว่าใคร

#### ตารางที่ 1.1 เปรียบเทียบความสามารถกับ ChatGPT

คุณสมบัติ	Gemini (Google)	ChatGPT (OpenAI)
Multimodal	ถูกออกแบบและฝึกให้เข้าใจข้อมูลหลายรูปแบบ	ถูกออกแบบพัฒนาจากโมเดลที่เก่งทางด้านข้อความ ต่อมาจึงเพิ่มความสามารถในด้านอื่น ๆ
Context window หรือ ความจำระยะสั้น	ขอบเขตข้อมูลที่ Gemini สามารถจำและใช้เหตุผลได้ในขณะสนทนาในโมเดลบางรุ่นสูงถึง 1 ล้านโทเค็น	โมเดลของ ChatGPT ในบางรุ่นรองรับประมาณ 128,000 โทเค็น

คุณสมบัติ	Gemini (Google)	ChatGPT (OpenAI)
การเชื่อมต่อบริการอื่น	Google Workspace, Search, YouTube	Plug-ins, OpenAI API, Web Search
ความสามารถในการใช้เหตุผล	ดีมีการเชื่อมโยงข้อมูลจากหลายแหล่ง เช่น ภาพ ข้อความ และบริบท เพื่อวิเคราะห์และสรุปผลอย่างมีเหตุผล ไม่ใช่เพียงการตอบสนองต่อข้อมูลรูปแบบเดียว	เด่นมากในด้านตรรกะและการเขียน
ความเข้าใจวิดีโอ	เด่นกว่า	ยังจำกัด GPT มีจุดแข็งด้านการใช้เหตุผลจากข้อความเป็นหลัก ทำให้การเข้าใจวิดีโอซึ่งมีมิติของเวลาและบริบทที่ซับซ้อนยังเป็นความท้าทาย
ความสามารถด้านภาษา	มีจุดเด่นในเรื่อง OCR หลายภาษา เพราะเป็น Multimodal AI ที่ไม่เพียงอ่านตัวอักษรจากภาพ แต่ยังเข้าใจบริบทและใช้เหตุผล ร่วมกับภาษาได้ดี โดยเฉพาะเมื่อมีหลายภาษาอยู่ในภาพเดียวกัน	เขียนสื่อสารได้เป็นธรรมชาติที่สุด
ความคิดสร้างสรรค์ในการเขียน	ปานกลาง	เด่นที่สุดในตลาดตอนนี้
การเขียนโค้ด	สามารถเขียนโค้ดและผสมผสานโค้ดกับบริบทจริง, เอกสาร, ภาพ, และระบบของ Google ได้	สามารถเขียนโค้ด โดยอธิบายแนวคิด การคิดเชิงอัลกอริทึม, สร้างโครงสร้างโค้ด, และการอธิบายที่ละเอียดที่อ่านเข้าใจง่าย

## 1.4 Roboflow

Roboflow คือแพลตฟอร์มสำหรับงานคอมพิวเตอร์วิทัศน์แบบครบวงจร ตั้งแต่การจัดการชุดข้อมูล (dataset) การติดป้ายกำกับ (annotation/labeling) การเพิ่มข้อมูล (augmentation) การฝึกโมเดล (train) ไปจนถึงการทดสอบ และใช้งานจริงทั้งบนคลาวด์ เว็บ และอุปกรณ์ขอบเครือข่าย (Edge เช่น Jetson, Raspberry Pi, มือถือ)

เวิร์กโฟลว์ในการทำงานกับ Roboflow เริ่มจากการเก็บข้อมูลซึ่งเป็นรูปภาพและอัปโหลดเข้าสู่ระบบ จากนั้นทำการติดป้ายกำกับข้อมูลโดยมีเครื่องมือช่วยกำหนดขอบเขตของวัตถุในภาพ เมื่อเตรียมข้อมูลเรียบร้อยแล้วจึงนำไปฝึกโมเดลและประเมินผลการทำงาน

Roboflow ยังมีเครื่องมือสำหรับจัดการและเตรียมภาพ เช่น การลดขนาด การครีโปก และการปรับค่าพิกเซลของภาพให้อยู่ในช่วงมาตรฐาน เพื่อช่วยให้โมเดลเรียนรู้ได้ง่ายและมีเสถียรภาพมากขึ้น นอกจากนี้ยังมีเทคนิค Data Augmentation ซึ่งเป็นการเพิ่มความหลากหลายของข้อมูลฝึกโดยไม่ต้องเก็บข้อมูลใหม่ ด้วยการแปลงข้อมูลเดิม เช่น การเปลี่ยนมุมมอง แสง หรือขนาดของภาพ เทคนิคนี้ช่วยลดปัญหา overfitting<sup>4</sup> และเพิ่มความทนทาน (robustness) ให้กับโมเดล ทำให้เมื่อใช้งานจริงโมเดลยังสามารถทำนายได้แม่นยำแม้ว่าภาพอินพุตจะมีการเปลี่ยนแปลงไปบ้าง

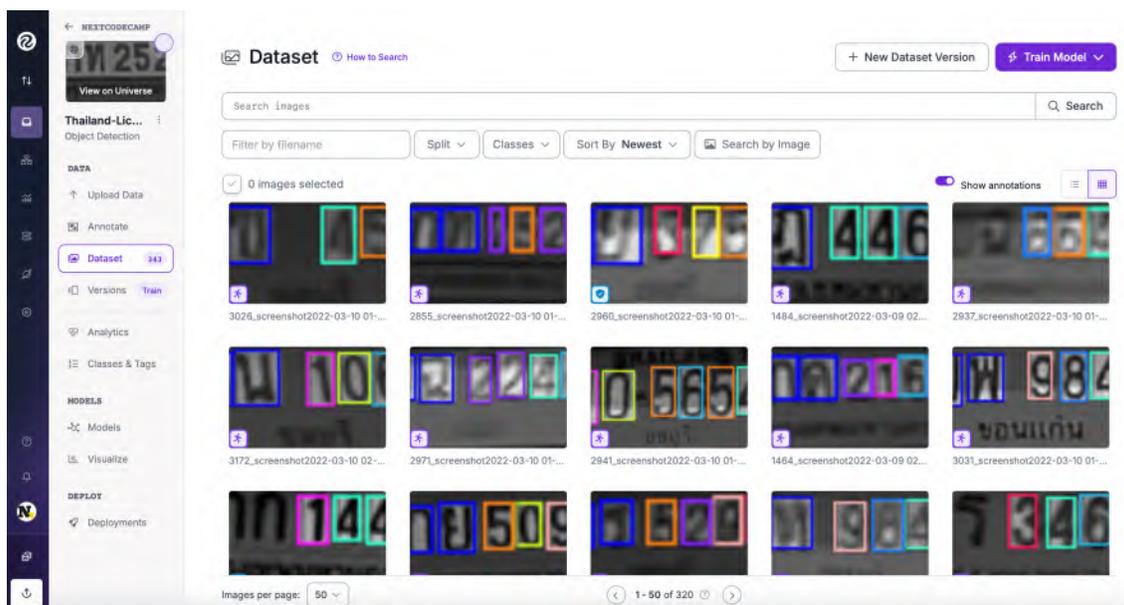
การอ่านป้ายทะเบียนรถยนต์ไทย (ภาพที่ 1.5) เป็นปัญหาด้านคอมพิวเตอร์วิทัศน์ ที่ประกอบด้วยสองขั้นตอนหลัก คือ Object Detection และ OCR (Optical Character Recognition) ซึ่งขั้นตอนแรกจะระบุ และครอบตำแหน่งของป้ายทะเบียนในภาพ ส่วนขั้นตอนหลังจะแยกแยะ และจำแนกตัวอักษร และตัวเลขบนป้ายทะเบียน

<sup>4</sup> ภาวะที่โมเดลเรียนรู้ข้อมูลฝึกจนจำรายละเอียดได้หมด ทำให้ทำนายข้อมูลใหม่ได้ไม่ดี



ภาพที่ 1.5 อ่านป้ายทะเบียน

ในภาพที่ 1.6 แสดงตัวอย่างชุดข้อมูล (Data set) ที่ใช้ในการฝึกฝนโมเดลซึ่งมีการติดป้ายกำกับแล้ว



ภาพที่ 1.6 ป้ายทะเบียนรถใน Roboflow

ผู้อ่านที่สนใจ สามารถเข้าใช้บริการ Roboflow ได้ฟรี แต่ต้องอนุญาตให้สามารถแชร์ชุดข้อมูลได้