



โครงการ การเรียนการสอนเพื่อเสริมประสบการณ์

ชื่อโครงการ การสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง
Article Summarization Using Machine Learning Techniques

ชื่อนิสิต นายธนเดช รุ่งศรีพัฒน์พร 603 36254 23
 นายชลิต ปฐมกลีวัฒนา 603 36145 23

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์
 สาขาวิชา วิทยาการคอมพิวเตอร์

ปีการศึกษา 2563

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง

นายธนเดช รุ่งศรีพัฒนพร
นายชลิต ปฐมกสิวัฒนา

โครงการนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2563
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Article Summarization Using Machine Learning Techniques

Tanadate Rungsripattanaporn

Chalit Pathomkasiwattana

A Project Submitted in Partial Fulfillment of the Requirements
for the Degree of Bachelor of Science Program in Computer Science

Department of Mathematics and Computer Science

Faculty of Science

Chulalongkorn University

Academic Year 2020

Copyright of Chulalongkorn University

นายธนเดช รุ่งศรีพัฒนพร, นายชลิต ปฐมกสิวัฒนา: การสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง. (Article Summarization Using Machine Learning Techniques)
 อ.ที่ปรึกษาโครงการหลัก : อาจารย์ ดร.นฤมล ประทานวณิช, 74 หน้า.

โครงการวิจัยเรื่อง “การสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง” มีวัตถุประสงค์เพื่อพัฒนาวิธีการสรุปบทความอัตโนมัติโดยใช้หลักการการเรียนรู้ของเครื่อง ที่ได้นำความรู้ และงานวิจัยที่เกี่ยวข้องมาประยุกต์ใช้ในการทำโครงการวิจัยฉบับนี้ การสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่องเป็นการสรุปบทความภาษาไทย โดยทำการเก็บรวบรวมข้อมูลมาจาก www.thepotential.org แล้วจึงนำบทความที่ได้มาเปลี่ยนแปลงเป็น Word Embedding ซึ่งเป็นการแปลงจากตัวอักษรไปเป็นเวกเตอร์จำนวนจริง หลังจากนั้นจึงนำเวกเตอร์ที่ได้มาใช้ในการจัดกลุ่มโดยใช้วิธีการจัดกลุ่ม แล้วจึงวัดผลด้วย BLEU Score, ROUGH Score และ F1 Score ผลการวิจัยแสดงให้เห็นว่า ผลสรุปที่ได้จากแบบจำลองนั้นดีกว่าผลสรุปที่ได้จากการสุ่ม

ภาควิชา...คณิตศาสตร์และวิทยาการคอมพิวเตอร์...ลายมือชื่อนิสิต..... ชนเดช รุ่งศรีพัฒนพร

ลายมือชื่อนิสิต..... ชลิต ปฐมกสิวัฒนา

สาขาวิชา...วิทยาการคอมพิวเตอร์.....ลายมือชื่อ อ.ที่ปรึกษาโครงการหลัก..... นฤมล ประทานวณิช

ปีการศึกษา.....2563.....

6033625423, 6033614523: MAJOR COMPUTER SCIENCE

KEYWORDS : MACHINE LEARNING / TEXT SUMMARIZATION / WORD EMBEDDING /
CLUSTERING

Tanadate Rungsripattanaporn, Chalit Pathomkasiwattana: Article Summarization
Using Machine Learning Techniques. ADVISOR : PROF. Naruemon Pratanwanich,
Ph.D., 74 pp.

The topic of this project is " Article Summarization Using Machine Learning
Techniques ". The objective of this research is to apply machine learning techniques to
summarize Thai articles. We started by gathering data of articles from
www.thepotential.com and then used word embedding techniques to transform words to
numerical vectors. After that, for each article, we clustered their word vectors to form a
summary. Finally, we evaluated our methods by using BLEU Score, ROUGH Score and F1
Score. Results that summarization from model is better than random.

Department : ..Mathematics and Computer Science.....Student's Signature.....

Student' Signature.....

Field of Study : ..Computer Science.....Advisor's Signature.....

Academic Year : 2020

กิตติกรรมประกาศ

โครงการวิจัยในชั้นเรียนในหัวข้อเรื่อง “การสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง” สามารถเสร็จสมบูรณ์ได้ เนื่องจากได้รับความอนุเคราะห์ และความช่วยเหลือจากอาจารย์ ดร.นฤมล ประทานวณิช อาจารย์ที่ปรึกษาโครงการ ที่ได้ให้ความรู้ คำแนะนำ ตลอดจนแก้ไขข้อบกพร่องต่าง ๆ มาโดยตลอด จนโครงการวิจัยฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี

ขอขอบพระคุณ รองศาสตราจารย์ ดร.ศรันญา มณีโรจน์ และ อาจารย์ โชติรส สุรพลชัย กรรมการคุมสอบทั้ง 2 ท่าน ผู้ที่สละเวลามาช่วยให้คำแนะนำแนวทางต่าง ๆ ที่เป็นประโยชน์ต่อโครงการวิจัยฉบับนี้

ขอขอบพระคุณครอบครัวที่คอยให้ความเชื่อมั่น และความช่วยเหลืออย่างสม่ำเสมอตลอดระยะเวลาในการพัฒนาโครงการวิจัย

สุดท้ายนี้ขอขอบคุณเพื่อน ๆ คณะวิทยาศาสตร์ ภาควิชาคณิตศาสตร์ สาขาวิชาวิทยาการคอมพิวเตอร์ทุกคนที่คอยให้คำปรึกษา และเป็นกำลังใจ แม้ในยามที่ประสบปัญหา หรืออุปสรรคต่าง ๆ

สารบัญ

หน้า

| | |
|---|----|
| บทคัดย่อภาษาไทย | จ |
| บทคัดย่อภาษาอังกฤษ | ฉ |
| กิตติกรรมประกาศ | ช |
| สารบัญ | ซ |
| สารบัญภาพ | ญ |
| | |
| บทที่ 1 บทนำ | 1 |
| 1.1 ความเป็นมาและเหตุผลการวิจัย | 1 |
| 1.2 วัตถุประสงค์ของการวิจัย | 1 |
| 1.3 ขอบเขตการวิจัย | 1 |
| 1.4 ขั้นตอนการวิจัย | 2 |
| 1.5 ประโยชน์ที่คาดว่าจะได้รับ | 3 |
| 1.6 โครงสร้างของรายงาน | 3 |
| | |
| บทที่ 2 ความรู้ และงานวิจัยที่เกี่ยวข้อง | 4 |
| 2.1 การตัดแบ่งข้อความ | 4 |
| 2.2 กระบวนการเปลี่ยนแปลงคำ หรือวลี ให้เป็นเวกเตอร์ตัวเลข | 6 |
| 2.3 กระบวนการเปลี่ยนแปลงคำ หรือวลี ให้เป็นเวกเตอร์ตัวเลขโดยใช้ BERT | 8 |
| 2.4 การจัดกลุ่มข้อมูล | 9 |
| 2.5 การสรุปข้อความ | 9 |
| 2.6 การประเมินผลการทำงานของแบบจำลอง | 10 |
| 2.7 งานวิจัยที่เกี่ยวข้อง : Text Summarization with Pretrained Encoders | 10 |
| 2.8 งานวิจัยที่เกี่ยวข้อง : Fine-tune BERT for extractive summarization | 11 |

| | |
|---|----|
| 2.9 งานวิจัยที่เกี่ยวข้อง : Leveraging BERT for Extractive Text Summarization on Lectures | 11 |
| บทที่ 3 วิธีการวิจัย..... | 13 |
| 3.1 การเก็บรวบรวมข้อมูลเพื่อนำมาใช้ในการพัฒนาแบบจำลอง..... | 13 |
| 3.2 พัฒนาแบบจำลองโดยใช้ความรู้ที่เกี่ยวข้อง | 13 |
| 3.3 การทดสอบประสิทธิภาพของแบบจำลอง | 15 |
| บทที่ 4 ผลการวิจัย | 17 |
| 4.1 ลักษณะข้อมูลที่ใช้..... | 16 |
| 4.2 การวัดประสิทธิภาพของแบบจำลอง | 18 |
| บทที่ 5 ข้อเสนอแนะ..... | 29 |
| 5.1 ข้อเสนอ..... | 30 |
| 5.2 ปัญหา และวิธีแก้ปัญหา..... | 30 |
| 5.3 ข้อเสนอแนะ | 28 |
| รายการอ้างอิง..... | 31 |
| ภาคผนวก ก แบบเสนอหัวข้อโครงการ รายวิชา 2301399 Project Proposal ปีการศึกษา 2563.. | 32 |
| ภาคผนวก ข ตัวอย่างบทความ และสรุปจากแบบจำลอง..... | 48 |
| ประวัติผู้เขียน | 64 |

สารบัญภาพ

| | หน้า |
|--|------|
| รูปภาพที่ 2.1 ตัวอย่างการตัดคำโดยใช้ newmm (บน) และ attacut (ล่าง)..... | 4 |
| รูปภาพที่ 2.2 ตัวอย่างการตัดคำโดยอาศัยการใช้ Longest matching กับ dictionary..... | 5 |
| รูปภาพที่ 2.3 ตัวอย่างการแปลงคำศัพท์ต่าง ๆ ภายในประโยคให้อยู่ในรูปของ one-hot vector.... | 6 |
| รูปภาพที่ 2.4 การทำงานของ CBOW เปรียบเทียบกับ Skip-gram | 7 |
| รูปภาพที่ 2.5 ขั้นตอนการทำงานของ ULMFiT | 8 |
| รูปภาพที่ 4.1 กราฟแสดงจำนวนคำในบทความ | 16 |
| รูปภาพที่ 4.2 กราฟแสดงจำนวนคำในบทความหลังจากตัด Stop Words ออกไป..... | 17 |
| รูปภาพที่ 4.3 กราฟแสดงจำนวนคำในสรุปจากเว็บไซต์ | 17 |
| รูปภาพที่ 4.4 Box Plot แสดงผล BLEU score ROUGE Score และ F1-Score แบบ 1-gram ของสรุปที่มาจากโมเดล (สีน้ำเงิน) สรุปที่มาจากการสุ่ม (สีแดง) และ ตัวอย่างบทความ (สีเขียว) เมื่อ เปรียบเทียบกับสรุปจากเว็บไซต์ (Gold Summary)..... | 19 |
| รูปภาพที่ 4.5 Box Plot แสดงผล BLEU score ROUGE Score และ F1-Score แบบ 1-gram ของสรุปที่มาจากโมเดล (สีน้ำเงิน) สรุปที่มาจากการสุ่ม (สีแดง) และ ตัวอย่างบทความ (สีเขียว) เมื่อ เปรียบเทียบกับสรุปจากเว็บไซต์ (Gold Summary)..... | 19 |

บทที่ 1

บทนำ

1.1 ความเป็นมาและเหตุผลการวิจัย

ในปัจจุบันมีบทความต่าง ๆ มากมายถูกเขียนขึ้นมาในแต่ละวันทั้งบทความข่าวสาร บทความที่ให้ความรู้ บทความที่เกี่ยวกับภาพยนตร์ เป็นต้น เนื่องด้วยจำนวนมากมายมหาศาลในปัจจุบัน การสรุปบทความจึงเป็นสิ่งที่จำเป็นอย่างยิ่ง เพื่อให้ได้บทความที่สั้น กระชับ ได้ใจความ เพื่อให้ผู้อ่านบทความใช้เวลาไม่มาก และเข้าใจเนื้อหาที่สำคัญได้อย่างรวดเร็ว

การสรุปบทความในปัจจุบัน ส่วนใหญ่ยังคงเป็นการสรุปด้วยฝีมือของมนุษย์ซึ่งจำเป็นต้องใช้เวลาเป็นอย่างมากในการทำสรุปแต่ละครั้ง เนื่องด้วยการสรุปโดยใช้คอมพิวเตอร์นั้นยังไม่เป็นที่แพร่หลายนัก เพราะไม่สามารถสรุปบทความได้ถูกต้องสมบูรณ์เท่าฝีมือของมนุษย์

คณะผู้จัดทำจึงต้องการที่จะพัฒนาการสรุปบทความแบบอัตโนมัติโดยใช้หลักการเรียนรู้ของเครื่อง (Machine Learning) เพื่อเป็นอีกก้าวหนึ่งที่น่าไปสู่การสรุปบทความที่สมบูรณ์ยิ่งขึ้นในอนาคตโดยไม่จำเป็นต้องพึ่งพาการสรุปของบุคคล สามารถใช้งานในการสรุปได้อย่างถูกต้อง สะดวก และรวดเร็ว

1.2 วัตถุประสงค์ของการวิจัย

เพื่อพัฒนาวิธีการสรุปบทความอัตโนมัติโดยใช้หลักการการเรียนรู้ของเครื่อง

1.3 ขอบเขตการวิจัย

1. ข้อความที่ใช้ในการทำสรุปเป็นข้อความประเภทบทความ (Article) จากเว็บไซต์ www.thepotential.org
2. การสรุปบทความจะเป็นการสรุปแบบ Extractive Summarization เท่านั้น
3. ข้อความที่ใช้ในการทำสรุปต้องเป็นภาษาไทยเท่านั้น
4. ข้อความที่ได้จากการสรุปบทความจะเป็นภาษาไทย

1.4 ขั้นตอนการวิจัย

1. ศึกษาเนื้อหาและบทความเกี่ยวกับการสรุปข้อความ
2. ศึกษาเนื้อหาและบทความเกี่ยวกับแบบจำลองต่าง ๆ ที่ใช้ในการทำสรุปข้อความ
3. เก็บข้อมูลจากหน้าเว็บไซต์ www.thepotential.org
4. ทำการพัฒนาแบบจำลอง
5. ทดสอบวัดประสิทธิภาพของแบบจำลองที่ได้
6. ปรับปรุงแบบจำลองเพื่อเพิ่มประสิทธิภาพ
7. สรุปผลการดำเนินงาน ข้อเสนอแนะ และการจัดทำเอกสาร

ตารางการดำเนินงาน

| ขั้นตอนการดำเนินงาน | ปีการศึกษา 2563 | | | | | | | | |
|---|-----------------|------|------|------|------|------|------|-------|-------|
| | ส.ค. | ก.ย. | ต.ค. | พ.ย. | ธ.ค. | ม.ค. | ก.พ. | มี.ค. | เม.ย. |
| 1.ศึกษาเนื้อหาและบทความเกี่ยวกับการสรุปข้อความ | | | | | | | | | |
| 2. ศึกษาเนื้อหาและบทความเกี่ยวกับแบบจำลองต่าง ๆ ที่ใช้ในการทำสรุปข้อความ | | | | | | | | | |
| 3. เก็บข้อมูลจากหน้าเว็บไซต์ www.thepotential.org | | | | | | | | | |
| 4. ทำการพัฒนาแบบจำลอง | | | | | | | | | |
| 5.ทดสอบวัดประสิทธิภาพของแบบจำลองที่ได้ | | | | | | | | | |
| 6.ปรับปรุงแบบจำลองเพื่อเพิ่มประสิทธิภาพ | | | | | | | | | |
| 7. สรุปผลการดำเนินงาน ข้อเสนอแนะ และการจัดทำเอกสาร | | | | | | | | | |

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ประโยชน์ในด้านความรู้ และประสบการณ์ต่อตัวนิสิตเอง
 - 1.1 ได้รับความรู้เกี่ยวกับหลักการเรียนรู้ของเครื่อง (Machine Learning)
 - 1.2 ได้รับความรู้เกี่ยวกับการประมวลผลภาษาธรรมชาติ (Natural Language Processing)
2. ประโยชน์ในด้านการแก้ไขปัญหาของสังคมหรือสภาพแวดล้อม
 - 2.1 สามารถทำการสรุปบทความได้โดยอัตโนมัติ
 - 2.2 เป็นตัวช่วยสำหรับผู้เขียนข่าว

1.6 โครงสร้างของรายงาน

- บทที่ 2 จะกล่าวถึงงานวิจัยที่เกี่ยวข้องกับการสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง
- บทที่ 3 จะกล่าวถึงวิธีการวิจัย ซึ่งจะประกอบไปด้วยการเก็บรวบรวมข้อมูล การพัฒนาแบบจำลอง และการทดสอบประสิทธิภาพของแบบจำลอง
- บทที่ 4 จะกล่าวถึงผลการวิจัย
- บทที่ 5 จะกล่าวถึงข้อสรุป และข้อเสนอแนะ

บทที่ 2

ความรู้ และงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงความรู้ และงานวิจัยที่เกี่ยวข้องกับการสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง

2.1 การตัดแบ่งข้อความ

2.1.1 Sentence Tokenization คือการแบ่งข้อความเป็นประโยค โดยมีวิธีในการแบ่งอยู่ 3 วิธีใหญ่ ๆ ได้แก่

1. การแบ่งข้อความเป็นประโยคโดยใช้ช่องว่าง (whitespace)
2. การแบ่งข้อความเป็นประโยคโดยใช้ช่องว่าง (whitespace) และการขึ้นบรรทัดใหม่ (new line)
3. การแบ่งข้อความเป็นประโยคโดยใช้ Conditional Random Field (CRF) หรือการใช้ sequence model ประเภทหนึ่ง ซึ่งใช้หลักการของ condition probability ในการทำงาน หลักการทำงานคือจะทำการแบ่งคำแล้วทำการทำนายความน่าจะเป็นในการกำกับคำ (label) สำหรับแต่ละคำ โดยอาศัยการกำกับคำของคำต่าง ๆ ก่อนหน้าทั้งหมด

2.1.2 Word Tokenization คือการแบ่งประโยคหรือข้อความไปเป็นคำ ๆ ซึ่งจะถูกรเรียกว่า token เพื่อที่จะสามารถนำไปประมวลผลต่อ ตัวอย่างของการแบ่งคำ จะเป็นไปตามภาพที่ 1

อัลกอริทึมแต่ละแบบนั้นใช้วิธีคิดในการแบ่งคำแตกต่างกัน ซึ่งจะทำให้ผลลัพธ์ของการแบ่งคำและเวลาในการประมวลผลแตกต่างกันออกไปอีกด้วย [5] [7]

```
[12] from pythainlp.tokenize import word_tokenize
      word_tokenize(text, engine="newmm")
↳ ['ช', 'ลิต', 'อยาก', 'จะ', 'เล่น', 'เกม', 'ตลอด', 'ทั้งวัน']

      word_tokenize(text, engine='attacut')
↳ ['ชลิต', 'อยาก', 'จะ', 'เล่น', 'เกม', 'ตลอด', 'ทั้ง', 'วัน']
```

รูปภาพที่ 2.1 ตัวอย่างการตัดคำโดยใช้ newmm (บน) และ attacut (ล่าง)

โดยมีอัลกอริทึมที่ใช้แบ่งคำนั้น แบ่งออกเป็น 2 กลุ่มใหญ่ ๆ คือ

1. Dictionary-based คือการแบ่งคำโดยอาศัยการใช้ดิกชันนารี (dictionary) [8] และ กริดีอัลกอริทึม (Greedy algorithm) ตัวอย่างของอัลกอริทึมเช่น Longest matching, Maximal matching โดยมีตัวอย่างการใช้งาน ดังนี้

Longest matching จะทำการนำตัวอักษรไปตรวจกับดิกชันนารี โดยทำจากซ้ายไปขวา ถ้าไม่พบคำ ดังกล่าวปรากฏใน dictionary ก็จะทำให้การตัดข้อความออก กล่าวคือ จะได้คำที่มีความหมายที่ยาวที่สุดใน dictionary (ดังรูปที่ 2) แต่ก็ยังมีปัญหาในเรื่องของคำที่ไม่อยู่ในดิกชันนารี (unknown words) เนื่องจากว่าคำสามารถเกิดขึ้นใหม่ได้ตลอดเวลา แต่สามารถแก้ไขได้โดยใช้ Customized dictionary นั่นคือการกำหนดดิกชันนารี เพิ่มเติมในส่วนของคำที่ต้องการเพิ่มขึ้นมา หรือ Thai Character Clustering Rule(TCC) คือ การนำตัวอักษรในประโยคมารวมกันให้เป็นกลุ่มของตัวอักษร โดยอาศัยกฎต่าง ๆ เพื่อใช้ในการรวมกลุ่มตัวอักษร [4]

| ลำดับการทำงาน | ส่วนของคำที่ยาวที่สุด | ส่วนที่เหลือ |
|---------------|-----------------------|---------------------------|
| 1 | แนวทาง | การแก้ปัญหาหนังสือเกษตรกร |
| 2 | การ | แก้ปัญหาหนังสือเกษตรกร |
| 3 | แก้ | ปัญหาหนังสือเกษตรกร |
| 4 | ปัญหา | หนังสือเกษตรกร |
| 5 | หนังสือ | เกษตรกร |
| 6 | เกษตรกร | |

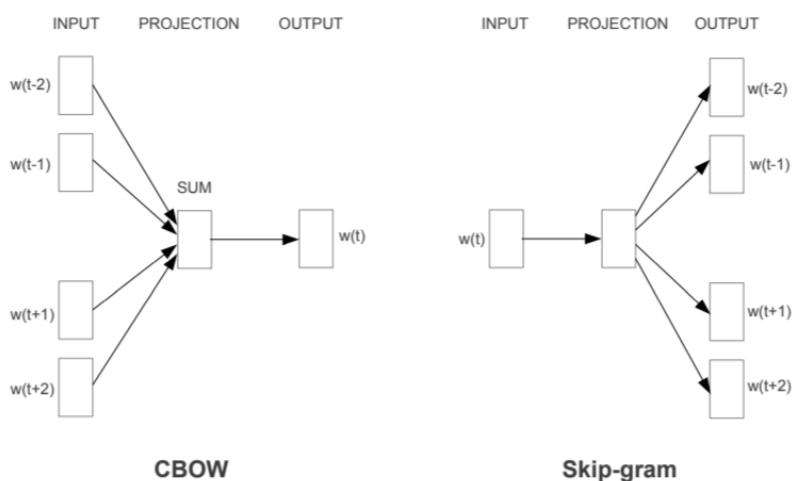
รูปภาพที่ 2.2 ตัวอย่างการตัดคำโดยอาศัยการใช้ Longest matching กับ dictionary

Source:http://thailang.nectec.or.th/halloffame/images/stories/best/download/best2010_12p34c002.pdf

รูปภาพที่ 2.3 ตัวอย่างการแปลงคำศัพท์ต่าง ๆ ภายในประโยคให้อยู่ในรูปของ one-hot vector

2.2.2 GloVe (Global Vectors) คือ Word Embedding โดยใช้วิธี unsupervised ในการได้เวกเตอร์ที่เป็นตัวแทนของคำมา ซึ่งทำให้ไม่ต้องอาศัยข้อมูลในการฝึกสอนมากนักแต่ก็ยังให้ผลที่ดี เป็นผลให้เวลาในการฝึกสอนนั้นรวดเร็วมากขึ้นอีกด้วย และยังอาศัยการใช้วิธีการทางสถิติ นั่นคือ การใช้อัตราส่วนความน่าจะเป็น (word co-occurrence) จากคลังข้อมูล (corpus)

2.2.3 Word2vec คือ word embedding ชนิดหนึ่ง โดยเป็นการเปลี่ยนคำให้อยู่ในรูปของตัวเลข แต่ตัวเลขเหล่านั้นจะเก็บบริบทของคำไว้ด้วย (เป็น feature สำหรับนำไปฝึกสอนโมเดล เพื่อให้โมเดลเกิดการเรียนรู้ต่อไป) โดยแรกเริ่มนั้นใช้วิธี one hot encoding แล้วจึงนำข้อมูลไปใส่ลงใน Neural Network โดยมีวิธีการฝึกสอน 2 รูปแบบ ได้แก่



รูปภาพที่ 2.4 การทำงานของ CBOW เปรียบเทียบกับ Skip-gram

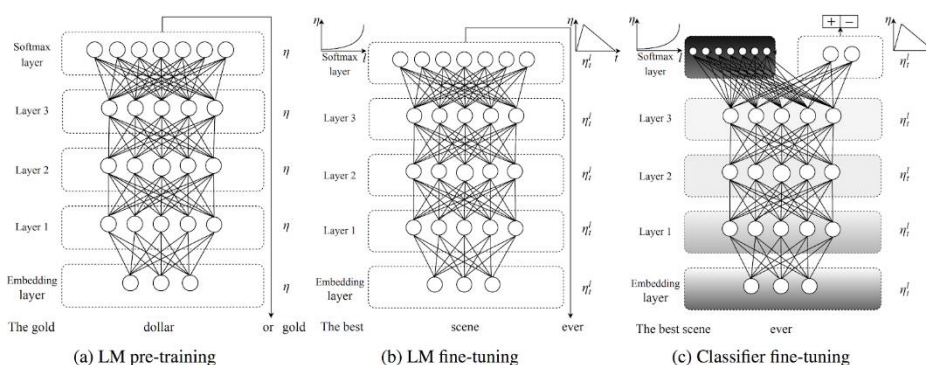
Source: *Exploiting Similarities among Languages for Machine Translation*

1. **CBOW (Continuous bag-of-words)** คือการทำนายคำโดยอาศัยบริบทต่าง ๆ กล่าวคือผลลัพธ์ที่ได้ออกมา นั้นอาศัยบริบทจากคำที่กระจายตัวอยู่รอบ ๆ หรือคำข้าง ๆ (ใช้หลาย

input ได้หนึ่ง output) โดยส่วนมากจะใช้เวลาเทรนน้อยและให้ผลที่ดีกว่า skip-gram เนื่องจากการที่ การที่อ้างอิงจากบริบทของหลาย ๆ คำนั้น จะสามารถทำนายได้ผลลัพธ์แม่นยำมากยิ่งขึ้น

2. Skip-gram คือการทำนายคำโดยอาศัยบริบทจากคำคำเดียว กล่าวคือผลลัพธ์ที่ได้ ออกมานั้นอาศัยบริบทจากคำคำเดียว (ใช้หนึ่ง input ได้หลาย output) ซึ่งจะทำงานได้ดีกับชุดข้อมูล ขนาดเล็ก และยังแสดงคำหรือวลีที่หายากได้เป็นอย่างดี

2.2.4 ULMFiT (Universal Language Model Fine-tuning)



รูปภาพที่ 2.5 ขั้นตอนการทำงานของ ULMFiT

Source: <https://thaikeras.com/2018/transfer-learning-nlp/>

Word vector ภายใน pythainlp นั้นมาจาก ULMFiT ซึ่งเป็น transfer learning ชนิดหนึ่ง มี ส่วนประกอบหลักคือ AWD-LSTM เมื่อทำการเรียนรู้ Language Model แล้ว จึงนำไปใช้งานต่าง ๆ ได้ โดยต้องทำการ fine-tuning ก่อน คือการเทรน Language Model สำหรับงานที่จะนำไปใช้ แล้ว จึงค่อยเทรนงานที่จะทำจริง ๆ

2.3 กระบวนการเปลี่ยนแปลงคำ หรือวลี ให้กลายเป็นเวกเตอร์ตัวเลข โดยใช้ BERT (BERT Sentence Embedding)

BERT เป็น Transfer Learning ประเภทหนึ่ง เป็นการเรียนรู้ Language Model ของภาษา นั้น ๆ เพื่อนำมาสร้างในส่วนของ Word Embedding, Sentence Embedding เพื่อที่จะนำไป ประมวลผลต่อไป โดย

2.3.1 Average Word Embedding within a Sentence คือการหา Embedding ของประโยคโดยการหาค่าเฉลี่ยของ Embedding ของคำที่อยู่ภายในประโยค

2.3.2 BERT for Sentence Embedding คือการใช้ Embedding ของประโยคที่นำมาจาก BERT

2.4 การจัดกลุ่มข้อมูล (Clustering)

2.4.1 K-mean Clustering เป็นวิธีการ Unsupervised-learning ซึ่งไม่ต้องอาศัยข้อมูลที่มี Label โดยทำการจัดกลุ่มของข้อมูลเป็นจำนวน k กลุ่ม มีขั้นตอนการทำงาน ดังนี้

1. กำหนดข้อมูลขึ้นมาเป็น 2 กลุ่ม
2. สุ่มเลือกจุด Centroid เพื่อใช้เป็นตัวแทนของแต่ละกลุ่ม
3. ทหาระยะห่างระหว่างจุดข้อมูล และจุด Centroid หากจุดข้อมูลใดอยู่ใกล้จุด Centroid ของกลุ่มนั้นๆ ถ้ามากกว่าให้ถือว่าเป็นสมาชิกของกลุ่มนั้น
4. ได้กลุ่มข้อมูลใหม่ที่ต่างจากเดิมของทั้ง 2 กลุ่ม
5. กำหนดจุด Centroid ของแต่ละกลุ่มใหม่
6. ทำวนไปเรื่อย ๆ จนกว่าจะไม่มีเปลี่ยนแปลงหรือเปลี่ยนแปลงน้อยที่สุด

2.5 การสรุปข้อความ (Text Summarization)

คือการสรุปข้อความจากข้อความที่ได้รับมาทั้งหมด โดยที่ยังสามารถแสดงใจความสำคัญของข้อความทั้งหมดได้ ข้อความที่นำมาใช้อาจจะมีมาจากหลายแหล่ง เช่น ข่าว บทความ เป็นต้น ดังนั้นข้อมูลที่ใช้นำเข้าคือข้อความที่ต้องการทำการสรุป และ ข้อมูลผลลัพธ์ที่ได้ก็คือข้อความที่ได้ทำการสรุปแล้ว ซึ่งสามารถนำมาใช้เพื่อให้การอ่านข่าวหรือบทความที่มีความยาวมีความสะดวก รวดเร็วที่มากยิ่งขึ้น โดยมีตัวอย่างเป็นการวิจัยในการทำสรุปข้อความในหนังสือเรียน แบ่งเป็น 2 ประเภท คือ

2.5.1 Extractive Summarization คือการสรุปข้อความโดยการใช้ข้อความเดิมที่มีอยู่แล้วนำมาสร้างสรุป โดยยังคงใจความสำคัญเอาไว้

2.5.2 Abstractive Summarization คือการสรุปข้อความโดยการสร้างสรุปขึ้นมาโดยใช้ Deep Learning ในการทำ ซึ่งสรุปที่ได้มานั้นอาจจะมีความเหมือนหรือต่างกับข้อความต้นฉบับก็ได้ แต่ยังคงใจความสำคัญเอาไว้

2.6 การประเมินผลการทำงานของแบบจำลอง

การประเมินผลของการทำสรุปนั้นใช้การวัดอยู่ 3 ประเภท ซึ่งมีความแตกต่างกัน คือ

2.6.1 ROUGE score โดย ROUGE score จะทำการวัดในส่วนของ Recall เป็นการบ่งบอกความแม่นยำของตัวระบบโดยสนใจว่ามีข้อความสรุปต้นฉบับปรากฏอยู่ในข้อความสรุปที่สร้างขึ้นมาเท่าไร

2.6.2 BLEU score โดย BLEU score จะทำการวัดในส่วนของ Precision เป็นการบ่งบอกความแม่นยำของตัวระบบโดยสนใจในส่วนของสรุปที่สร้างขึ้นมานั้นปรากฏอยู่ในข้อความสรุปต้นฉบับเท่าไร

2.6.3 F-1 Score คือการหาค่าเฉลี่ยระหว่าง Precision และ Recall เพื่อใช้วัดความสามารถของแบบจำลอง

2.7 งานวิจัยที่เกี่ยวข้อง : Text Summarization with Pretrained Encoders

จากงานวิจัยที่ศึกษาโดย Yang Liu และ Mirella Lapata ในปี ค.ศ. 2019 [1] ได้ทำการทดลองสรุปข้อความ แบ่งเป็น 2 กลุ่ม คือ Extractive และ Abstractive โดยใช้ชุดข้อมูลที่มีเอกลักษณ์เฉพาะของแต่ละชุดข้อมูลที่แตกต่างกัน รวมทั้งหมด 3 ชุดข้อมูล ได้แก่ 1.CNN/Daily mail ซึ่งเป็นข้อความประเภทข่าว ทำการไฮไลท์สรุปให้แล้ว จึงเหมาะแก่การทำสรุปข้อความแบบ Extractive มาก 2.New York Times (NYT) ซึ่งเป็นข้อความประเภทบทความ และมีการสรุปข้อความที่เป็นไปในแบบ abstractive และ 3.XSum เป็นข้อความประเภทข่าว ทำการสรุปโดยใช้ประโยคเดียวในการสรุปข้อความทั้งหมด โดยทั้งหมดนี้นำมาใช้กับ pretrained encoders ที่ชื่อว่า

BERT (Bidirectional Encoder Representations from Transformers) จากนั้นจึงนำผลมาเปรียบเทียบกับแบบจำลองก่อน ๆ นั่นคือจากที่เคยใช้ในการทำสรุปข้อความผ่านการวัดประสิทธิภาพด้วยคะแนน ROUGE Score ผลปรากฏว่า BERT นั้นให้ผลที่ดีเมื่อเทียบกับแบบจำลองอื่น อีกทั้งชุดข้อมูลในแต่ละชุดนั้นก็มิผลกับการทำสรุปด้วยเช่นกัน กล่าวคือในส่วนของชุดข้อมูล CNN/Dailymail นั้น BERT สามารถทำคะแนนได้ดีกว่าแบบจำลองแบบอื่น ๆ และได้ผลของคะแนนที่ใกล้เคียงกันทั้งในส่วนของแบบจำลอง Abstractive และ Extractive, ชุดข้อมูล NYT นั้น BERT ก็ยังสามารถทำคะแนนได้ดีกว่าแบบจำลองแบบอื่น ๆ และแบบ Abstractive ได้คะแนนที่ดีกว่าแบบ Extractive, XSum ในส่วนของ BERT แบบ Extractive ทำคะแนนได้ค่อนข้างแย่มากจึงไม่ได้นำมาผลมาแสดง ต่างจากแบบ Abstractive ที่ยังให้ผลที่ดีกว่าแบบจำลองอื่น ๆ และดีกว่าแบบจำลอง Oracle ซึ่งเป็นขอบเขตบนของงานนี้อีกด้วย

2.8 งานวิจัยที่เกี่ยวข้อง : Fine-tune BERT for extractive summarization

จากงานวิจัยของ Yang Liu ในปีค.ศ. 2019 [2] เน้นไปที่การทำสรุปข้อความแบบ Extractive โดยอาศัย BERT กับชุดข้อมูล CNN/Dailymail และ New York Times (NYT) โดยในการทดลองนั้นจะเป็นการเพิ่มชั้น (layer) ให้กับ BERT มี 3 แบบ หลัก ๆ 1.Classifier โดยการใช้ Sigmoid Function ในชั้น BERT Output 2.Inter-sentence Transformer เป็นการเพิ่มชั้นของ Transformer ขึ้นมาอีกชั้นในระดับประโยค และยังใช้ Sigmoid Function ในชั้น BERT Output เหมือนเดิม และ 3.Recurrent Neural Network (RNN) โดยทำการใช้ RNN แบบ LSTM(Long Short Term Memory) โดยจะใช้ input มาจาก BERT Output แล้วจึงให้ผลลัพธ์สุดท้ายด้วย Sigmoid Function เหมือนเดิม

2.9 งานวิจัยที่เกี่ยวข้อง : Leveraging BERT for Extractive Text Summarization on Lectures

จากงานวิจัยของ Derek Miller ในปีค.ศ. 2019 [3] นี้จะเน้นการทำสรุปข้อความแบบ Extractive และใช้ข้อมูลที่เป็นข้อความประเภทหนังสือเรียน ขั้นตอนวิธีที่ใช้ในการสรุปข้อความคือการใช้ Clustering โดยใช้กับ Word Embedding ของ BERT แล้วจึงนำไปเทียบกับการใช้ TextRank ซึ่งเป็นวิธีการสรุปข้อความแบบ Extractive อีกแบบ โดยทำการประเมินผลด้วยการเปรียบเทียบผลลัพธ์ของการสรุปทั้งสองแบบโดยมนุษย์ ผลพบว่าการทำ Clustering จาก Word Embedding ของ BERT นั้นให้ผลที่ดีกว่าแต่ก็ยังประสบปัญหาที่เหมือนกับตัวแบบจำลองอื่น ๆ ใน

กรณีที่ข้อความยาวเกินไป โดยในโครงการนี้ คณะผู้วิจัยจะยึดแบบแผนจากงานวิจัยของ Derek Miller ในการทำสรุปบทความ

บทที่ 3

วิธีการวิจัย

ในบทนี้จะกล่าวถึงวิธีการวิจัยการสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง เพื่อใช้ในการพัฒนาแบบจำลอง โดยมีทั้งหมด 3 ขั้นตอนหลักๆ ดังนี้

1. การเก็บรวบรวมข้อมูลเพื่อนำมาใช้ในการพัฒนาแบบจำลอง
2. พัฒนาแบบจำลองโดยใช้ความรู้ที่เกี่ยวข้อง
3. ทดสอบประสิทธิภาพการทำงานของแบบจำลอง

3.1 การเก็บรวบรวมข้อมูลเพื่อนำมาใช้ในการพัฒนาแบบจำลอง

ในการวิจัยนี้ คณะผู้วิจัยได้ใช้ข้อมูลจากเว็บไซต์ www.thepotential.org โดยการใช้ Python Library ที่มีชื่อว่า BeautifulSoup ร่วมกับกับ Google Colaboratory จากนั้นจึงบันทึกข้อมูลที่ได้รับมาในรูปแบบไฟล์เอกสาร และจัดเก็บลงใน Google Drive เพื่อความสะดวกในการจัดเก็บข้อมูล

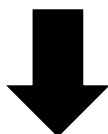
3.2 พัฒนาแบบจำลองโดยใช้ความรู้ที่เกี่ยวข้อง

3.2.1 การจัดเตรียมข้อมูล

หลังจากที่ทำการเก็บรวบรวมข้อมูลเสร็จสิ้น ทางคณะผู้วิจัยจึงได้เริ่มทำการพัฒนาแบบจำลอง โดยเริ่มจากการนำข้อมูลที่ได้มาทำการตัดข้อความให้อยู่ในระดับประโยค และคำ โดยใช้การทำ Sentence Tokenization และ Word Tokenization หลังจากนั้นจึงทำการแบ่งข้อมูลออกเป็น 2 ชุด ได้แก่ ชุดข้อมูลเดิม และชุดข้อมูลที่ถูกตัดคำต่างๆไปที่เราพบบ่อย (Stop Words) ออกไป

ตัวอย่างการตัดบทความให้อยู่ในระดับประโยค

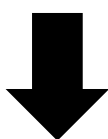
ชีวิตที่ผ่านมาของคุณเป็นอย่างไรบ้าง? ตั้งแต่ปฐมวัย-ประถมศึกษา-มัธยมศึกษา-มหาวิทยาลัย-ทำงาน-เปลี่ยนงาน-มีลูก ฯลฯ ลองนำความทรงจำช่วงนั้นขึ้นมาทบทวนตัวเองกันอีกครั้ง เพราะตัวตนที่ถูกหล่อหลอมขึ้นมาจากเหตุการณ์สำคัญในทุก 7 ปี อาจสะท้อนให้เห็นอีกครั้งโดยที่คุณก็ไม่รู้ตัวในอีกหลายปีให้หลัง



[ชีวิตที่ผ่านมาของคุณเป็นอย่างไรบ้าง? , 'ตั้งแต่ปฐมวัย-ประถมศึกษา-มัธยมศึกษา-มหาวิทยาลัย-ทำงาน-เปลี่ยนงาน-มีลูก ฯลฯ ', 'ลองนำความทรงจำช่วงนั้นขึ้นมาทบทวนตัวเองกันอีกครั้ง เพราะตัวตนที่ถูกหล่อหลอมขึ้นมาจากเหตุการณ์สำคัญในทุก 7 ปี ', 'อาจสะท้อนให้เห็นอีกครั้งโดยที่คุณก็ไม่รู้ตัวในอีกหลายปีให้หลัง']

ตัวอย่างการตัดบทความให้อยู่ในระดับคำ

[ชีวิตที่ผ่านมาของคุณเป็นอย่างไรบ้าง? , 'ตั้งแต่ปฐมวัย-ประถมศึกษา-มัธยมศึกษา-มหาวิทยาลัย-ทำงาน-เปลี่ยนงาน-มีลูก ฯลฯ ', 'ลองนำความทรงจำช่วงนั้นขึ้นมาทบทวนตัวเองกันอีกครั้ง เพราะตัวตนที่ถูกหล่อหลอมขึ้นมาจากเหตุการณ์สำคัญในทุก 7 ปี ', 'อาจสะท้อนให้เห็นอีกครั้งโดยที่คุณก็ไม่รู้ตัวในอีกหลายปีให้หลัง']



[[ชีวิต', 'ที่ผ่านมา', 'ของ', 'คุณ', 'เป็น', 'อย่างไรบ้าง', '?'], [ตั้งแต่', 'ปฐมวัย', '-', 'ประถมศึกษา', '-', 'มัธยมศึกษา', '-', 'มหาวิทยาลัย', '-', 'ทำงาน', '-', 'เปลี่ยน', 'งาน', '-', 'มีลูก', ' ฯลฯ'], [ลอง', 'นำ', 'ความทรงจำ', 'ช่วง', 'นั้น', 'ขึ้น', 'มา', 'ทบทวน', 'ตัวเอง', 'กัน', 'อีกครั้ง', 'เพราะ', 'ตัวตน', 'ที่', 'ถูก', 'หล่อหลอม', 'ขึ้น', 'มาจาก', 'เหตุการณ์', 'สำคัญ', 'ใน', 'ทุก', '7', 'ปี'], [อาจ', 'สะท้อน', 'ให้', 'เห็น', 'อีกครั้ง', 'โดยที่', 'คุณ', 'ก็', 'ไม่', 'รู้ตัว', 'ใน', 'อีก', 'หลาย', 'ปี', 'ให้หลัง']]

3.2.2 การแปลงประโยคให้กลายเป็นเวกเตอร์

หลังจากนั้นจึงนำชุดข้อมูลที่เตรียมมานั้นแปลงเป็น Word Embedding โดยใช้ Embedding 2 ประเภท โดยแบ่งเป็นแบบ Average Pooling เป็นการหาค่าเฉลี่ย และนำค่าเฉลี่ยที่ได้มาเป็นตัวแทนของประโยค โดยจะเรียก Embedding ว่า Embedding1 และในส่วนของ Embedding ที่ได้จาก BERT จะเรียกว่า Embedding2 ที่ได้มาเป็นตัวแทนของประโยค

3.2.3 การทำ Clustering และเลือกตัวแทนประโยคมาใช้ทำสรุปบทความ

นำ Embedding1 และ Embedding2 ที่ได้มาใช้งานร่วมกับอัลกอริทึม K-mean Clustering เพื่อการจัดกลุ่มข้อความในระดับประโยค โดยเหตุผลที่ใช้การทำ Clustering เนื่องจากว่าการจัดกลุ่มของ แต่ละ Embedding ภายในบทความนั้น ทำให้ได้กลุ่มของ Embedding ต่าง ๆ ที่จะสามารถสื่อได้ถึงความหมายใจความหลักของประโยคในแต่ละกลุ่มได้ โดยใช้ประโยคแรกของแต่ละย่อหน้าเป็น seeds เริ่มต้นในการทำ K-mean clustering ในการจับกันเป็นกลุ่มก่อนนั้นได้มีการจัดกลุ่มข้อความออกเป็นกลุ่ม ๆ ให้จำนวนกลุ่มเท่ากับจำนวนย่อหน้าของแต่ละบทความ โดยเหตุผลที่ใช้การกำหนดให้จำนวนกลุ่มเท่ากับจำนวนย่อหน้าและประโยคแรกในแต่ละย่อหน้าเป็นจุดเริ่มต้นของแต่ละกลุ่มนั้น เพื่อกำหนดให้ผลสรุปที่ได้จากบทความเดิมนั้นมีจำนวนกลุ่มที่เท่าเดิมและให้ผลที่คงเดิมมากที่สุดไม่ว่าจะทำการจัดกลุ่มกี่ครั้งก็ตาม หลังจากนั้นจึงเลือกประโยคที่อยู่ใกล้จุด Centroid มากที่สุดของแต่ละกลุ่มออกมา ซึ่งประโยคที่เลือกมานั้นจะสามารถแสดงความหมายของแต่ละกลุ่มได้ และนำประโยคนั้นมาเป็นตัวแทนประโยคของแต่ละกลุ่มจากนั้นจึงนำมารวมกันเพื่อให้ได้สรุปบทความออกมา

3.3 การทดสอบประสิทธิภาพของแบบจำลอง

ทางคณะผู้วิจัยวัดผลประสิทธิภาพการทำงานของแบบจำลองโดยนำการสรุปที่ได้จากเว็บไซต์ (Gold Summary) มาเปรียบเทียบกับสรุปที่ได้จากแบบจำลองเพื่อคำนวณหาค่า BLEU Score, ROUGE Score และ F-1 Score และนำมาเปรียบเทียบกับสรุปที่ได้มาจากการสุ่ม โดยการคิดคะแนนทั้งสามส่วนนี้ใช้ n-gram ในการคิดคะแนน n-gram คือส่วนประกอบของคำที่ติดกัน n คำ ยกตัวอย่างประโยคคือ ‘ฉันทินข้าว’ เพราะฉะนั้นแล้ว 1 gram คือ [ฉันทิน, ข้าว] 2 gram คือ [ฉันทิน, ข้าว] และ 3 gram คือ [ฉันทินข้าว] เป็นต้น นอกจากนี้คณะผู้วิจัยยังเปรียบเทียบความที่ยังไม่สรุปกับสรุปที่ได้จากเว็บไซต์และวัดประสิทธิภาพด้วยค่าทั้ง 3 แบบข้างต้น เพื่อใช้เป็นค่าขอบเขตบน (Upper bound) ของประสิทธิภาพที่เป็นไปได้ในการทำสรุปแบบ Extractive

บทที่ 4

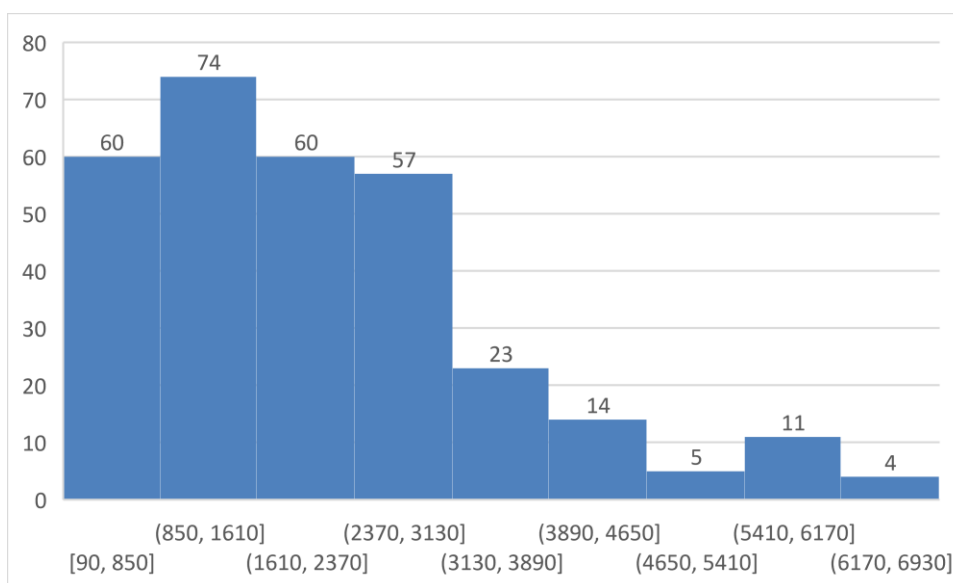
ผลการวิจัย

ในบทนี้จะกล่าวถึง ผลของการดำเนินการวิจัยการสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง

4.1 ลักษณะข้อมูลที่ใช้

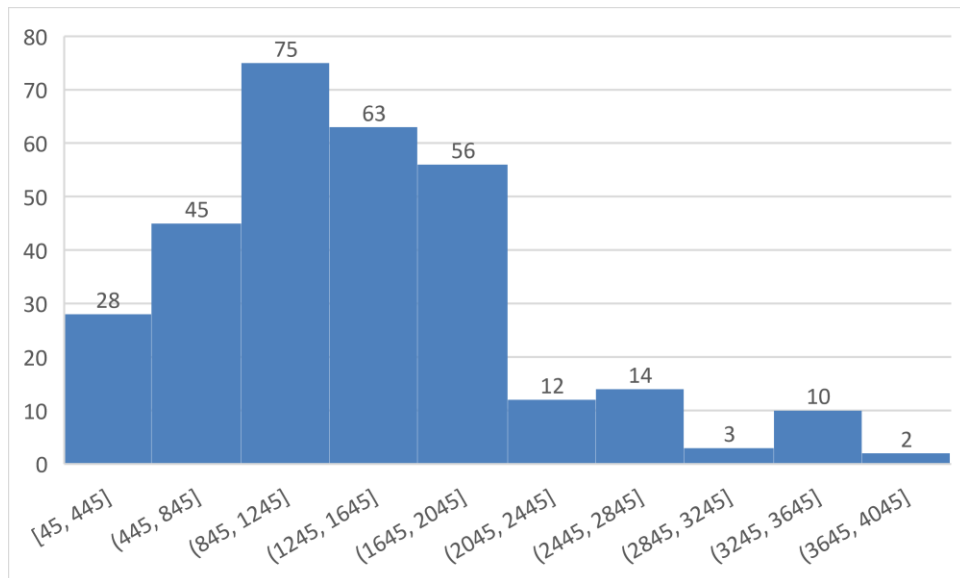
ข้อมูลบทความที่เก็บรวบรวมมาจากเว็บไซต์ www.thepotential.org จากหัวข้อ 3.1 ก่อนนำมาผ่านการจัดเตรียมข้อมูลด้วยการนำบทความที่ได้มาทั้งหมดจำนวน 308 บทความ (โดยมีจำนวนคำในบทความเฉลี่ยอยู่ที่ประมาณ 850 – 1600 คำ) มาตัดบทความให้อยู่ในระดับประโยค และคำ หลังจากนั้นจึงนำข้อมูลไปตัด Stop Words (ตัวอย่าง Stop Words เช่น 'กล่าว', 'จาก', 'ด้วย', 'ตาม', 'ต่าง', 'ถ้า', 'ทั้ง' เป็นต้น) ออกไป เพื่อให้แต่ละประโยคมีแต่คำที่จะสื่อถึงประโยคนั้น ๆ ไม่มีคำฟุ่มเฟือย และยังเป็นการลดจำนวนคำในบทความ ไม่ให้มีจำนวนคำที่มากเกินไปจนความจำเป็น

จำนวนบทความ



จำนวนบทความ

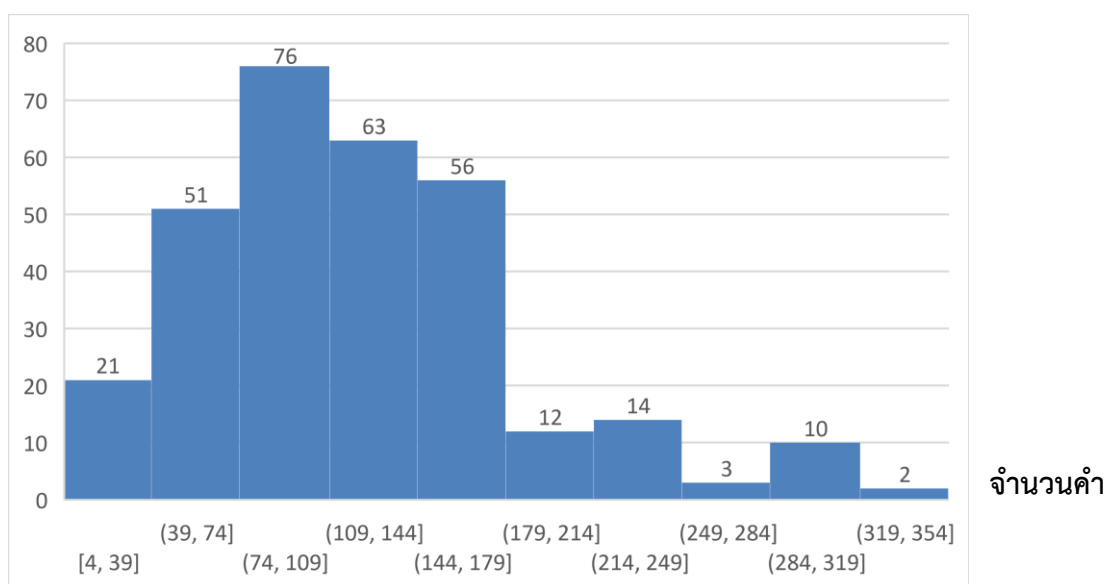
รูปภาพที่ 4.1 กราฟแสดงจำนวนคำในบทความ



จำนวนคำ

รูปภาพที่ 4.2 กราฟแสดงจำนวนคำในบทความหลังจากตัด Stop Words ออกไป

จำนวนบทความ



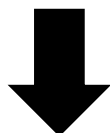
รูปภาพที่ 4.3 กราฟแสดงจำนวนคำในสรุปลจากเว็บไซต์

ส่วนหนึ่งของบทความที่ถูกเก็บรวบรวมข้อมูลมาก่อนการตัด Stop Words ออกไป

ชีวิตที่ผ่านมาของคุณเป็นอย่างไรบ้าง ? ตั้งแต่ปฐมวัย-ประถมศึกษา-มัธยมศึกษา-มหาวิทยาลัย-ทำงาน-เปลี่ยนงาน-มีลูก ฯลฯ ลองนำความทรงจำช่วงนั้นขึ้นมาทบทวนตัวเองกันอีกครั้ง เพราะตัวตนที่ถูกหล่อหลอมขึ้นมาจากเหตุการณ์สำคัญในทุก 7 ปี อาจสะท้อนให้เห็นอีกครั้งโดยที่คุณก็ไม่วู้ตัวในอีกหลายปีให้หลัง

ส่วนหนึ่งของบทความเดียวกันที่ผ่านการตัด Stop Words ออกไป

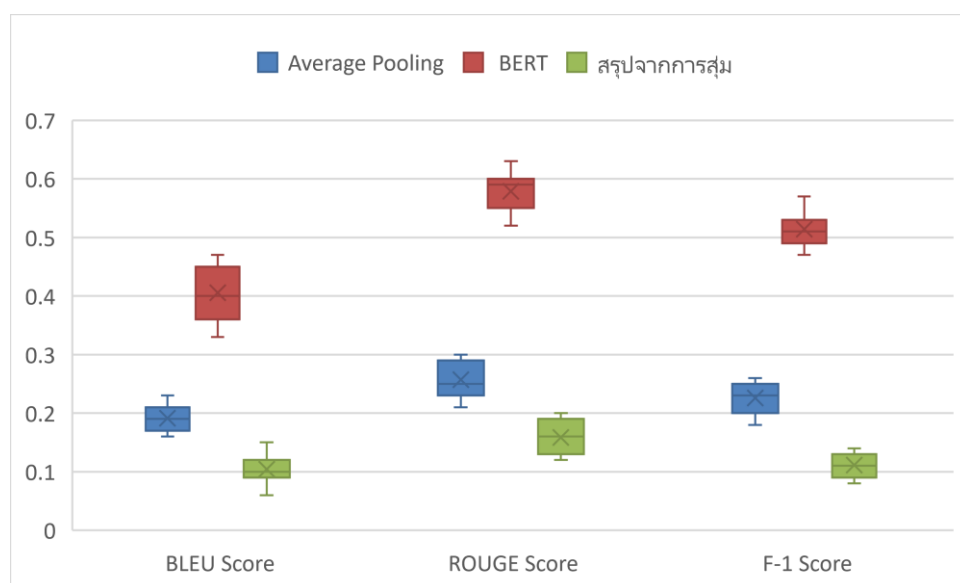
[['ชีวิต', 'ที่ผ่านมา', 'ของ', 'คุณ', 'เป็น', 'อย่างไรบ้าง', '?'], ['ตั้งแต่', 'ปฐมวัย', '-', 'ประถมศึกษา', '-', 'มัธยมศึกษา', '-', 'มหาวิทยาลัย', '-', 'ทำงาน', '-', 'เปลี่ยน', 'งาน', '-', 'มีลูก', ' ฯลฯ'], ['ลอง', 'นำ', 'ความทรงจำ', 'ช่วง', 'นั้น', 'ขึ้น', 'มา', 'ทบทวน', 'ตัวเอง', 'กัน', 'อีกครั้ง', 'เพราะ', 'ตัวตน', 'ที่', 'ถูก', 'หล่อหลอม', 'ขึ้น', 'มาจาก', 'เหตุการณ์', 'สำคัญ', 'ใน', 'ทุก', '7', 'ปี'], ['อาจ', 'สะท้อน', 'ให้', 'เห็น', 'อีกครั้ง', 'โดยที่', 'คุณ', 'ก็', 'ไม่', 'รู้ตัว', 'ใน', 'อีก', 'หลาย', 'ปี', 'ให้หลัง']]



[['ชีวิต', 'ที่ผ่านมา', 'อย่างไรบ้าง', '?'], ['ปฐมวัย', '-', 'ประถมศึกษา', '-', 'มัธยมศึกษา', '-', 'มหาวิทยาลัย', '-', 'ทำงาน', '-', 'งาน', '-', 'มีลูก'], ['ลอง', 'ความทรงจำ', 'ทบทวน', 'ตัวเอง', 'อีกครั้ง', 'ตัวตน', 'หล่อหลอม', 'มาจาก', 'เหตุการณ์', '7', 'ปี'], ['สะท้อน', 'อีกครั้ง', 'โดยที่', 'รู้ตัว', 'ปี', 'ให้หลัง']]

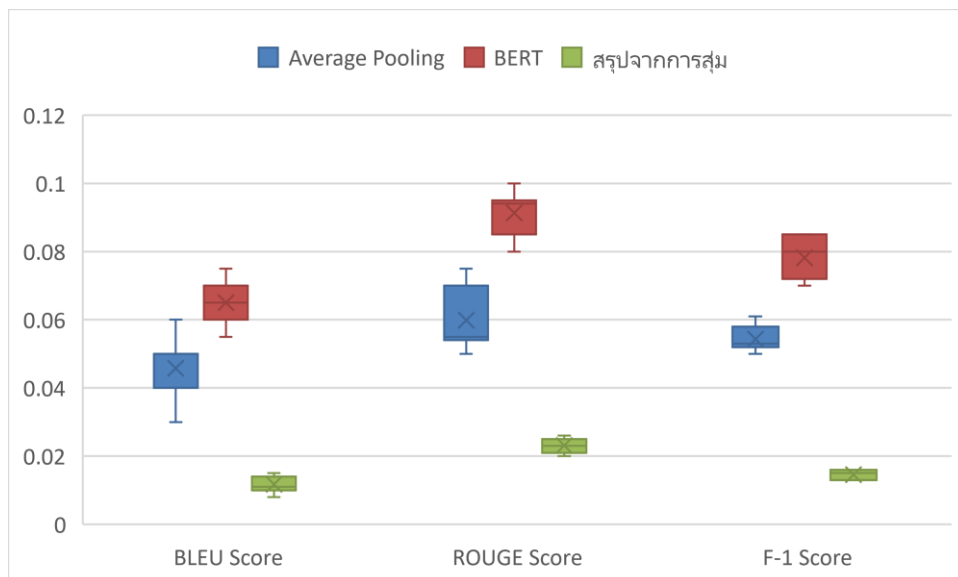
4.2 การวัดประสิทธิภาพของแบบจำลอง

หลังจากที่ได้สรุปจากแบบจำลองออกมาแล้ว จึงใช้ BLEU Score ROUGE Score และ F-1 Score ในการวัดประสิทธิภาพของแบบจำลอง โดยใช้ทั้งแบบ 1-gram และ 2-gram เพื่อวัดผล และเปรียบเทียบสรุปของแบบจำลอง ว่ามีความใกล้เคียงมากน้อยแค่ไหนกับสรุปจากเว็บไซต์ (Gold Summary) พบว่าคะแนนที่วัดได้จากการใช้ BERT ให้ผลคะแนนที่ดีกว่าการใช้ Average Pooling เนื่องจาก Embedding ของ Sentence ที่ได้มาจาก BERT นั้นสามารถเก็บ และให้ความหมายจากคำอื่น ๆ ภายในประโยคได้ด้วย ซึ่งแตกต่างจาก Average Pooling ที่เป็นการนำมาหาค่าเฉลี่ย ซึ่งจะให้ความหมายของ



แต่ค่าภายในประโยคจะถูกลดทอนบทบาทลงไป

รูปภาพที่ 4.4 Box Plot แสดงผล BLEU score ROUGE Score และ F1-Score แบบ 1-gram ของสรุปที่มาจากแบบจำลองที่ใช้ Average Pooling (สีน้ำเงิน) สรุปที่มาจากแบบจำลองที่ใช้ BERT (สีแดง) และสรุปที่มาจากการสุ่ม (สีเขียว) เมื่อเปรียบเทียบกับสรุปจากเว็บไซต์ (Gold Summary)



รูปภาพที่ 4.5 Box Plot แสดงผล BLEU score ROUGE Score และ F1-Score แบบ 2-gram ของสรุปที่มาจากแบบจำลองที่ใช้ Average Pooling (สีน้ำเงิน) สรุปที่มาจากแบบจำลองที่ใช้ BERT (สีแดง) และสรุปที่มาจากการสุ่ม (สีเขียว) เมื่อเปรียบเทียบกับสรุปจากเว็บไซต์ (Gold Summary)

ตัวอย่างบทความ และประโยคที่ถูกเลือกโดยแบบจำลอง

ตัวอย่างบทความที่ถูกเก็บรวบรวมข้อมูลมา

ชีวิตที่ผ่านมาของคุณเป็นอย่างไรบ้าง ? ตั้งแต่ปฐมวัย-ประถมศึกษา-มัธยมศึกษา-มหาวิทยาลัย-ทำงาน-เปลี่ยนงาน-มีลูก ฯลฯ ลองนำความทรงจำช่วงนั้นขึ้นมาทบทวนตัวเองกันอีกครั้ง เพราะตัวตนที่ถูกหล่อหลอมขึ้นมาจากเหตุการณ์สำคัญในทุก 7 ปี อาจสะท้อนให้เห็นอีกครั้งโดยที่คุณก็ไม่ได้รู้ตัวในอีกหลายปีให้หลัง

ตลอดชีวิตที่ผ่านมา หลายคนอาจไม่เคยมีเวลาสงบๆ ให้ตัวเองได้นั่งพักแล้วคิดทบทวนการเติบโตของชีวิตเลยสักครั้ง อาจไม่เคยกลับมาตั้งหลักพิจารณาอดีตก่อนจะตัดสินใจก้าวต่อไปในอนาคต อาจพบกับปัญหาเดิมซ้ำซากในทุกความสัมพันธ์ที่แก้เท่าไรก็ไม่หาย นั่นอาจเพราะเราไม่เคยมองเห็นนิสัยบางอย่างของตัวเองที่สะสมมาตั้งแต่เด็กจนโตอย่างแนบเนียน และถ้าเรายังไม่ให้ความสำคัญกับการสำรวจตัวเอง เราอาจต้องใช้ชีวิตอยู่กับคำถามในใจที่ไม่เคยมีใครตอบไปตลอดหรือเปล่า?

ตัวผู้เขียนเองก็เคยอยู่ในจุดที่มองเห็นเพียงแค่สิ่งรอบกายแล้วตอบสนองกับมันไปตามสถานการณ์อย่างร่างที่ไร้จิตวิญญาณ ไม่เคยย้อนถามตัวเองว่าทำไมถึงทำแบบนั้น? ทำไมถึงคิดแบบนี้? หรือแท้จริงแล้วต้องการอะไรกันแน่? แต่มีวิชาหนึ่งที่ตอบปัญหาที่ค้างคาในใจจนหมดสิ้นคือ วิชาการสื่อสารภายในบุคคล (Intrapersonal Communication) ซึ่งทำให้ได้รู้จักกับครุคนหนึ่งที่มีารื้อฟื้นความทรงจำในวัยเด็กที่สะท้อนมาถึงตัวตนของเราในปัจจุบันผ่านการสำรวจ Biography หรือชีวประวัตินั่นเอง

เมื่อมีโอกาสจึงอยากชวน ครูอีก ๓ ญาติ ญาติครู คุณครูผู้สอนการทำความรู้จักตัวเองผ่าน Biography มาแบ่งปันผู้อ่าน เพราะคิดว่าอาจช่วยให้หลายคนสามารถค้นพบตัวตนบางอย่างจากการสำรวจ Biography และอาจคลี่คลายปมบางอย่างที่ตัวเองซ่อนไว้จนหาต้นเหตุไม่เจอ

ปัจจุบันครูอีกเป็นอาจารย์สอนนิสิตครุศาสตร์และนิสิตนิเทศศาสตร์ รวมทั้งทำงานเกี่ยวกับเด็กปฐมวัยด้วย แต่ก่อนหน้านั้นเธอผ่านการประกอบอาชีพมาอย่างหลากหลาย ไม่ว่าจะเป็น แอร์โฮสเตส ดิวเตอร์ นักแสดง ถ่ายโฆษณา แอคติ้งโค้ช หรือแม้กระทั่งเด็กเสิร์ฟ แต่ท้ายที่สุดเธอก็ได้

คำตอบกับตัวเองว่าอาชีพในฝันของเธอจริงๆก็คือการเป็นครูสอนเด็กปฐมวัย จึงลาออกจากการเป็นแอร์โฮสเตสเพื่อมาทำตามความฝัน

ครูอีกเรียนจบปริญญาตรีคณะนิเทศศาสตร์ ภาควิชาวาริชวิทยาและสื่อสารการแสดง เอกสื่อสารการแสดงมาก่อนที่จะไปเรียนต่อทางครุศาสตร์ปฐมวัย นั้นทำให้เธอมีความรู้สำหรับการสอนนักเรียนทั้งสองคณะ แต่สิ่งที่น่าสนใจคือการที่เธอนำ “ละคร” มาใช้สอน “เด็กปฐมวัย”

เธอเล่าเกี่ยวกับสาขาวิชาสื่อสารการแสดงที่เธอเรียนจบออกมา แต่อีกวิชาหนึ่งที่เธอให้ความสำคัญไม่แพ้กันก็คือวาริชวิทยา

เธออธิบายว่าวาริชวิทยาและการละครไม่ได้แยกออกจากกัน ละครเป็นสิ่งที่ช่วยให้ครูเข้าใจนักเรียน ขณะเดียวกันวาริชวิทยาก็เป็นเหมือนเครื่องมือในการสื่อสารกับนักเรียน ซึ่งทั้งสองศาสตร์แทรกซึมอยู่ในชีวิตประจำวันของคนอยู่แล้ว ละครจะช่วยให้เรามองเห็นความคิดและความต้องการของอีกฝ่ายผ่านการสวมบทบาทให้เขาเป็นตัวละครหนึ่ง ส่วนวาริชศาสตร์คือการเลือกว่าจะใช้วิธีการไหนเพื่อให้สื่อสารได้ตรงวัตถุประสงค์

เธอยกตัวอย่างเด็กเล็กคนหนึ่งที่ตั้งของจากมือเพื่อน แล้วเพื่อนคนนั้นก็เตรียมมาละวาด ครูจึงต้องใช้ศาสตร์ทางวาริชวิทยามาอธิบายให้เด็กๆ ฟังว่า “เหตุผลที่เพื่อนดึงของจากมือเราเพราะ ‘เขายังใช้ภาษาไม่ค่อยได้ จึงต้องใช้ภาษากายแทน’ เด็กคนที่แย่งของไปไม่ใช่เด็กนิสัยไม่ดีแต่เพราะเขามีข้อจำกัดบางอย่างนี่คือการใช้ละครในการทำความเข้าใจเขา”

ในการสอนเด็กโตอย่างนิสิตคณะครุศาสตร์ เอกปฐมวัย เธอก็ได้ประยุกต์ศาสตร์แห่งวาริชวิทยาและสื่อสารการแสดงมาสอนนิสิตในการวางแผนการสอนให้มี “ความเป็นมนุษย์” มากขึ้น ครูอีกแบ่งปันกระบวนการวางแผนการสอนว่า เธอจะออกแบบเหมือนพีระมิดของพล็อตละคร มีตั้งแต่จุดเริ่มต้นได้ไปจนถึงไคลแมกซ์ เจอปมขัดแย้งต่างๆ ค่อยๆ คลี่คลาย แล้วก็จบ เพื่อให้ผู้เรียนรู้สึกสนุกมากกว่าแผนการแบบทั่วไป และเธอก็ได้แบ่งปันเทคนิคนี้ให้กับนิสิตคณะครุศาสตร์เพื่อไปสอนนักเรียนของพวกเขาให้สนุกกับบทเรียนมากขึ้น

นอกจากนี้เธอยังฝากคำถามชวนคิดทิ้งท้ายให้นิสิตได้หาคำตอบว่าพวกเขา “คุณเป็นครูแบบไหน” โดยให้นิสิตลองนึกย้อนกลับไปว่าวิธีการสอนของพวกเขาได้รับอิทธิพลมาจากไหน แล้วสไตล์การสอนแบบนี้มีผลดีผลเสียกับเด็กอย่างไรบ้าง ครูอีกให้เหตุผลว่า “ครูต้องทบทวนบทบาทของตัวเองให้ดี เพราะการจะเป็นครูที่ดีนั้นต้องเข้าใจตัวเองก่อน แล้วจึงจะเข้าใจคนอื่นได้”

บทเรียนเรื่อง Biography นั้นอยู่ในวิชาการสื่อสารภายในบุคคล เป็นวิชาของภาควิชวาริชวิทยา ซึ่งภาควิชานี้ถูกจัดรวมอยู่ในสาขาเดียวกับสื่อสารการแสดงเพราะทั้งสองวิชามีความคาบเกี่ยวกันในด้านการทำงานเข้าใจเกี่ยวกับมนุษย์ ภาควิชวาริชวิทยาจะมุ่งเน้นการสอนเรื่องการสื่อสารภายใน

ตัวเอง การสื่อสารกับคนรอบข้าง หรือแม้แต่การสื่อสารต่อสาธารณะ ส่วนภาควิชาสื่อสารการแสดง สอนเรื่องการสื่อสารกับผู้ชมผ่านบทบาทสมมติจากบทที่แต่งขึ้น

Biography หรือทฤษฎีการสำรวจชีวประวัติของตัวเอง ถือเป็นเครื่องมือหนึ่งที่ใช้ทำความเข้าใจเกี่ยวกับตนเอง คล้ายการเขียนไดอารีรายวัน แต่เป็นก่อนย้อนอ่านชีวิตตั้งแต่เกิดมาจนถึงปัจจุบัน เป็นทฤษฎีที่อยู่ในรากฐานของมนุษยปรัชญา ที่เชื่อว่าทุก 7 ปี ชีวิตมนุษย์จะเปลี่ยน โดยจะเปลี่ยนไปอย่างไรขึ้นอยู่กับ การเติมเต็มความต้องการในช่วงวัยหนึ่งๆ

ครูอีกยกกรณีศึกษาที่เห็นในสังคมมาตั้งคำถาม ทำไมคนที่มีทุกอย่างพร้อมจึงลุกขึ้นมาทำสิ่งที่เราคาดไม่ถึง หรือสิ่งที่ไม่ควรทำ นั่นอาจเป็นผลมาจากประสบการณ์วัยก่อนหน้าที่เขาเคยพบเจอมากก็ได้

ครูอีกแสดงความคิดเห็นอย่างมีความหวังว่า “วิชาการสื่อสารภายในบุคคลควรเป็นวิชาที่ถูกบรรจุให้ทุกหลักสูตรของทุกคณะ เพราะไม่ว่าโตไปเป็นอาชีพใด ก็ไม่ควรละเลยการทำความเข้าใจตัวเอง” โดยช่วงวัยที่ครูอีกแนะนำให้เรียนการทบทวน Biography นั้น อย่างน้อยที่สุดผู้เรียนควรอยู่ในช่วง 15-21 ปี แต่ถ้าจะให้ดีที่สุดควรจะเรียนในวัยที่ผ่านโลกมามากกว่า 20 ปีขึ้นไป

จากนั้นเธอก็พานิสิตร่วมเดินทางไปกับทฤษฎี Biography ว่าทุกๆ 7 ปีชีวิตคนเราจะเปลี่ยนแปลงไปอย่างไรบ้าง โดยสิ่งหลักๆ ที่เปลี่ยนไปคือวิธีการมองโลกและความคาดหวังที่จะได้รับจากคนรอบข้าง รวมถึงความคาดหวังต่อชีวิตของตัวเอง

0-7 ปี The world is good (โลกนี้ดี) : ช่วงวัยแห่งการสะท้อน ทำอะไรให้ลูก ลูกจะทำสิ่งนั้นตอบกลับมา พ่อแม่จึงต้องเป็นแบบอย่างที่ดีให้เขา แสดงให้เห็นว่าโลกนี้มีสิ่งดีๆ ที่รอเขาอยู่ เขาต้องการรู้ขอบเขตว่าสิ่งใดทำได้สิ่งใดทำไม่ได้ และต้องการจะรู้ว่าโลกนี้ไว้วางใจได้ คนใกล้ชิดจึงสำคัญมาก นอกจากนี้สิ่งที่เด็กซึมซับในช่วงวัยนี้จะสะท้อนให้เห็นเมื่อเขาอายุ 56-63 ปี

8-14 ปี The world is beautiful (โลกนี้ช่างงดงาม) : ช่วงวัยนี้จะเริ่มมีความรัก เริ่มเข้าใจเรื่องเพศมากขึ้น ช่วงอายุ 9 ปีเป็นช่วงที่เด็กตื่นตัวต่อสิ่งรอบข้าง เขาเริ่มสังเกตว่าทุกสิ่งไม่ใช่เรื่องจริงทั้งหมด แต่โลกก็ยังคงงดงาม

15-21 ปี The world is true (โลกนี้คือความจริง) : ช่วงวัยนี้หากพ่อแม่ทำผิดต้องแสดงให้เห็นให้ลูกเห็นว่าควรแก้ไขยังไง ไม่ใช่ผิดแล้วหมกเม็ด การโกหกหรือผิดสัญญาจะทำให้ลูกเริ่มหันหลังให้ เพราะสิ่งที่เขาตามหาในช่วงวัยนี้คือแบบอย่างแท้ๆ ในการดำเนินชีวิต หากครูสอนผิดหรือทำผิดแล้วไม่ยอมรับจะทำให้เด็กปิดใจไม่เปิดรับอีก เพราะเขารู้สึกว่าครูไม่สามารถมอบความจริงให้เขาได้ นอกจากนี้วัย 16 ปียังเป็นช่วงวัยที่เด็กเริ่มขบถ เช่น การโดดเรียน การทำสีผม การฉีกกรอบบังคับบางอย่าง เพื่อที่จะแสดงให้เห็นว่าเขามีตัวตนอยู่ในโลกใบนี้

22-28 ปี The world is interesting (โลกนี้น่าสนใจ) : เป็นช่วงวัยที่อยากจะออกไปสำรวจโลกภายนอก อยากค้นหาสีสันในชีวิต อยากลองผจญภัยและใช้ชีวิตในแบบของตัวเอง

29-35 ปี The world is feasible (โลกนี้มีความเป็นไปได้) : เป็นช่วงเวลาที่เราใช้จิตวิญญานในการรับรู้โลก เป็นวัยที่มีทั้งกำลังและสติปัญญา หากจุดสมดุลระหว่างความฝันและความจริง แสวงหาคำตอบอื่นๆในชีวิต เป็นช่วงที่คนเริ่มหางานที่จะอยู่กับมันตลอดไป

36-42 ปี The world is contradictory (โลกกำลังกลับด้าน) : เป็นช่วงวัยที่ค้นพบสิ่งสำคัญหรือความหมายที่แท้จริงของชีวิต อยู่ในช่วงหัวเลี้ยวหัวต่อว่าจะเลือกชีวิตแบบไหนให้ตัวเอง

ครูอีกชี้ให้เห็นว่าคำตอบของแต่ละคนจะออกมาหลากหลายแทบไม่ซ้ำกัน นั่นเพราะทุกคนได้รับการเติมเต็มความต้องการที่ไม่เท่ากัน ต่อให้มีพ่อแม่คนเดียวหรือประสบสถานการณ์เดียวกันทุกอย่าง แต่เกิดขึ้นคนละช่วงวัยผลลัพธ์ก็อาจจะออกมาไม่เหมือนกัน

หากอิงตามทฤษฎี Biography เด็กที่พ่อแม่หย่ากันในช่วง 0-7 ปี The world is good (โลกนี้มันดี) ซึ่งเป็นวัยที่ต้องการรับรู้ว่ามีสิ่งดีๆ รอเขาอยู่และต้องการเชื่อใจใครสักคน อาจทำให้โลกของเขาทั้งใบอาจพังทลายลง แต่หากเหตุการณ์เดียวกันนี้เกิดขึ้นกับเด็กในวัย 8-14 ปี The world is beautiful (โลกนี้ช่างงดงาม) ซึ่งเขาเริ่มแยกแยะความจริงต่างๆ ได้แล้ว เขาอาจจะรับมือกับการหย่าร้างของพ่อแม่ได้ และอาจมองเห็นความงดงามของความรักที่พ่อแม่ยังคงมอบให้เขาต่อไป แม้พวกเขาจะแยกทางกัน

ถึงแม้ว่าช่วงวัยของนิสิตจะอยู่ในช่วง 20 ปี แต่การสำรวจ Biography ก็สามารถทำให้พวกเขามองเห็นตัวเองในอดีตและมองเห็นความเป็นไปได้ในอนาคตจากการสังเกตและสวมบทบาทเป็นคนในครอบครัว

ตอนที่ครูอีกเป็นนักแสดง เธอได้ลองสวมบทบาทของตัวเองลงจึงทำให้เธอเข้าใจการกระทำของตัวเองที่คนอื่นอาจไม่เข้าใจ หรือตัดสินใจว่าเป็นเรื่องผิด เธอคิดว่านั่นเป็นเพราะทุกคนก็ต่างมองในมุมของตัวเอง **จึงมีเหตุผลและการกระทำที่ต่างกันออกไป**

ในห้องเรียนครูอีกจะให้นิสิตที่สมัครใจช่วยแบ่งปันเกี่ยวกับการกระทำของพ่อแม่ที่พวกเขาเคยตั้งคำถามหรือรู้สึกต่อต้าน แต่ปัจจุบันรู้สึกเข้าใจพ่อแม่มากขึ้นเพราะค้นพบเหตุผลที่แท้จริงของพวกเขา

มีนิสิตแสดงความคิดเห็นกันหลากหลาย เช่น พ่อแม่ของบางคนในช่วงชีวิต 0-21 ปีประสบกับความยากลำบากในชีวิตหลายประการ จึงนำความคาดหวังที่จะเติมเต็มความสุขสบายของเขามาลงที่ลูก ในยามที่เป็นพ่อแม่ โดยที่เขาอาจไม่รู้ตัวที่กำลังใช้ลูกเพื่อเติมเต็มสิ่งที่ขาดหรือกำลังเอาความฝันของตัวเองมายัดให้ลูก

ครูอีกเองก็แบ่งปันให้นิสิตฟังว่า “คุณแม่ของครูมักเกรงใจ จนไม่ร้องขอเพื่อความต้องการของตัวเอง” ซึ่งทำให้บางครั้งครูก็รู้สึกว่าเป็นเรื่องยากในการคาดเดาความต้องการของคุณแม่ **เมื่อได้รู้จักกับ Biography** เธอไม่เพียงแต่ย้อนความทรงจำไปยังวัยเด็กของตัวเองอย่างเดียว แต่พยายามย้อนไป

มองว่าทำไมแม่ของเธอถึงชอบเสียสละ จนได้พบคำตอบว่าแม่ของเธอมีพี่น้องหลายคน จึงต้องแสดงออกว่าตัวเองไม่ต้องการเพื่อให้ยายสบายใจ เมื่อเธอทราบข้อเท็จจริงนี้จึงพยายามปรับมุมมองใหม่ แล้วเริ่มบอกความต้องการที่แท้จริงกับลูกมากขึ้น เริ่มปรึกษาเรื่องต่างๆ กับลูก และเธอก็ได้ค้นพบว่าปัญหาบางอย่างที่ผู้ใหญ่กังวล เด็กๆ อาจแก้ไขมันได้อย่างง่ายดาย

ครูอึ้งก็กล่าวอีกว่า “พ่อแม่มักหลงลืมว่าตอนที่ตัวเองเป็นเด็กนั้นรู้สึกยังไง เพราะความคาดหวังจากทั้งตัวเองและคนรอบข้างทำให้พ่อแม่กลัวที่จะเป็นคนอ่อนแอ กลัวว่าจะรู้่น้อยกว่าลูก กลัวที่จะยอมรับกับลูกว่าตัวเองกำลังเสียใจ บางครั้งผู้ใหญ่ก็คิดมากเกินไป ลูกอาจจะไม่ได้ต้องการพ่อแม่ที่สมบูรณ์แบบทุกอย่าง เขาอาจจะต้องการใครสักคนที่สามารถพูดคุยกันได้”

นอกจากกิจกรรมชวนมองพ่อแม่ที่มีจุดประสงค์ให้นิสิตมองเห็นจุดเริ่มต้นของผู้เลี้ยงดูพวกเขาแล้ว ครูอึ้งก็ชวนให้นิสิตแบ่งปัน “การขบถในวัย 15-21 ปี” ของพวกเขา เพื่อให้นิสิตลองเปิดใจทำความเข้าใจกับตัวเอง ทำความเข้าใจกับบริบทของครอบครัวหรือสังคมที่ตัวเองเติบโตมา

โดยคำตอบของนิสิตออกมาหลากหลาย ไม่ว่าจะเป็นการโดดเรียน การปลอมลายเซ็น ผู้ปกครอง การทำสีผม หรือแม้แต่การกลับบ้านช้า ซึ่งครูอึ้งให้ความเห็นว่า “การแบ่งปันวิถีชีวิตของตัวเองนี้ไม่ได้ชี้ให้เห็นว่าวีรกรรมของใครดีหรือแย่กว่ากัน แต่แสดงให้เห็นถึงกฎเกณฑ์และขอบเขตในวัยเด็กที่ทำให้เด็กๆ รู้สึกว่าการลงทำบางอย่างเป็นเรื่องต้องห้าม เช่น การกลับบ้านช้าดูเป็นเรื่องจืดจางแต่อาจเป็นขบถใหญ่หลวงของเด็กบางคน”

ผู้เขียนเองก็ได้แบ่งปันการขบถในวัย 15-21 ปีในห้องเรียนเหมือนกัน มีอยู่ 2 ข้อคือ การออกจากบ้านที่ต่างจังหวัดมาเรียนที่กรุงเทพฯ และ “การเปลี่ยนศาสนา” เรื่องทั้งหมดเกิดขึ้นจากความรู้สึกอึดอัดใจที่ไม่สามารถเป็นตัวของตัวเองได้เวลาอยู่บ้าน และรู้สึกว่าบ้านไม่ใช่พื้นที่ปลอดภัย เราสังเกตตัวเองว่าเวลาอยู่โรงเรียนเราจะมีความเป็นผู้นำ มีความกล้าแสดงออก ผิดกับตอนอยู่บ้านที่กลายเป็นคนเก็บตัวเงียบ ซึ่งคิดว่าการแสดงออกที่แตกต่างออกไปเวลาอยู่บ้านเกิดขึ้นจากนิสัยของคุณพ่อที่ชอบเอาชนะ ยกตัวอย่างเช่น เราชอบบะหมี่น้ำ พ่อก็จะบอกว่าบะหมี่แห้งอร่อยกว่าทำไมไม่กิน เราต้มบะหมี่กึ่งสำเร็จรูปโดยใส่เส้นก่อนผงปรุงรส พ่อก็จะบอกว่าต้มไม่เป็นแบบนี้เส้นมันก็จืดสิ

ผู้เขียนจึงตัดสินใจเดินทางมาสอบเข้าโรงเรียนมัธยมปลายในกรุงเทพฯ เพื่อจะได้อยู่หอพักคนเดียวและใช้ชีวิตตามใจตัวเอง เมื่อย้ายมาอยู่ที่ตัดสินใจเข้าโบสถ์ไปเรียนศาสนาอยู่ 1 ปี จนได้รับการบัพติศมาหรือพิธีรับเข้าเป็นคริสต์ศาสนิกชน ทั้งๆที่คุณแม่เป็นคนเคร่งศาสนา(พุทธ)มาก ซึ่งนี่ถือเป็นความภาคภูมิใจอย่างหนึ่งที่ได้เลือกทำตามความชื่นชอบและความต้องการของตัวเอง

ในห้องเรียนแห่งนี้ นิสิตทุกคนได้ย้อนกลับไปสู่การเติบโตของพ่อแม่ และอดีตอันไกลของพวกเขา ซึ่งนั่นก็ทำให้พวกเขาก็เริ่มเห็นภาพของชีวิตตัวเองแต่ละฉากชัดเจนขึ้น ภาพชีวิตเหล่านั้นนี่เองที่เป็นจิ๊กซอร์ว่บอกรับตัวตนของเขาในปัจจุบัน ซึ่งการจะระบุตัวเองว่าเป็นคนแบบไหน ได้รับอิทธิพล

มาตั้งแต่ตอนไหน หรือเหตุการณ์ใดที่พัฒนาให้นิสัยบางอย่างปรากฏขึ้นกับตัวเรา ก็ขึ้นอยู่กับวิถีชีวิตของแต่ละคน แต่อย่างน้อยสิ่งที่มีนิสัยกลุ่มนี้ได้รับจากคาบเรียน คือชิ้นส่วนความทรงจำที่เคยเลื่อนรางหรืออาจเป็นชิ้นส่วนของความทรงจำที่เขาเผลอทำหล่นหายไปในอดีต

การเรียนรู้ Biography ของตัวเองไม่ใช่เพื่อทำนายอนาคต แต่เพื่อย้อนไปมองตัวเองในอดีตทุกๆ 7 ปี แล้วเตรียมตัวรับมือกับสิ่งที่อาจสะท้อนในช่วงวัยต่อมา เป้าหมายหลักในการสอนเรื่องนี้ให้นิสิตก็เพื่อให้เขาทำความเข้าใจชีวิตตัวเอง น้อยวิชาที่สอนให้เด็กรู้จักตัวเอง แต่คนทั้งโลกกลับถามซัดเขาไม่หยุดว่า “คุณทำอะไรได้บ้าง” โดยที่ไม่มีใครสักคนพาพวกเขาย้อนกลับไปทบทวนตัวเอง

ตัวอย่างสรุปจากแบบจำลอง

เป็นช่วงคนหางานที่จะ 36-42 ปี The world is contradictory (โลกกลับด้าน): เป็นช่วงวัยค้นพบความหมายชีวิตหัวเลี้ยวหัวต่อเลือกชีวิตตัวเอง

ครูอีกชี้ให้เห็นคำตอบแต่ละคนออกมาหลากหลายแทบซ้ำ เตรียมตัวรับมือสะท้อนวัยต่อมา เป้าหมายหลักการสอนเรื่องนิสัยทำความเข้าใจชีวิตตัวเองวิชาสอนเด็กรู้จักตัวเอง คนทั้งโลกถามซัดหยุดทำโดยที่สักคนย้อนกลับทบทวนตัวเอง

ให้ความสำคัญการสำรวจตัวเอง รู้จัก Biography ครูสอนทำผิคนักเรียนใจเปิดรับ ปฐมวัย-ประถมศึกษา-มัธยมศึกษา-มหาวิทยาลัย-ทำงาน-เปลี่ยนงาน-มีลูก มีเหตุการณ์กระทำต่างกัน

ห้องเรียนครูอีกนิสัยสมัครใจแบ่งปันการกระทำพ่อแม่ตั้งคำถามรู้สึกต่อต้าน รู้สึกอกพ่อแม่มากขึ้นค้นพบท่าน

นิสัยแสดงความคิดเห็นหลากหลายพ่อแม่คนในช่วงชีวิต 0-21 ปีความยากลำบากชีวิตอาชีพได้

ตัวอย่างการเปรียบเทียบคำที่ตรงกันระหว่างสรุปจากโมเดล และสรุปจากเว็บไซต์ (Gold Summary) แบบ 1-gram

ตัวอย่างสรุปจากแบบจำลอง

เป็นช่วงคน(1)ทำงานที่จะ(2) 36-42ปี The world is contradictory (โลกกลับด้าน): เป็นช่วงวัยค้นพบความหมายชีวิต(3)หัวเลี้ยวหัวต่อเลือกชีวิตตัวเอง(4)

ครู(5)อีก(6)ชี้ให้เห็นคำตอบแต่ละคนออกมาหลากหลายแทบซ้ำเตรียมตัวรับมือสะท้อน(7)วัยต่อมาเป้าหมายหลักการสอนเรื่องนิสิตทำความเข้าใจชีวิตตัวเอง(8)วิชา(9)สอนเด็ก(10)รู้จักตัวเอง(11)คน(12)ทั้งโลก(13)ถาม(14)ซัก(15)หยุด(16)ทำ(17)โดยที่(18)สัก(19)คน(20)ย้อนกลับ(21)พบทวน(22)ตัวเอง(23)

ให้ความสำคัญการสำรวจตัวเองรู้จัก Biography(24) ครูสอนทำผิคนักเรียนใจเปิดรับ ปฐมวัย-ประถมศึกษา-มัธยมศึกษา-มหาวิทยาลัย-ทำงาน-เปลี่ยนงาน-มีลูก มีเหตุผลการกระทำต่างกัน

ห้องเรียนครูอีกนิสิตสมัครใจแบ่งปันการกระทำพ่อแม่ตั้งคำถามรู้สึกต่อต้านรู้สึกอกพ่อแม่มากขึ้นค้นพบท่าน

นิสิตแสดงความคิดเห็นหลากหลายพ่อแม่คนในช่วงชีวิต(25)0-21ปีความยากลำบากชีวิตอาชีพใด

ตัวอย่างสรุปจากเว็บไซต์ (Gold Summary)

คน(1)ทั้งโลก(13)ถาม(14)ซัก(15)เด็ก(10)ไม่หยุด(16)ว่า “คุณทำ(17)อะไรได้บ้าง” โดยที่(18)ไม่มีใครสัก(19)คน(12) หรือวิชา(9)ไหนพาพวกเขาย้อนกลับ(21)ไปพบทวน(22)ตัวเอง(4)ชวนไปสำรวจชีวประวัติ (Biography(24)) ในวิชาการสื่อสารภายในบุคคล พร้อมเดินทางย้อนเวลาไปสู่จุดเริ่มต้นของการหล่อหลอมตัวตนตั้งแต่วัยเด็ก มาจนถึงการสะท้อน(7)ตัวตนในปัจจุบันไปพร้อมๆกัน กับครู(5)อีก(6)ณัฐณี เจียรกุลในช่วงชีวิต(25)ของทุกคน(20)ควรมีโอกาสได้พบทวนชีวิต(3)ตัวเอง(8) เพื่อกลับมาตั้งหลักและยืนอยู่บนเท้าของตัวเอง(11)อีกครั้ง เตรียมพร้อมตัวเอง(23)วันนี้ในการที่จะ(2)ก้าวต่อไป ไม่งั้นเราก็จะไม่รู้เลยว่าเราก้าวต่อไปเรื่อยๆ เพื่ออะไร

บทที่ 5

ข้อสรุปและข้อเสนอแนะ

ในบทนี้จะกล่าวถึง สรุปผลจากข้อมูลผลการวิจัยในบทที่ 4 ข้อสรุปของการสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง รวมถึงข้อเสนอแนะอื่นๆ ที่ได้จากการวิจัย

5.1 ข้อสรุป

จากผลการวิจัยพบว่า เมื่อเปรียบเทียบประสิทธิภาพของแบบจำลองสรุปบทความของคณะผู้วิจัย กับการสุ่ม และนำมาต่อกันเป็นบทความ จะพบว่าค่า BLEU Score, ROUGE Score และ F-1 Score ของแบบจำลองสรุปบทความของคณะผู้จัดให้คะแนนที่ดีกว่าการสุ่ม และสามารถบอกได้ว่าผลสรุปที่ได้จากแบบจำลองของคณะผู้จัดนั้นมีประสิทธิภาพที่ดีกว่าการสุ่ม และผลสรุปได้จากแบบจำลองที่ใช้ BERT นั้นให้ผลคะแนนที่ดีกว่าการทำ Average Pooling ซึ่งสามารถบอกได้ถึงประสิทธิภาพของการทำ Sentence Embedding ที่ดีกว่า

5.2 ปัญหาที่พบ และวิธีแก้ปัญหา

5.2.1 ระหว่างการทำงานของแบบจำลองที่ใช้ BERT นั้น ใช้ระยะเวลาในการทำงานเป็นระยะเวลาที่นานมาก

วิธีแก้ปัญหา: พยายามปรับแก้แบบจำลองเพื่อให้แบบจำลองใช้เวลาในการทำงานที่น้อยที่สุดเท่าที่เป็นไปได้

5.2.2 ระหว่างการทำงานของแบบจำลอง การทำ Clustering ในบางกรณี พบว่าไม่สามารถทำงานต่อไปได้

วิธีแก้ปัญหา: แก้ไขข้อผิดพลาดของแบบจำลองในการทำ Clustering ให้สามารถทำงานต่อไปได้

5.3 ข้อเสนอแนะ

5.3.1 ทดลองใช้แบบจำลองกับ Embedding ต่างๆที่มีให้ใช้ในภาษาไทย

5.3.2 นำชุดข้อมูลที่จะใช้ทำสรุปนำไปเรียนรู้ให้กับตัวแบบจำลองที่จะนำ Embedding มาใช้ เพื่อปรับปรุงเพิ่มประสิทธิภาพของ Embedding ที่จะนำมาใช้

5.3.3 ปรับปรุงและเพิ่มประสิทธิภาพการทำงานของ การตัดคำและประโยค

5.3.4 ใช้มนุษย์ในการช่วยประเมินประสิทธิภาพของบทสรุปที่ได้จากแบบจำลอง

รายการอ้างอิง

- [1] Liu, Yang, and Mirella Lapata. "Text Summarization with Pretrained Encoders." *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*.
Available from : <https://arxiv.org/pdf/1908.08345.pdf> [2019]
- [2] Liu, Yang. "Fine-tune BERT for extractive summarization." *arXiv preprint arXiv:1903.10318*
Available from : <https://arxiv.org/pdf/1903.10318.pdf> [2019]
- [3] Derek Miller. Leveraging BERT for Extractive Text Summarization on Lectures.
Available from : <https://arxiv.org/ftp/arxiv/papers/1906/1906.04165.pdf> [2019, June 7]
- [4] Peeradech Bangchareons. A Machine-Translation based Approach to Word Boundary Identification: A Projective Analogy of Bilingual Translation. [2010]
- [5] Keng Surapong. Python ตัดคำภาษาไทยด้วย PyThaiNLP.
Available from : <https://www.bualabs.com/archives/3740/python-word-tokenize-pythainlp-example-algorithm-deepcut-newmm-longest-python-pythainlp-ep-2/>
[2020, January 29]
- [6] Pattarawat Chormai. AttaCut : A Fast and Accurate Neural Thai Word Segmenter.
Available from : <https://arxiv.org/ftp/arxiv/papers/1911/1911.07056.pdf>
[2019, August 26]
- [7] Wirote Aroonmanakun. TLTK: Thai Language Toolkit.
Available from : <https://pypi.org/project/tltk> [2020, July 30]

[8] Y. Poowarawan. Dictionary-based Thai Syllable Separation

Available from : <https://www.semanticscholar.org/paper/Dictionary-based-Thai-Syllable-Separation-Poowarawan/d15dca32d2bd685c12fb9a5cdb5bbdf6ccc366b2>

[1986]

[9] Choochart Haruechaiyasak, Sarawoot Kongyoung and Matthew N. Dailey.

A Comparative Study on Thai Word Segmentation Approaches. [2008, June]

[10] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems*. [2017]

ภาคผนวก

ภาคผนวก ก

แบบเสนอหัวข้อโครงการ รายวิชา 2301399 Project Proposal

ปีการศึกษา 2563

| | |
|--------------------------|--|
| ชื่อโครงการ (ภาษาไทย) | การสรุปบทความโดยใช้หลักการเรียนรู้ของเครื่อง |
| ชื่อโครงการ (ภาษาอังกฤษ) | Article Summarization Using Machine Learning Techniques |
| อาจารย์ที่ปรึกษา | อ.ดร.นฤมล ประทานวณิช |
| ผู้ดำเนินการ | 1. นายธนเดช รุ่งศรีพัฒน์พร เลขประจำตัวนิสิต 6033625423 2. นายชลิต ปฐมกสิวัฒนา เลขประจำตัวนิสิต 6033614523 |
| | สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย |

หลักการและเหตุผล

ในปัจจุบันมีบทความต่าง ๆ มากมายถูกเขียนขึ้นมาในแต่ละวันทั้งบทความข่าวสาร บทความที่ให้ความรู้ บทความที่เกี่ยวกับภาพยนตร์ เป็นต้น เนื่องด้วยจำนวนบทความที่มีจำนวนมากมหาศาลในปัจจุบัน การสรุปบทความจึงเป็นสิ่งจำเป็นเป็นอย่างยิ่ง เพื่อให้ได้บทความที่สั้น กระชับ ได้ใจความ เพื่อให้ผู้อ่านบทความใช้เวลาไม่มาก และเข้าใจเนื้อหาที่สำคัญได้อย่างรวดเร็ว

การสรุปบทความในปัจจุบัน ส่วนใหญ่ยังคงเป็นการสรุปด้วยฝีมือของมนุษย์ ซึ่งจำเป็นต้องใช้เวลาเป็นอย่างมากในการทำสรุปแต่ละครั้ง เนื่องด้วยการสรุปโดยใช้คอมพิวเตอร์นั้น ยังไม่เป็นที่แพร่หลายนัก เพราะไม่สามารถสรุปบทความได้ถูกต้องสมบูรณ์เท่าฝีมือของมนุษย์

คณะผู้จัดทำจึงต้องการที่จะพัฒนาการสรุปบทความแบบอัตโนมัติโดยใช้หลักการเรียนรู้ของเครื่อง (Machine Learning) เพื่อเป็นอีกก้าวหนึ่งที่น่าสนใจไปสู่การสรุปบทความที่สมบูรณ์ยิ่งขึ้นในอนาคต โดยไม่จำเป็นต้องพึ่งพาการสรุปของบุคคล สามารถใช้งานในการสรุปได้อย่างถูกต้อง สะดวก และรวดเร็ว

วัตถุประสงค์

เพื่อพัฒนาวิธีการสรุปบทความอัตโนมัติโดยใช้หลักการการเรียนรู้ของเครื่อง

ความรู้ที่เกี่ยวข้อง

1. **Sentence Tokenization** คือการแบ่งข้อความเป็นประโยค แบ่งออกเป็น 2 วิธีใหญ่ ๆ คือ

1. การแบ่งข้อความเป็นประโยคโดยใช้ช่องว่าง (whitespace)

2. การแบ่งข้อความเป็นประโยคโดยใช้ช่องว่าง และการขึ้นบรรทัดใหม่ (new line) การแบ่งข้อความไปเป็นประโยคโดยใช้ Conditional Random Field (CRF) นั่นคือการใช้ sequence model ประเภทหนึ่ง ซึ่งใช้หลักการของ condition probability ในการทำงาน หลักการทำงานคือจะทำการแบ่งคำแล้วทำการทำนายความน่าจะเป็นในการกำกับคำ (label) สำหรับแต่ละคำ โดยอาศัยการกำกับคำของคำต่างๆก่อนหน้าทั้งหมด

2. **Word Tokenization** คือการแบ่งประโยคหรือข้อความไปเป็นคำ ๆ ซึ่งเรียกว่า token เพื่อที่จะสามารถนำไปประมวลผลต่อ ตัวอย่างของการแบ่งคำเป็นดังภาพที่ 1

อัลกอริทึมแต่ละแบบนั้นใช้วิธีคิดในการแบ่งคำแตกต่างกัน ซึ่งจะทำให้ผลลัพธ์ของการแบ่งคำและเวลาในการประมวลผลแตกต่างกันออกไปอีกด้วย [5] [7]

```
[12] from pythainlp.tokenize import word_tokenize
      word_tokenize(text, engine="newmm")
↳ ['ช', 'ลิต', 'อยาก', 'จะ', 'เล่น', 'เกม', 'ตลอด', 'ทั้งวัน']

      word_tokenize(text, engine='attacut')
↳ ['ชลิต', 'อยาก', 'จะ', 'เล่น', 'เกม', 'ตลอด', 'ทั้ง', 'วัน']
```

รูปที่ 1: ตัวอย่างการตัดคำโดยใช้ newmm (บน) และattacut (ล่าง)

โดยอัลกอริทึมที่ใช้แบ่งคำนั้น แบ่งออกเป็น 2 กลุ่มใหญ่ ๆ คือ

2.1 Dictionary-based

แบ่งโดยอาศัยการใช้ดิกชันนารี (dictionary) [8] และ กริดีอัลกอริทึม (Greedy algorithm) ตัวอย่างของอัลกอริทึม มีดังต่อไปนี้ Longest matching, Maximal matching ยกตัวอย่างคือ

Longest matching จะทำการนำตัวอักษรไปตรวจกับดิกชันนารี โดยทำจากซ้ายไปขวา ถ้าไม่พบคำดังกล่าวปรากฏใน dictionary ก็จะมีการตัดข้อความออก กล่าวคือ จะได้คำที่มีความหมายที่ยาวที่สุดใน dictionary (ดังรูปที่ 2) แต่ก็ยังมีปัญหาในเรื่องของคำที่ไม่อยู่ในดิกชันนารี (unknown words) เนื่องจากว่าคำสามารถเกิดขึ้นใหม่ได้ตลอดเวลา แต่สามารถแก้ได้โดยใช้ Customized dictionary นั่นคือ การกำหนดดิกชันนารี เพิ่มเติมในส่วนของคำที่ต้องการเพิ่มขึ้นมา หรือ Thai Character Clustering Rule(TCC) คือ การนำตัวอักษรในประโยคมารวมกันให้เป็นกลุ่มของตัวอักษร โดยอาศัยกฎต่างๆเพื่อใช้ในการรวมกลุ่มตัวอักษร [4]

| ลำดับการทำงาน | ส่วนของคำที่ยาวที่สุด | ส่วนที่เหลือ |
|---------------|-----------------------|---------------------------|
| 1 | แนวทาง | การแก้ปัญหาหนี้สินเกษตรกร |
| 2 | การ | แก้ปัญหาหนี้สินเกษตรกร |
| 3 | แก้ | ปัญหาหนี้สินเกษตรกร |
| 4 | ปัญหา | หนี้สินเกษตรกร |
| 5 | หนี้สิน | เกษตรกร |
| 6 | เกษตรกร | |

รูปที่ 2: ตัวอย่างการตัดคำโดยอาศัยการใช้Longest matchingกับdictionary

Source:http://thailang.nectec.or.th/halloffame/images/stories/best/download/best2010_12p34c002.pdf

2.2 Model-based

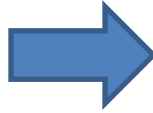
แบ่งคำโดยการอาศัยโมเดลต่าง ๆ โดยที่แต่ละโมเดลนั้นจะมีความแตกต่างกันของอัลกอริทึมที่ใช้ ทำให้ประสิทธิภาพรวมถึงระยะเวลาในการใช้นั้นก็จะแตกต่างกันออกไป ทุกโมเดลจำเป็นที่จะต้องใช้ข้อมูลในการเรียนรู้ (train) เพื่อที่จะเรียนรู้ language model ของภาษานั้น ๆ โดยข้อมูลที่น่าเข้ามา (input) เพื่อให้โมเดลเรียนรู้นั้นจะแบ่งเป็นส่วนของตัวข้อความและส่วนของคำที่แบ่งได้ (label data) หลังจากเรียนรู้เสร็จแล้วจึงนำข้อมูลอื่นๆที่โมเดลไม่เคยพบ นำมาลองให้ตัวโมเดลแบ่งคำ จะได้ข้อมูลออกมาเป็นคำต่าง ๆ (output) ซึ่งถูกแบ่งโดยโมเดล Traditional models เช่น SVM, CRF [9] , ยุคของ deep learning เช่น Sertis Bi-GRU [10] , deepcut และ attacut [6]

3. **Word Embedding** คือการกระบวนการเปลี่ยนแปลงคำ หรือวลี ให้กลายเป็นเวกเตอร์ตัวเลข เพราะว่าคอมพิวเตอร์นั้นไม่สามารถประมวลผลข้อความได้เช่นเดียวกับมนุษย์ จึงต้องมีการเปลี่ยนแปลงข้อความให้กลายเป็นตัวเลข เพื่อให้คอมพิวเตอร์สามารถประมวลผลได้ มีวิธีการดังนี้

3.1 One-hot vector คือการเปลี่ยนแปลงคำ หรือวลี ให้กลายเป็นเวกเตอร์ตัวเลขวิธีหนึ่ง โดยใช้ตัวเลขเป็นแบบ binary กล่าวคือจะใช้เลข 0 , 1 เป็นตัวแทนของการปรากฏของคำแต่ละคำภายในประโยค เป็นดังรูปที่ 3

ข้อดีคือคณะผู้จัดทำสามารถที่จะนำ one-hot vector ที่ได้มานี้ไปให้คอมพิวเตอร์ประมวลผลได้ แต่ก็ยังมีข้อเสียอยู่คือ one-hot vector นั้นมีความ sparse กล่าวคือมีการเก็บค่า 0 อยู่เยอะมาก ซึ่งทำให้ไม่มีประสิทธิภาพในการเก็บและคำนวณ นอกจากนี้ระยะห่างของแต่ละคำมีค่าเท่ากันคือ 1 ทั้ง ๆ ที่ควรจะมี ความแตกต่างกัน ยกตัวอย่างเช่น การใช้คลังคำศัพท์ที่ไม่ค่อยเหมาะสมกับตัวข้อมูลที่ต้องการจะทำ one-hot vector จะทำให้เกิดปัญหาดังกล่าวได้ ดังนั้นจึงมีการ embed แบบเวกเตอร์แบบแน่น (dense vect)

คำศัพท์: ผู้ชาย, ผู้หญิง, รัก,
เกลียด, ดี, เลว, กบ, หมูป่า, คิดถึง

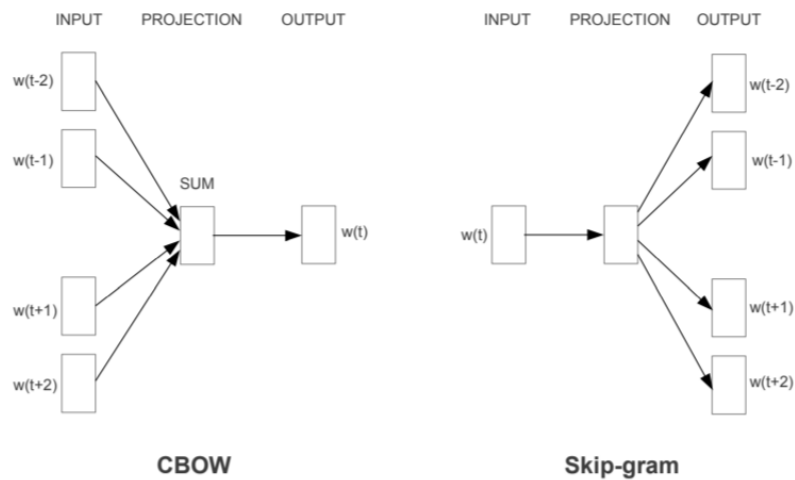


| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---------|---|---|---|---|---|---|---|---|---|
| ผู้ชาย | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ผู้หญิง | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| รัก | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| เกลียด | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| ดี | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| เลว | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| กบ | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| หมูป่า | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| คิดถึง | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

รูปที่ 3: ตัวอย่างการแปลงคำศัพท์ต่าง ๆ ภายในประโยคให้อยู่ในรูปของ one-hot vector

3.2 GloVe (Global Vectors) คือ Word Embedding โดยใช้วิธี unsupervised ในการได้เวกเตอร์ที่เป็นตัวแทนของคำมา ซึ่งทำให้ไม่ต้องอาศัยข้อมูลในการฝึกสอนมากนักแต่ก็ยังให้ผลที่ดี เป็นผลให้เวลาในการฝึกสอนนั้นรวดเร็วมากขึ้นอีกด้วย และยังอาศัยการใช้วิธีการทางสถิตินั้นคือการใช้อัตราส่วนความน่าจะเป็น (word co-occurrence) จากคลังข้อมูล (corpus)

3.3 Word2vec คือ word embedding ชนิดนี้ โดยเป็นการเปลี่ยนคำให้อยู่ในรูปของตัวเลข แต่ตัวเลขเหล่านั้นจะเก็บบริบทของคำไว้ด้วย (เป็น feature สำหรับนำไปให้โมเดลเรียนรู้ต่อไป) โดยแรกเริ่มนั้นใช้วิธี one hot encoding แล้วจึงนำข้อมูลไปใส่ลงใน Neural Network



รูปที่ 4: การทำงานของ CBOW เปรียบเทียบกับ Skip-gram

Source: *Exploiting Similarities among Languages for Machine Translation*

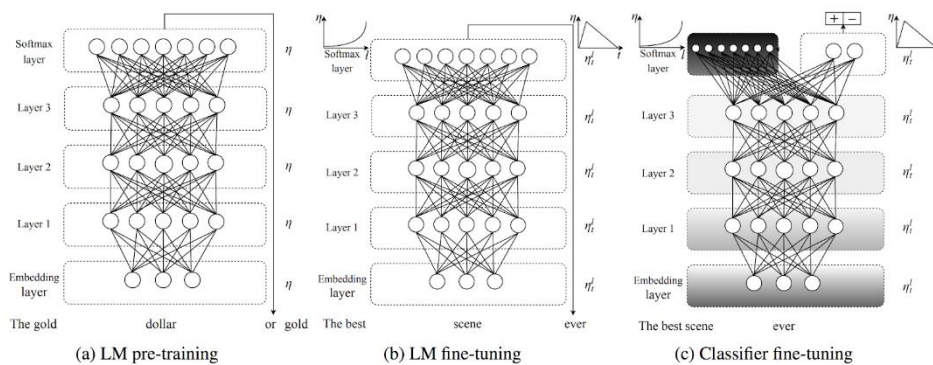
3.3.1 CBOW (Continuous bag-of-words)

CBOW คือการทำนายคำโดยอาศัยบริบทต่าง ๆ กล่าวคือผลลัพธ์ที่ได้ออกมา นั้นอาศัยบริบทจากคำที่กระจายตัวอยู่รอบ ๆ หรือคำข้าง ๆ (ใช้หลาย input ได้หนึ่ง output) โดยส่วนมากจะใช้เวลาเทรนน้อย และให้ผลที่ดีกว่า skip-gram เนื่องจากการที่การที่อ้างอิงจากบริบทของหลาย ๆ คำนั้น จะสามารถทำนายได้ผลลัพธ์แม่นยำมากยิ่งขึ้น

3.3.2 Skip-gram

Skip-gram คือการทำนายคำโดยอาศัยบริบทจากคำคำเดียว กล่าวคือผลลัพธ์ที่ได้ออกมา นั้นอาศัยบริบทจากคำคำเดียว (ใช้หนึ่ง input ได้หลาย output) ซึ่งจะทำงานได้ดีกับชุดข้อมูลขนาดเล็ก และยังสามารถแสดงคำหรือวลีที่หายากได้เป็นอย่างดี

3.4 ULMFiT (Universal Language Model Fine-tuning)



รูปที่ 5: ขั้นตอนการทำงานของ ULMFiT

Source: <https://thaikeras.com/2018/transfer-learning-nlp/>

Word vector ภายใน pythainlp นั้นมาจาก ULMFiT ซึ่งเป็น transfer learning ชนิดหนึ่ง มีส่วนประกอบหลักคือ AWD-LSTM เมื่อทำการเรียนรู้ Language Model แล้ว จึงนำไปใช้งานต่าง ๆ ได้ โดยต้องทำการ fine-tuning ก่อน คือการเทรน Language Model สำหรับงานที่จะนำไปใช้ แล้วจึงค่อยเทรนงานที่จะทำจริง ๆ

4. BERT Sentence embedding

BERT เป็น Transfer Learning ประเภทหนึ่ง เป็นการเรียนรู้ Language Model ของภาษานั้น ๆ เพื่อนำมาสร้างในส่วนของ Word Embedding, Sentence Embedding เพื่อที่จะนำไปประมวลผลต่อไป

4.1 Average Word Embedding within a Sentence

คือการหา embedding ของประโยคโดยการหาค่าเฉลี่ยของ embedding ของคำที่อยู่ภายในประโยค

4.2 BERT for Sentence Embedding

คือการใช้ Embedding ของประโยคที่นำมาจาก BER

5. K-mean clustering

เป็นวิธีแบบ Unsupervised-learning ซึ่งไม่ต้องอาศัยข้อมูลที่มี Label โดยจะทำการจัดกลุ่มของข้อมูลเป็นจำนวน k กลุ่ม มีขั้นตอน ดังนี้

1. กำหนดข้อมูลขึ้นมาเป็น 2 กลุ่ม
2. สุ่มเลือกจุด Centroid เพื่อใช้เป็นตัวแทนของแต่ละกลุ่ม
3. ทหาระยะห่างระหว่างจุดข้อมูล และจุด Centroid หากจุดข้อมูลใดอยู่ใกล้จุด Centroid ของกลุ่มนั้นๆ ถ้ามากกว่าให้ถือว่าเป็นสมาชิกของกลุ่มนั้น
4. ได้กลุ่มข้อมูลใหม่ที่ต่างจากเดิมของทั้ง 2 กลุ่ม
5. กำหนดจุด Centroid ของแต่ละกลุ่มใหม่
6. ทำวนไปเรื่อย ๆ จนกว่าจะไม่มีเปลี่ยนแปลงหรือเปลี่ยนแปลงน้อยที่สุด

6. Text summarisation

คือการสรุปข้อความจากข้อความที่ได้รับมาทั้งหมด โดยที่ยังสามารถแสดงใจความสำคัญของข้อความทั้งหมดได้ ข้อความที่นำมาใช้อาจจะมีมาจากหลายแหล่ง เช่น ข่าว บทความ เป็นต้น ดังนั้นข้อมูลที่ใช้นำเข้าคือข้อความที่ต้องการทำการสรุป และ ข้อมูลผลลัพธ์ที่ได้ก็คือข้อความที่ได้ทำการสรุปแล้ว ซึ่งสามารถนำมาใช้เพื่อให้การอ่านข่าวหรือบทความที่มีความยาวมีความสะดวก รวดเร็วที่มากยิ่งขึ้น โดยมีตัวอย่างเป็นการวิจัยในการทำสรุปข้อความในหนังสือเรียน แบ่งเป็น 2 ประเภท คือ

6.1 Extractive คือการสรุปข้อความโดยการใช้ข้อความเดิมที่มีอยู่แล้ว นำมาสร้างสรุป โดยยังคงใจความสำคัญเอาไว้

6.2 Abstractive คือการสรุปข้อความโดยการสร้างสรุปขึ้นมาโดยใช้ deep learning ในการทำ ซึ่งสรุปที่ได้มานั้นอาจจะมีเหมือนหรือต่างกับข้อความต้นฉบับก็ได้ แต่ยังคงใจความสำคัญเอาไว้ด้วย

7. Evaluation metrics

การประเมินผลของการทำสรุปนั้นใช้การวัดอยู่ 2 ประเภท ซึ่งมีความแตกต่างกัน คือ

7.1 ROUGE score โดย ROUGE score จะทำการวัดในส่วนของ Recall เป็นการบ่งบอกความแม่นยำของตัวระบบโดยสนใจว่ามีข้อความสรุปต้นฉบับปรากฏอยู่ในข้อความสรุปที่สร้างขึ้นมากเท่าไร

7.2 BLEU score โดย BLEU score จะทำการวัดในส่วนของ Precision เป็นการบ่งบอกความแม่นยำของตัวระบบโดยสนใจในส่วนของสรุปที่สร้างขึ้นมานั้นปรากฏอยู่ในข้อความสรุปต้นฉบับเท่าไร

งานวิจัยที่เกี่ยวข้อง

จากงานวิจัยที่ศึกษาโดย Yang Liu และ Mirella Lapata ในปี ค.ศ. 2019 [1] ได้ทำการทดลองสรุปข้อความ แบ่งเป็น 2 กลุ่ม คือ Extractive และ Abstractive โดยใช้ชุดข้อมูลที่มีเอกลักษณ์เฉพาะของแต่ละชุดข้อมูลที่แตกต่างกัน รวมทั้งหมด 3 ชุดข้อมูล ได้แก่ 1.CNN/Daily mail ซึ่งเป็นข้อความประเภทข่าว ทำการไฮไลท์สรุปให้แล้ว จึงเหมาะแก่การทำการสรุปข้อความแบบ Extractive มาก 2.New York Times (NYT) ซึ่งเป็นข้อความประเภทบทความ และมีการสรุปข้อความที่เป็นไปในแบบ abstractive และ 3.XSum เป็นข้อความประเภทข่าว ทำการสรุปโดยใช้ประโยคเดียวในการสรุปข้อความทั้งหมด โดยทั้งหมดนี้นำมาใช้กับ pretrained encoders ที่ชื่อว่า BERT (Bidirectional Encoder Representations from Transformers) จากนั้นจึงนำผลมาเปรียบเทียบกับแบบจำลองก่อน ๆ นั่นคือจากที่เคยใช้ในการทำสรุปข้อความผ่านการวัดประสิทธิภาพด้วยคะแนน ROUGE Score ผลปรากฏว่า BERT นั้นให้ผลที่ดีเมื่อเทียบกับแบบจำลองอื่น อีกทั้งชุดข้อมูลในแต่ละชุดนั้นก็ยังมีผลกับการทำสรุปด้วยเช่นกัน กล่าวคือในส่วนของชุดข้อมูล CNN/Dailymail นั้น BERTสามารถทำคะแนนได้ดีกว่าแบบจำลองแบบอื่น ๆ และได้ผลของคะแนนที่ใกล้เคียงกันทั้งในส่วนของแบบจำลอง Abstractive และ Extractive, ชุดข้อมูล NYT นั้น BERT ก็ยังสามารถทำคะแนนได้ดีกว่าแบบจำลองแบบอื่น ๆ และแบบ Abstractive ได้คะแนนที่ดีกว่าแบบ Extractive, XSum ในส่วนของ BERT แบบ Extractive ทำคะแนนได้ค่อนข้างแย่มากจึงไม่ได้นำผลมาแสดง ต่างจากแบบ Abstractive ที่ยังให้ผลที่ดีกว่าแบบจำลองอื่น ๆ และดีกว่าแบบจำลอง Oracle ซึ่งเป็นขอบเขตบนของงานนี้อีกด้วย

จากงานวิจัยของ Yang Liu ในปีค.ศ. 2019 [2] เน้นไปที่การทำสรุปข้อความแบบ Extractive โดยอาศัย BERT กับชุดข้อมูล CNN/Dailymail และ New York Times (NYT) โดยในการทดลองนั้นจะเป็นการเพิ่มชั้น (layer) ให้กับ BERT มี 3 แบบ หลัก ๆ 1.Classifier โดยการใช้ Sigmoid Function ในชั้น BERT Output 2.Inter-sentence Transformer เป็นการเพิ่มชั้นของ Transformer ขึ้นมาอีกชั้นในระดับประโยค และยังใช้ Sigmoid Function ในชั้น BERT Output เหมือนเดิม และ 3.Recurrent Neural Network (RNN) โดยทำการใช้ RNN แบบ LSTM(Long Short Term Memory) โดยจะใช้ input มาจาก BERT Output แล้วจึงให้ผลลัพธ์สุดท้ายด้วย Sigmoid Function เหมือนเดิม

จากงานวิจัยของ Derek Miller ในปีค.ศ. 2019 [3] นี้จะเน้นการทำสรุปข้อความแบบ Extractive และใช้ข้อมูลที่เป็นข้อความประเภทหนังสือเรียน ขั้นตอนวิธีที่ใช้ในการสรุปข้อความคือการใช้ Clustering โดยใช้กับ Word Embedding ของ BERT แล้วจึงนำไปเทียบกับการใช้ TextRank ซึ่งเป็นวิธีการสรุปข้อความแบบ Extractive อีกแบบ โดยทำการประเมินผลด้วยการเปรียบเทียบผลลัพธ์ของการสรุปทั้งสองแบบโดยมนุษย์ ผลพบว่าการทำ Clustering จาก Word Embedding ของ BERT นั้นให้ผลที่ดีกว่าแต่ก็ยังประสบปัญหาที่เหมือนกับตัวแบบจำลองอื่น ๆ ในกรณีที่ข้อความยาวเกินไป

ในโครงการนี้ ผู้พัฒนาจะยึดจากงานวิจัยของ Derek Miller ในการทำสรุปบทความ

คลังโปรแกรมที่เกี่ยวข้อง

- **BeautifulSoup** คือคลังโปรแกรมที่ใช้ในการดึงข้อมูลจากหน้าเว็บไซต์ (Web Scraping) ในส่วนที่เราต้องการนำมาใช้งาน
- **Pythainlp** คือคลังโปรแกรมภาษาของ Python ที่ใช้สำหรับการประมวลผลข้อความ การวิเคราะห์ทางภาษา ถูกพัฒนาขึ้นเพื่อทำงานร่วมกับภาษาไทยโดยเฉพาะ
- **TLTK** คือคลังโปรแกรมภาษาของ Python ที่ใช้สำหรับการตัดคำ จดจำเอนทิตี ตรวจสอบความถูกต้อง และมีการเก็บคำลงไปคลังเก็บข้อมูล ถูกพัฒนาขึ้นที่จุฬาลงกรณ์มหาวิทยาลัย
- **Gensim** คือคลังโปรแกรมที่ใช้สำหรับการสร้างแบบจำลอง และการประมวลผลภาษาธรรมชาติ (Natural Language Processibg) โดยใช้หลักการเรียนรู้ของเครื่องที่ทันสมัย ประกอบด้วยอัลกอริทึมที่หลากหลาย ถูกออกแบบมาเพื่อจัดการการเก็บข้อความจำนวนมาก

ภาพรวมของโครงการ

แหล่งที่มาของข้อมูล

ข้อมูลที่นำมาใช้คือเนื้อหาข่าวที่ได้มาจาก www.thepotential.org

การดึงข้อมูล และจัดเก็บข้อมูล

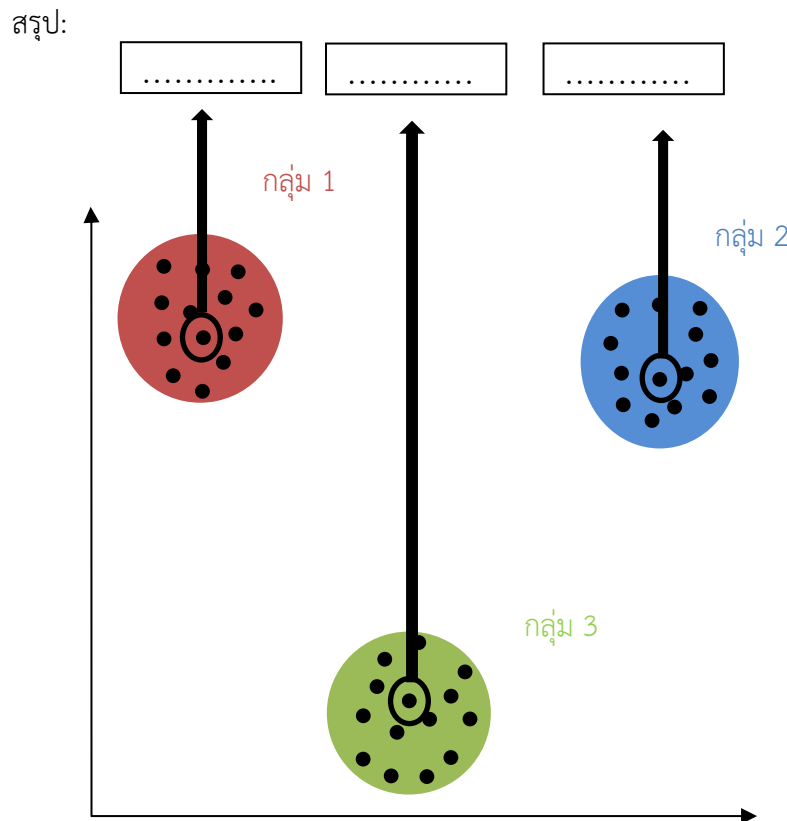
ดึงข้อมูลจาก www.thepotential.org ใช้ Python Library ที่มีชื่อว่า BeautifulSoup ร่วมกันกับ Google Colaboratory จากนั้นจึงบันทึกข้อมูลที่ได้รับมาในรูปแบบไฟล์เอกสาร และจัดเก็บลงไปใน Google Drive เพื่อความสะดวกในการจัดเก็บข้อมูล

การพัฒนาแบบจำลอง

มี 3 ขั้นตอนหลัก ๆ ดังนี้

1. **Word Embedding** เลือกวิธีในการทำ word embedding ที่เหมาะสมจากคลังข้อมูล Pythainlp
2. **Sentences Embedding** โดยอาจจะใช้วิธี BERT เพื่อหา Sentence Embedding โดยตรง หรือนำ word embedding มาทำ Sentences Embedding โดยใช้ Aggregate Function ประเภทใดประเภทหนึ่งมาใช้

3. **Clustering** นำ Sentence Embedding ที่ได้จากข้อ 2 มาจัดกลุ่ม และเลือกประโยคที่ใกล้จุด Centroid ของแต่ละกลุ่มออกมา เพื่อนำมาต่อกันจนเป็นสรุป (รูปที่ 6)



รูปที่ 6: การทำ Clustering

การทดสอบประสิทธิภาพของแบบจำลอง

1. นำการสรุปที่ได้จากเว็บไซต์ (Gold Summary) มาเปรียบเทียบกับสรุปที่ได้จากโมเดลโดยตรง
2. ใช้สรุปที่ได้จากเว็บไซต์ (Gold Summary) มาเลือกประโยคในบทความเพื่อนำไปเปรียบเทียบกับสรุปที่ได้จากโมเดล
3. ให้นักช้เลือกประโยคในบทความเพื่อนำมาเป็นสรุป และนำมาเปรียบเทียบกับสรุปที่ได้จากโมเดล

โดยในแต่ละวิธี เราจะหาค่า Precision และค่า Recall เพื่อใช้ในการประเมินประสิทธิภาพของโมเดลสรุปข่าวที่สร้างขึ้น

ขอบเขตของโครงการ

1. ข้อความที่ใช้ในการทำสรุปเป็นข้อความประเภทบทความ (Article) จากเว็บไซต์ www.thepotential.org
2. การสรุปบทความจะเป็นการสรุปแบบ Extractive Summarization เท่านั้น
3. ข้อความที่ใช้ในการทำสรุปต้องเป็นภาษาไทยเท่านั้น
4. ข้อความที่ได้จากการสรุปบทความจะเป็นภาษาไทย

วิธีการดำเนินงาน

1. ศึกษาเนื้อหาและบทความเกี่ยวกับการสรุปข้อความ
2. ศึกษาเนื้อหาและบทความเกี่ยวกับแบบจำลองต่าง ๆ ที่ใช้ในการทำสรุปข้อความ
3. เก็บข้อมูลจากหน้าเว็บไซต์ www.thepotential.org
4. ทำการพัฒนาแบบจำลอง
5. ทดสอบวัดประสิทธิภาพของแบบจำลองที่ได้
6. ปรับปรุงแบบจำลองเพื่อเพิ่มประสิทธิภาพ
7. สรุปผลการดำเนินงาน ข้อเสนอแนะ และการจัดทำเอกสาร

ตารางการดำเนินงาน

| ขั้นตอนการดำเนินงาน | ปีการศึกษา 2563 | | | | | | | | |
|---|-----------------|------|------|------|------|------|------|-------|-------|
| | ส.ค. | ก.ย. | ต.ค. | พ.ย. | ธ.ค. | ม.ค. | ก.พ. | มี.ค. | เม.ย. |
| 1. ศึกษาเนื้อหาและบทความเกี่ยวกับการสรุปข้อความ | | | | | | | | | |
| 2. ศึกษาเนื้อหาและบทความเกี่ยวกับแบบจำลองต่าง ๆ ที่ใช้ในการทำสรุปข้อความ | | | | | | | | | |
| 3. เก็บข้อมูลจากหน้าเว็บไซต์ www.thepotential.org | | | | | | | | | |
| 4. ทำการพัฒนาแบบจำลอง | | | | | | | | | |
| 5. ทดสอบวัดประสิทธิภาพของแบบจำลองที่ได้ | | | | | | | | | |
| 6. ปรับปรุงแบบจำลองเพื่อเพิ่มประสิทธิภาพ | | | | | | | | | |
| 7. สรุปผลการดำเนินงาน ข้อเสนอแนะ และการจัดทำเอกสาร | | | | | | | | | |

ประโยชน์ที่คาดว่าจะได้รับ

1. ประโยชน์ในด้านความรู้ และประสบการณ์ต่อตัวนิสิตเอง
 - 1.1 ได้รับความรู้เกี่ยวกับหลักการเรียนรู้ของเครื่อง (Machine Learning)
 - 1.2 ได้รับความรู้เกี่ยวกับการประมวลผลภาษาธรรมชาติ (Natural Language Processing)
2. ประโยชน์ในด้านการแก้ไขปัญหาของสังคมหรือสภาพแวดล้อม
 - 2.1 สามารถทำการสรุปบทความได้โดยอัตโนมัติ
 - 2.2 เป็นตัวช่วยสำหรับผู้เขียนข่าว

อุปกรณ์และเครื่องมือที่ใช้

1. ฮาร์ดแวร์(Hardware)
 - 1.1 Lenovo thinkpad notebook L440
 - 1.2 ASUS TUF FX505DV
2. ซอฟต์แวร์(Software)
 - 2.1 Google Colaboratory
 - 2.2 Google Drive
 - 2.3 Google Document
 - 2.4 Python 3.6
 - 2.5 Machine Learning Libraries Keras

งบประมาณ

| | |
|---------------------------------------|----------------|
| 1. Cloud Service เป็นระยะเวลา 9 เดือน | 7,500 บาท |
| 2. RAM DDR3L BUS1600 8GB จำนวน 2 ชิ้น | 2,500 บาท |
| | รวม 10,000 บาท |

เอกสารอ้างอิง

[1] Liu, Yang, and Mirella Lapata. "Text Summarization with Pretrained Encoders." *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*.

Available from : <https://arxiv.org/pdf/1908.08345.pdf> [2019]

[2] Liu, Yang. "Fine-tune BERT for extractive summarization." *arXiv preprint arXiv:1903.10318*

Available from : <https://arxiv.org/pdf/1903.10318.pdf> [2019]

[3] Derek Miller. Leveraging BERT for Extractive Text Summarization on Lectures.

Available from : <https://arxiv.org/ftp/arxiv/papers/1906/1906.04165.pdf> [2019, June 7]

[4] Peeradech Bangchareons. A Machine-Translation based Approach to Word Boundary Identification: A Projective Analogy of Bilingual Translation. [2010]

[5] Keng Surapong. Python ตัดคำภาษาไทยด้วย PyThaiNLP.

Available from : <https://www.bualabs.com/archives/3740/python-word-tokenize-pythainlp-example-algorithm-deepcut-newmm-longest-python-pythainlp-ep-2/> [2020, January 29]

[6] Pattarawat Chormai. AttaCut: : A Fast and Accurate Neural Thai Word Segmenter.

Available from : <https://arxiv.org/ftp/arxiv/papers/1911/1911.07056.pdf> [2019, August 26]

[7] Wirote Aroonmanakun. TLTK: Thai Language Toolkit.

Available from : <https://pypi.org/project/tltk> [2020, July 30]

[8] Y. Poowarawan. Dictionary-based Thai Syllable Separation

Available from : <https://www.semanticscholar.org/paper/Dictionary-based-Thai-Syllable-Separation-Poowarawan/d15dca32d2bd685c12fb9a5cdb5bbdf6ccc366b2> [1986]

[9] Choochart Haruechaiyasak, Sarawoot Kongyoung and Matthew N. Dailey.

A Comparative Study on Thai Word Segmentation Approaches. [2008, June.]

[10] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems*. [2017]

ภาคผนวก ข

ตัวอย่างบทความ และสรุปจากแบบจำลอง

ตัวอย่างที่ 1

ตัวอย่างบทความ

การแพร่ระบาดของไวรัสโคโรนาสายพันธุ์ใหม่ (โควิด-19) ก่อให้เกิดผลกระทบต่อวิถีชีวิตดั้งเดิมเป็นวงกว้าง หากพิจารณาในภาพย่อยๆ ระดับครอบครัว พ่อแม่จำนวนไม่น้อยกลายเป็นผู้ประสพภัยกันถ้วนหน้า ทั้งมิติการทำงาน เศรษฐกิจ วิถีชีวิต และสำคัญที่สุดคือการศึกษาของลูกซึ่งพ่อแม่อาจต้องเป็นผู้รับผิดชอบเป็นหลักไปก่อน

แม้นโยบายล่าสุด กระทรวงศึกษาธิการจะประกาศว่าให้เลื่อนเปิดภาคเรียนของโรงเรียนทั่วประเทศออกไปเป็นวันที่ 1 กรกฎาคม 2563 แต่ในความเป็นจริงแล้ว ไม่อาจยืนยันได้ด้วยซ้ำว่าเมื่อถึงเวลานั้น การเรียนการสอนจะเป็นไปได้ตามปกติจริงไหม เพราะการเรียนออนไลน์มีองค์ประกอบและปัจจัยสนับสนุนหลายประการเหลือเกิน ยังไม่นับความเหลื่อมล้ำทางสังคมที่ทำให้ปัญหาภัยทวีความรุนแรงมากขึ้นไปอีก และคนที่ได้รับผลกระทบหนักหนากที่สุดก็คือครอบครัวและพ่อแม่นั่นเอง

The Potential ได้มีโอกาสพูดคุยกับคุณพ่อคุณแม่และเด็กที่มีประสบการณ์การจัดการเรียนรู้ด้วยตัวเอง(บ้านเรียน)มาก่อน ได้แก่ พ่อปุ๊-วิวัฒน์ กังวานนวกุล, แม่เซ็ง-ศุภรัตน์ ศุภชัยศิริกุล, แม่จี๊-ธัญชนก สีหาพล, ป้อมปิ่น-วรวิศ สบายใจ และถอดบทเรียนมาเป็นแนวทางสำหรับพ่อแม่ว่าเราอาจมีตัวช่วยในสถานการณ์ “Learn From Home” นี้ได้อย่างไรบ้าง

การศึกษาในความคุ้นเคยของพ่อแม่อาจหมายถึงการส่งลูกไปโรงเรียน เรียนเป็นวิชา วัตถุประสงค์จากเกณฑ์การประเมินที่มี แต่อันที่จริงแล้วการศึกษาเป็นเพียงส่วนย่อยที่เล็กมากของคำว่า “การเรียนรู้” ในวิกฤติครั้งนี้ ก่อให้เกิดการล่มสลายของหลายอย่าง แต่สิ่งหนึ่งที่ช่วยให้คนรอดได้คือการมีทักษะชีวิต บางคนตกงาน แต่ยังอยู่รอดได้เพราะทำอาหารได้ ค้าขายออนไลน์เก่ง นำเอางานอดิเรกมาทำเป็นอาชีพเป็นรายได้หลักในช่วงเวลานี้ คนจำนวนมาก ใช้องกาศนี้สร้างแ

หลังอาหารประจำบ้านของตัวเอง ด้วยการปลูกผักสวนครัว รื้อพื้นที่ทักษะการถนอมอาหาร แปรรูปอาหาร เป็นความรู้ที่ช่วยชีวิตได้จริง

ดังนั้นพ่อแม่อาจต้องมีมุมมองใหม่ว่า “การเรียนรู้ไม่ได้จำกัดแค่เพียงการไปโรงเรียนเท่านั้น” และเกิดขึ้นได้จากทุกสิ่งที่ลูกสนใจ

นายแพทย์ประเสริฐ ผลิตผลการพิมพ์ กล่าวว่า เด็กเรียนรู้ในทุกวันอยู่แล้วนับตั้งแต่เกิดโดยมีแม่หรือผู้เลี้ยงดูหลักเป็นของเล่นที่ดีที่สุด เป้าหมายการเรียนรู้ของเด็กเล็กในช่วง 0-7 ปี คือ “การเล่น” เล่นดิน เล่นทราย สืบราชบรรพชาติ ต่อบล็อกไม้ ให้เด็กได้ใช้กล้ามเนื้อขาและนิ้วมือมากๆ และ “การฟังนิทาน” ทุกวัน เพื่อให้เด็กมีความมั่นคงภายใน(แม่มีอยู่จริง)และส่งเสริมพัฒนาการสมองทางอ้อมด้วย

คุณหมอยืนยันว่าเด็กเล็กช่วงปฐมวัยจึงไม่ต้องไปโรงเรียน(ในวิกฤติ)ก็ได้ สามารถเรียนรู้จากการเล่น ฟังนิทาน และการทำงานบ้านง่ายๆได้อย่างไม่มีปัญหาใดเลย

แม่จิบ-ธัญชนก สีหาพล เป็นคุณแม่ของลูกวัย 5 ปี และ 3 ปี เป็นเจ้าของเพจเลี้ยงลูกให้โตไปกับหนังสือ แบ่งปันการดูแลลูกผ่านการเล่นและอาชีพของพ่อแม่ว่า

“ที่บ้านคุณพ่อปลูกผัก เราจึงปล่อยให้ลูกเล่นดินเล่นทรายเต็มที่ ชวนลูกคุ้ยถ้าปลูกผักต้องใช้ อุปกรณ์อะไร ต้องรดน้ำแบบไหน ผักหน้าตาเป็นอย่างไร เขาอาจแค้ค้ยเขี่ยดิน แต่ได้ฝึกการใช้กล้ามเนื้อ ได้อยู่กับธรรมชาติ เป็นโอกาสที่พ่อแม่ได้พักด้วย เด็กอยู่นิ่งไม่ได้ ต้องหากิจกรรมเล่นให้เขาได้ใช้พลังและไม่เบื่อ แค้ได้ไหนตัวกับต้นไม้ บางทีเขาก็สนุกมากแล้ว ที่บ้านเราอ่านนิทานก่อนนอนทุกคืน แล้วชวนลูกคิดว่าพรุ่งนี้อยากทำอะไร เช่น ทำแคมป์ไฟ เล่นทราย ทำอาหาร ตื่นเข้ามาคนโตจะวาดรูปนำเสนอความคิดว่าที่เขาอยากได้คือแบบนี้ มีอะไรบ้าง แม่จะเตรียมของให้ เช่น ปิ้งลูกชิ้น มาชมเมลโล ชวนลูกคุยว่าได้เรียนรู้ อะไรบ้าง หรือมีสิ่งไหนที่เราอยากทำให้ดีขึ้น เราขายหนังสือเด็กมือสองมาได้สองสามปีแล้วเพราะทำห้องสมุดที่บ้าน ลูกๆ มีโอกาสไปช่วยขายหนังสือด้วย เขาได้พบบรรยากาศที่เด็กๆ มาอ่านหนังสือ คนมาซื้อของ ได้ชิมซบสิ่งเหล่านี้ จนวันหนึ่งเขาก็พูดว่า ‘โตขึ้นภูจะช่วยแม่ขายหนังสือนะ’ ทุกอย่างที่เราทำจึงช่วยการเรียนรู้ของลูกเสมอ”

เป้าหมายการเรียนรู้ของเด็กระดับประถมศึกษา (7-12ปี) คือการฝึกฝนความตื่นตัว ตั้งใจ คัดกรอง คัดเลือกข้อมูลเพื่อลงมือทำให้สำเร็จ ผู้ปกครองทำหน้าที่เพียงออกแบบโจทย์ปัญหาพร้อมกับลูกเพื่อนำไปคิดวิเคราะห์ และลงมือทำเพื่อแก้ไขปัญหาให้ได้ตามเป้าหมาย เด็กที่เริ่มโตขึ้นจำเป็นต้องทำตามหน้าที่ ขณะเดียวกันก็เริ่มต้องการพื้นที่และเวลาของตัวเอง

ป้อมปิ่น-วรรษวิศ สบายใจ อายุ 27 ปี เป็นผู้ที่มิประสบความสำเร็จการเรียนรู้แบบบ้านเรียนมาก่อน และทุกวันนี้ยังคงเรียนรู้และทำงานเกี่ยวข้องการกับวิเคราะห์และพัฒนาการเรียนรู้อยู่เสมอ เขาแบ่งปันแนวทางการจัดสรรเวลาสำหรับพ่อแม่ในช่วงการเรียนรู้ที่บ้านว่า

“การกลับมาเรียนรู้ที่บ้าน(สำหรับคนที่ไม่เคยทำบ้านเรียน) เป็นไปได้ว่าเราอาจต้องนิยามความหมายของ Space&Time คำว่า ‘บ้าน’ เสียใหม่ ต้องยกระดับขึ้นเป็น ‘พื้นที่เรียนรู้’ ได้ด้วยแนวคิดการออกแบบ 3 กิจกรรม 3 พื้นที่ คือ การทำกิจกรรมในบ้าน (งานบ้าน) กิจกรรมนอกบ้าน (การเชื่อมโยงการเรียนรู้กับปรากฏการณ์ทางสังคม) และกิจกรรมตามความสนใจส่วนตัว จะช่วยให้การอยู่ร่วมกันของคนในบ้านผ่อนคลายลง ทุกคนมีโอกาสเป็นเจ้าของเวลาของตัวเอง ในพื้นที่ของตัวเองได้ด้วย”

“ไม่อยากให้พ่อแม่คิดว่าลูกต้องเรียนรู้ตลอดเวลา การปล่อยให้ได้เรียนรู้ด้วยตัวเองหรือทำสิ่งที่สนใจด้วยความสนุก ความอยากรู้ ก็ช่วยเพิ่มทักษะได้เหมือนกัน การออกแบบการศึกษาที่บ้านเราสามารถใช้ทรัพยากรที่เรามีอยู่อย่างจำกัดในบ้าน สร้างโอกาสใหม่ๆ ให้สนุกได้”

1. ออกแบบตารางเวลาร่วม

เมื่อพ่อแม่และลูกจำเป็นต้องใช้เวลาอยู่ร่วมกันมากขึ้น ในขณะที่พ่อแม่หลายคนจำเป็นต้องทำงานที่บ้าน การออกแบบตารางเวลาร่วมกันเป็นเครื่องมือช่วยให้สมาชิกในบ้านเห็นภาพรวมที่ร่วมกันมากขึ้น โดยระบุเวลาที่ตายตัวของแต่ละคน เช่น พ่อแม่ต้องทำงาน ลูกต้องเข้าชั้นเรียน แล้วมาบริหารเวลาว่างที่เหลือที่มาร่วมกันว่าสามารถทำอะไรในช่วงไหนได้บ้าง หรือ บริหารเวลาระหว่างพอกับแม่ว่าเวลาไหน ใครต้องทำอะไร พ่อแม่ที่มีลูกเล็กอาจต้องแบ่งกันดูแลลูก หรือนั่งเรียนออนไลน์พร้อมกับลูก เป็นต้น

2. ออกแบบกิจกรรมร่วม

1) ใช้งานบ้านและวิถีชีวิตเป็นการเรียนรู้ในบ้าน โดยแบ่งกิจกรรมงานบ้านมอบหมายให้ทุกคนร่วมกันทำ ครอบครัวของ แม่เซ็ง-ศุภรัตน์ ศุภชัยศิริกุล ใช้งานบ้านช่วยออกแบบการใช้ชีวิตร่วมกันกับลูกสาวทั้งสามคน วัย 13, 18 และ 20 ปี

“ลูกทุกคนมีหน้าที่ทำงานบ้านโดยเลือกจากที่ชอบ อย่างพิกซี่ (ลูกคนเล็ก) ชอบเรื่องสัตว์ จึงมอบหมายให้มีหน้าที่ดูแลสัตว์ทั้งหมดในบ้าน แต่ไม่ชอบซักผ้าตากผ้า แพรี่ (ลูกคนที่สอง) ก็เป็นคนทำ คนน้องจะรู้หน้าที่ตัวเอง ตื่นมาให้อาหารไก่ ปล่อยสัตว์ไปเดิน ปีนี้เขาได้รับมอบหมายโปรเจกต์ที่ซับซ้อนขึ้น โดยมีหน้าที่สำรวจธรรมชาติทุกอย่างในรั้วบ้าน เช่น นกอพยพ แมลง สัตว์เลื้อยคลานต่างๆ แล้วตรวจสอบข้อมูล นำถ่ายภาพมาวิเคราะห์เทียบเคียง ตรวจสอบอ้างอิงกับกลุ่มอนุรักษ์ความหลากหลายทางชีวภาพ และสิ่งแวดล้อม ถ้ายังไม่ได้คำตอบให้สอบถามผู้เชี่ยวชาญ เป็นต้น

“งานบ้านเป็นกิจกรรมที่เชื่อมโยงการเรียนรู้ได้เยอะมาก ทั้งศิลปะ วิทยาศาสตร์ การทดลอง เราใช้เวลาช่วงกินข้าวโยนคำถามให้กัน เช่น กินแกงเทโพ แล้วสงสัยว่าผักบุงมีกี่แบบ ที่ใส่เย็นตาโฟ กินกับส้มตำที่ทำผัดผักบุงต่างกันไหม แค่นี้ก็เป็นคำถามให้ค้นคว้าได้แล้ว สายพันธุ์ผักบุงมีอะไรบ้าง อะไรคือพันธุ์พื้นบ้าน อะไรคือพันธุ์ต่างถิ่น แล้วเต้าเจี้ยวที่ใส่ผัดผักบุง ทำยังไง ใช้อะไรหมัก ถ้าพ่อแม่เคารพลูกนะ เราจะเห็นเลยว่าลูกโดยเฉพาะเด็กโตมีความสามารถในการหาข้อมูลที่ลึกมากเลย บางทีเค้าตีข้อมูลเรากระจายเลย (หัวเราะ) นี่แหละคือการเรียนรู้ที่เราทำให้มันเกิดขึ้น ไม่ใช่แค่การเรียนแบบที่เราคุ้นชินเป็นวิชาๆ”

2) ใช้ปรากฏการณ์ทางสังคมเชื่อมโยงกับการเรียนรู้ เช่น ตั้งคำถามว่าเราจะช่วยผู้อื่นในสถานการณ์โรคระบาดนี้ได้อย่างไร อาจชวนกันแบ่งปันอาหาร สิ่งของให้ผู้ที่ได้รับผลกระทบ การออกแบบวิถีชีวิตใหม่หลังผ่านพ้นสถานการณ์โควิด เพื่อฝึกฝนการคิดวิเคราะห์ ฝึกการเชื่อมโยงตัวเองเข้ากับสังคม ฝึกความเห็นอก

เห็นใจเพื่อนมนุษย์ ซึ่งเป็นสิ่งสำคัญที่สุดในยุค AI ทำงานแทนคนได้มากขึ้น บางครอบครัวใช้สถานการณ์ไฟไหม้ป่าภาคเหนือเป็นจุดเริ่มต้นในการสอนลูกเรื่องการแบ่งปัน การทำเพื่อผู้อื่น เช่น การร่วมบริจาคสิ่งของ ให้เจ้าหน้าที่ดับไฟป่า การช่วยทำแนวกันไฟ และต่อเนื่องถึงการช่วยเหลือดูแลสัตว์ป่าที่ได้รับผลกระทบจากไฟป่า เป็นต้น

3. มีพื้นที่และเวลาส่วนตัวของแต่ละคนด้วย

การอยู่ร่วมกันไม่จำเป็นต้องทำทุกอย่างด้วยกันไปทั้งหมด ควรมีเวลาอิสระสำหรับแต่ละคนได้ด้วยเพื่อผ่อนคลายและทำสิ่งที่สนใจตามความต้องการของตัวเอง เด็กๆ อาจใช้เวลาว่างของเขาเล่นเกม ดูทีวี พ่อแม่ก็ต้องเคารพช่วงเวลานี้ของเขาด้วย สิ่งสำคัญที่สมาชิกทุกคนจะได้ออกมาจากการจัดการเวลาร่วมกัน คือการสร้างจังหวะชีวิตที่สอดคล้องกัน มีทั้งพื้นที่ส่วนรวม พื้นที่ส่วนตัว และเรียนรู้การเคารพเวลาและสิทธิของกันและกันด้วย

วิธีการที่ง่ายที่สุดสำหรับพ่อแม่คือการสำรวจทรัพยากรในบ้านและชุมชนแล้วนำมาสร้างการมีส่วนร่วมเรียนรู้และสร้างโจทย์ปัญหาอย่างง่าย ๆ เพื่อเล่นสนุกไปด้วยกันกับลูก

พ่อปุ๋-วีรวุฒิ กังวานนวกุล ผู้จัดการโรงเรียน พิพิธภัณฑ์ของเล่นเพื่อการเรียนรู้ จ.เชียงราย ใช้บรรยากาศชุมชน พ่อครูแม่ครูเป็นต้นทุนให้ลูกชายสองคนเรียนรู้ในสิ่งที่สนใจ ในกรณีของลูกชายคนโต เขาได้เติบโตจากเด็กที่เล่นของเล่นที่บ้าน มาสู่การรื้อซ่อมของเล่น ออกแบบของเล่นด้วยตัวเอง จนปัจจุบันได้พัฒนาความซับซ้อนกลายเป็นการออกแบบรถไฟสามล้อที่เคลื่อนที่ไปได้ในชีวิตจริงด้วยการลงมือทำเองทั้งหมด ส่วนลูกคนเล็กก็พัฒนาสินค้าเป็นของเล่นและงานฝีมือจากผ้าในแบรนด์ปักด้ายปักดิ้นร่วมกับคุณแม่ โดยยกระดับสู่การขายออนไลน์ร่วมกันทั้งครอบครัว (The Potential จะนำเสนอบทความในรายละเอียดต่อไป)

“การเริ่มต้นจากทรัพยากรที่มีในบ้าน ทำให้เราไม่ต้องลงทุนใหม่ ลองดูว่าในชั้นหนังสือ ห้องครัว ในสวน พื้นที่รอบชุมชน เราน่าอะไรเป็นเครื่องมือการเรียนรู้ได้บ้าง เช่น เราชวนลูกคุยว่าถ้าในหนึ่งสัปดาห์นี้ เราจะไม่ออกไปตลาดเลย เราควรจะต้องมีวัตถุดิบอาหารอะไรบ้าง หรือถ้าหนึ่งเดือนเราจะไม่ออกไปตลาดเลย เราต้องปลูกผักอะไรไว้บ้าง มีอะไรที่เราจะสร้างเป็นอาหารได้บ้าง ก็เกิดเป็นรูปแบบกิจกรรมอย่างง่าย ๆ ขึ้น ทดลองกันออกแบบเมนูอาหาร เช้ากลางวันเย็น น่าจะมีเมนูอะไรบ้าง

“การที่เราชวนลูกทำเรื่องอาหาร ไม่ได้แค่อาหารที่กินแล้วอิม้อย่างเดียว แต่ได้เรื่องปฏิสัมพันธ์ การแลกเปลี่ยนสนทนากับลูก อย่างเช่นเมนูกะเพรา เด็กที่ไม่เคยลงมือทำมาก่อนอาจจะสงสัยว่าทำไมต้องใส่กระเทียม พริก แล้วพริกมีกี่แบบ ไฟอ่อนหรือแรงได้ผลต่างกันไหม ทำไมต้องใส่ใบกะเพรา ถ้าเปลี่ยนเป็นใบอย่างอื่นได้หรือเปล่า เครื่องปรุงต้องใช้อะไร ทุกอย่างเป็นทั้งคำถามและการเรียนรู้การทำอาหารหรือการ

เรียนรู้ร่วมกับลูกในสถานการณ์แบบนี้ เราไม่ได้หวังผลระยะสั้นทันที แต่เป็นจุดเริ่มต้นของการปรับแนวคิดของทั้งพ่อแม่และเด็กๆว่า การเรียนรู้เกิดขึ้นทุกที่ทุกเวลา”

ทรัพยากรในที่นี้ยังอาจหมายถึงความสนใจของเด็ก อุปกรณ์ เครื่องมือที่พ่อแม่จะสนับสนุนให้ได้ เช่น หัดถ่ายภาพ หัดตัดต่อวิดีโอ การลองถ่ายสารคดี การวาดสติ๊กเกอร์ไลน์ หรือความเป็นไปได้ทุกอย่างที่เด็กอยากรู้อยากเห็นที่จะลงมือทำด้วยตัวเอง

ความสนุกเป็นสิ่งสำคัญสำหรับเด็ก การทดลองริเริ่มทำบางอย่าง ไม่จำเป็นต้องได้ผลสัมฤทธิ์ที่ดีที่สุดเสมอไป เพราะในระหว่างการทำพลาดนั้น เด็กๆ ก็เกิดการเรียนรู้ด้วยเหมือนกัน พ่อแม่อาจใช้โอกาสของการเรียนที่บ้านเป็นสนามเด็กเล่นหรือเป็นห้องทดลองขนาดใหญ่ที่เราสามารถใช้ความอยากรู้อยากเห็นมาออกแบบสู่การแสวงหาคำตอบอย่างเป็นขั้นเป็นตอนได้เหมือนกับการทดลองทางวิทยาศาสตร์ ทุกๆ ขั้นตอนของการลองผิดลองถูก เราถือว่าเป็นการเรียนรู้ทั้งสิ้น

แม่เชิง-ศุภรัตน์ ศุภชัยศิริกุล มองว่าวิกฤตินี้ทำให้ทุกคนตั้งเครียดอยู่แล้ว พ่อแม่ต้องลดความคาดหวังลง แล้วหันมายืนอยู่ข้างเดียวกับลูกเพื่อสามารถเปลี่ยนผ่านไปด้วยกัน

“อยากให้พ่อแม่มองว่าเรามีโอกาสเรียนรู้จากทุกๆ อย่าง ถ้าไม่มองเรื่องความสำเร็จ ทุกสิ่งที่เราทำพลาด เราเรียนรู้และพัฒนาอะไรจากมันได้ นี่มีความหมายมากแล้ว เพราะฉะนั้น พ่อแม่ไม่ต้องเก่งที่สุด ดีที่สุด แต่เราก็จะเรียนรู้ไปด้วยกันกับลูกได้ทุกครั้ง”

วิถีชีวิตที่ต้องทั้งทำงานจากบ้านและเลี้ยงลูก (ให้เรียน) ที่บ้าน อาจไม่ใช่เรื่องง่ายสำหรับหลายคน จำเป็นต้องอาศัยการปรับตัวกันทุกคน ข้อนี้เชื่อมโยงกับข้อก่อนหน้า ถ้าเราลดความคาดหวังไปบ้าง ทั้งจากตัวเองในฐานะพ่อแม่และจากลูก ชีวิตจะโปร่งเบาไปได้มาก ให้อยอมรับและเข้าใจว่าในการเริ่มต้นใหม่ ทุกเรื่องต้องใช้เวลาเรียนรู้ด้วยกันทั้งสิ้น

แม่จีบ-ธัญชนก สีหาพล แบ่งปันมุมมองของการเป็นคุณแม่เด็กปฐมวัยว่า “บางทีพ่อแม่แบกคำว่า การเรียนรู้มากเกินไป บางครั้งเราเห็นลูกทำด้วยความสนุกด้วยตัวเอง ไม่แบก ไม่ต้องมาคิดว่าจะเรียนรู้อะไร เขาจึงมีความสุข เราต้องเรียนรู้สิ่งเหล่านี้จากลูกด้วย”

เช่นเดียวกับที่ พ่อปุ้-วิรวุฒน์ กังวาลนวกุล ส่งพลังใจกำลังใจแก่พ่อแม่ทุกคนว่า “ทุกสิ่งทุกอย่างเป็นการเรียนรู้ได้ทั้งนั้น แม้แต่วิถีชีวิตของพ่อแม่ สิ่งนั้นสอดคล้องกับการศึกษาใหม่ เราไม่รู้ แต่ที่แน่ๆ คือทำให้สอดคล้องกับวิถีชีวิต ความสุขจึงจะเกิดขึ้นได้จริง ถ้าปล่อยให้ชีวิตรวน ต่อให้ลูกได้การศึกษาที่ดีแค่ไหนมันกลายเป็นปัญหาอยู่ดี ทำให้สมดุล ยืดหยุ่นและเหมาะสมในแบบของเราเอง เราจะมีความสุขในแบบของเราเอง ขอให้เริ่มด้วยความสนุก ถ้าเริ่มต้นจากสิ่งนี้ได้ ทุกอย่างจะไปต่อได้ง่ายมาก”

หากเหนื่อยหนักหนักหนากว่านั้น นายแพทย์ประเสริฐ ผลิตผลการพิมพ์ ได้เขียนไว้ในเพจเฟซบุ๊กของท่านเพื่อสื่อสารกับพ่อแม่อย่างเข้าใจต่อสถานการณ์การเปลี่ยนผ่านคราวนี้ว่า ได้การศึกษาที่ดีแค่ไหนมันกลายเป็นปัญหาอยู่ดี ทำให้สมดุล ยืดหยุ่นและเหมาะสมในแบบของเราเอง เราจะมีความสุขในแบบของเราเอง ขอให้เริ่มด้วยความสนุก ถ้าเริ่มต้นจากสิ่งนี้ได้ ทุกอย่างจะไปต่อได้ง่ายมาก”

ตัวอย่างสรุปจากแบบจำลอง

การปล่อยให้ได้เรียนรู้ด้วยตัวเองหรือทำสิ่งที่สนใจด้วยความสนุก ความอยากรู้ ก็ช่วยเพิ่มทักษะได้เหมือนกัน อาจไม่ใช่เรื่องง่ายสำหรับหลายคน จำเป็นต้องอาศัยการปรับตัวกันทุกคน ส่วนลูกคนเล็กก็พัฒนาสินค้าเป็นของเล่นและงานฝีมือจากผ้าในแบรนด์ปักด้ายปักดิ้ร่วมกับคุณแม่ โดยยกระดับสู่การขายออนไลน์ร่วมกันทั้งครอบครัว (The Potential จะนำเสนอบทความในรายละเอียดต่อไป)

“การเริ่มต้นจากทรัพยากรที่มีในบ้าน เสียใหม่ คนจำนวนมาก สายพันธุ์ผักบุงมีอะไรบ้าง อะไรคือพันธุ์พื้นบ้าน นี่มีความหมายมากแล้ว จ.เชียงราย ใครต้องทำอะไร ด้วยการปลูกผักสวนครัว รื้อฟื้นทักษะการถนอมอาหาร แปรรูปอาหาร เป็นความรู้ที่ช่วยชีวิตได้จริง

ดังนั้นพ่อแม่อาจต้องมีมุมมองใหม่ว่า “การเรียนรู้ไม่ได้จำกัดแค่เพียงการไปโรงเรียนเท่านั้น” และเกิดขึ้นได้จากทุกสิ่งที่ลูกสนใจ

นายแพทย์ประเสริฐ เด็กอยู่ไม่ได้ เขาจึงมีความสุข เราไม่รู้ หากพิจารณาในภาพย่อยๆ ระดับครอบครัว ถ้าพ่อแม่เคารพลูกนะ เขาแบ่งปันแนวทางการจัดสรรเวลาสำหรับพ่อแม่ในช่วงการเรียนรู้ที่บ้านว่า

“การกลับมาเรียนรู้ที่บ้าน(สำหรับคนที่ไม่เคยทำบ้านเรียน) เป็นไปได้ว่าเราอาจต้องนิยามความหมายของ Space&Time ปล่อยสัตว์ไปเดิน ทำให้สมดุล ยืดหยุ่นและเหมาะสมในแบบของเราเอง ลองดูว่าในชั้นหนังสือ ห้องครัว ในสวน พื้นที่รอบชุมชน สามารถเรียนรู้จากการเล่น ทุกคนมีโอกาสเป็นเจ้าของเวลาของตัวเอง คำว่า ‘บ้าน’ ทดลองกันออกแบบเมนูอาหาร เช้ากลางวันเย็น ไฟอ่อนหรือแรงได้ผลต่างกันไหม เครื่องปรุงต้องใช้อะไร ในกรณีของลูกชายคนโต เรานำอะไรเป็นเครื่องมือการเรียนรู้ได้บ้าง เช่นเราชวนลูกคุยว่าถ้าในหนึ่งสัปดาห์นี้เราจะไม่ออกไปตลาดเลย ใช้บรรยากาศชุมชน

ตัวอย่างที่ 2

ตัวอย่างบทความ

ย้อนกลับไปเมื่อปี 2561 คงไม่มีนักเรียนชั้นไหนเครียดไปกว่าน้อง ป.1 ป.4 และพี่ ม.1 ม.4 อีกแล้ว เพราะนั่นคือปีแรกที่พวกเขาต้องเผชิญกับวิชาบังคับใหม่จากกระทรวงศึกษาธิการที่ชื่อ ‘วิชาวิทยาการคำนวณ’ และไม่ใช่เพียงนักเรียนหรือที่มันติดกับวิชานี้ คุณครูที่จับพลัดจับผลูต้องมาสอนวิชานี้ก็มันไม่ต่างกัน หนึ่งในนั้นคือ ‘ครูฝ้าย’ ศรา หรุจิตตวิวัฒน์ แห่ง โรงเรียนเซนต์ฟรังซิสซาเวียร์คอนแวนต์

เพราะนี่คือวิชาใหม่แกะกล่องสำหรับระดับประถมและมัธยม ครูเองก็ไม่เคยสอน ตำราที่ได้มาก็ยากแสนยาก สอนไปสอนมากังงกันไปทั้งคาบ ทั้งครูทั้งนักเรียน

แต่แทนที่จะนั่งงง ครูฝ้ายจึงเปลี่ยนวิกฤตินี้เป็นโอกาส...

เมื่อนวัตกรรมหน้าใสอย่าง หลิง-สุภาวดี ภูสนาม (ม.6), ปังจิง-ณามา หิรัญญการ (ม.6) และ เหม-เหม-จิรนนท์ อนันต์ประภากรณ์ (ม.4) กำลังหาหัวข้อทำโครงงานอยู่ ครูฝ้ายจึงไม่รอช้า เสนอให้ทีมทำสื่อการเรียนการสอนวิชานี้ขึ้นมาใหม่ นำความยากมาयीใหม่ให้ง่ายขึ้นสำหรับครู และสนุกขึ้นสำหรับนักเรียน

ฟังดูแล้วน่าสนุก แต่เมื่อได้ลงมือทำจริง 3 นวัตกรรมหน้าใสของเราก็พบว่า การทำเรื่องยากให้ง่ายขึ้น...ไม่ง่ายเลย

คนเราเกี่ยวข้องสัมพันธ์กับสิ่งใด ย่อมมีพื้นฐานความเข้าใจในสิ่งนั้นเป็นอย่างดี นักพัฒนาที่สามารถพัฒนานวัตกรรมและตอบโจทย์ผู้ใช้ ส่วนใหญ่จึงมักเริ่มจากสิ่งที่อยู่ใกล้ตัวหรือเป็นสิ่งที่เขาสนใจ ต้องเกี่ยวข้องสัมพันธ์อยู่ในชีวิตประจำวัน

3 สาวแห่งโรงเรียนเซนต์ฟรังซิสซาเวียร์คอนแวนต์ หลิง-ปังจิง-เหม-เหม ก็เป็นเช่นนั้น ในฐานะนวัตกรรมที่ยังเป็นนักเรียน การหาหัวข้อผลงานที่ผ่านๆ มาของพวกเธอจึงเน้นไปที่เรื่องการศึกษา โดยเฉพาะการสร้างโปรแกรมสื่อการเรียนรู้อื่นๆ เพื่อตอบสนองผู้ใช้ที่ส่วนใหญ่ก็เป็นเพื่อนนักเรียนเหมือนๆ กับเธอ หากแต่สำหรับผลงาน Algolaxy แอปพลิเคชันสื่อการเรียนการสอนเรื่องอัลกอริธึม (Algorithm) ในครั้งนี้เป็นความต้องการจากผู้ใช้ที่ไม่ใช่ นักเรียน หน้าซำยังอยู่ใกล้ตัวพวกเธอมาก จนคาดไม่ถึง

“เราจะทำผลงานส่งประกวด NSC ตั้งใจว่าอยากทำโปรแกรมที่เป็นสื่อการเรียนรู้ แต่ยังไม่รู้ว่าจะทำเกี่ยวกับวิชาอะไรดี ลองถามครูฝ้ายซึ่งเป็นทีปรึกษาโครงงานว่ามีวิชาอะไรที่น่าสนใจ ครูฝ้ายบอกว่ามีวิชาใหม่คือวิชาวิทยาการคำนวณ ยังไม่มีใครทำสื่อการเรียนรู้อะไรเลยหาเนื้อหาหามาลองทำ” หลิงเล่าถึงที่มาของหัวข้อผลงาน ซึ่งเกิดมาจากครูฝ้ายที่อยู่ในฐานะผู้ใช้โดยตรง

“อัลกอริธึมมันเกี่ยวกับเรื่องลำดับความคิด ขั้นตอนการทำงาน ที่ผ่านมาถ้าอ่านแต่ตัวอักษรจะเข้าใจยาก เลยมองว่าถ้าทำเป็นสื่อการเรียนการสอนน่าจะเป็นประโยชน์มาก เพราะครูหลายคนก็เพิ่งเริ่มต้นสอนวิชานี้เหมือนกัน” ครูฝ้ายส่ายหัว

Algolaxy จึงตั้งต้นขึ้นมาเพื่อตอบสนองกลุ่มเป้าหมาย 2 กลุ่ม หนึ่งคืออาจารย์ที่สอนวิชาอัลกอริธึม และอีกหนึ่งคือนักเรียนที่เรียนวิชานี้

โจทย์ของทีมในการพัฒนาผลงานเพื่อตอบสนองผู้ใช้กลุ่มอาจารย์ คือการพัฒนาโปรแกรมที่มีเนื้อหาวิชาอัลกอริธึมที่ถูกต้อง ซึ่งถือเป็นงานหนักสำหรับ 3 สาวและครูฝ้ายเอง ที่ต้องไปศึกษาเก็บข้อมูลความรู้เกี่ยวกับอัลกอริธึมทั้งเชิงกว้างและเชิงลึก เพื่อออกแบบเนื้อหาโปรแกรม พร้อมทั้งตรวจสอบความถูกต้องของเนื้อหาต่างๆ ด้วยการขอความช่วยเหลือจากผู้รู้

“ด้วยความที่เป็นวิชาใหม่ เราจึงต้องเริ่มศึกษาตั้งแต่พื้นฐาน แต่เราก็อยากให้แอปพลิเคชันนี้ครอบคลุมเนื้อหาเชิงลึกมากกว่าแค่ระดับพื้นฐาน เราจึงไปขอความรู้จากอาจารย์มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี และรุ่นพี่ศิษย์เก่าเซนต์ฟรังฯ ที่ไปเรียนมหาวิทยาลัยในสาขาที่เกี่ยวข้อง ให้ช่วยมาทดสอบเนื้อหา ทั้งลำดับความรู้และความถูกต้อง” ปังจึงอธิบายกระบวนการค้นคว้าและตรวจสอบข้อมูลของทีม

โจทย์อีกข้อของทีมในการพัฒนาผลงานเพื่อตอบสนองผู้ใช้กลุ่มนักเรียน ก็คือการทำเนื้อหาที่ยากให้ง่ายขึ้น ซึ่งนอกจากการนำเสนอด้วยรูปแบบแอนิเมชัน 2 มิติแล้ว การเลือกสรรและแปรรูปเนื้อหาวิชาให้เชื่อมโยงถึงชีวิตประจำวันของเด็กนักเรียน ก็เป็นสิ่งที่ทีมเลือกทำ แม้จะเป็นงานที่หนักเอาเรื่องไม่น้อย

“เนื้อหาอัลกอริธึมมันมีเยอะมาก พวกหนูต้องพยายามเลือกเรื่องที่เกี่ยวข้องกับชีวิตประจำวันมาสอนเด็ก นำมาประยุกต์ว่าจะทำยังไงให้เขาเข้าใจได้ง่ายๆ เห็นภาพ และไม่ไกลเกินตัวเขา เขาจะได้รู้ว่าอัลกอริธึมมันอยู่ในชีวิตเรา แต่เราไม่รู้ว่าถ้าใช้มันแล้วจะทำให้ชีวิตดีขึ้น เช่น การคิดเป็นลำดับ ถ้าเราจะทำอะไรสักอย่างเราต้องคิดเป็นสเต็ปในหัว 1 2 3 4 แล้วเราจะรู้ว่าในลำดับเดียวกันอันไหนควรทำก่อนหรือหลัง” หญิงยกตัวอย่าง

“มันยากตรงการคิดแบบฝึกหัด เพราะพอเราเข้าใจแล้ว เราต้องปรับให้คนอื่นเข้าใจตามเราให้ได้ด้วย พยายามที่จะทำให้เข้าใจง่ายที่สุด เช่น เรื่องการจัดเรียงตัวเลขจากน้อยไปหามาก มันมีหลายวิธีมากเลย จะทำให้เขารู้ได้ยังไงว่าการจัดเรียงแบบนี้มันชื่อว่าอะไร กระบวนการไหนดีกว่าหรือเร็วกว่า จากเนื้อหาในหนังสือเราก็แปลงเป็นรูปภาพ ทำเป็นแบบฝึกหัดให้เขาได้ทำ” ปังจึงเสริมถึงความยากที่ทีมต้องฟันฝ่ามาด้วยกัน

Algolaxy ถูกพัฒนาส่งเข้าประกวด NSC ตามที่หญิงและปังจ้งตั้งใจ โดยแรกที่เดียวนั้นหญิงรับหน้าที่เขียนโค้ด รุ่นพี่อีกคนหนึ่งรับหน้าที่ทำกราฟิก ส่วนปังจ้งช่วยงานทั้ง 2 ส่วน กระทั่งผลงานต่อ ยอด

มาสู่โครงการต่อกล้าให้เติบโตใหญ่ ปีที่ 6 ปังจังกี้หันมารับงานกราฟิกเต็มตัว ส่วนเหม่เหม่เข้ามาช่วยหญิงเขียนโค้ด

“ที่เรามาทำต่อกล้าฯ เพราะอยากได้คำแนะนำเพิ่มเติม และอยากให้ผลงานเอาไปใช้ได้จริงๆ ตอน NSC ยังไม่ได้เอาไปใช้จริง ก็ต้องปรับเยอะมาก เหมือนเป็นเวอร์ชันใหม่เลย (หัวเราะ) หลักๆ ก็คือปรับเนื้อหาให้เข้าใจง่ายขึ้น เพิ่มแบบฝึกหัด และแก้ UI UX” หญิงเล่าพลางอมยิ้ม

เหตุแห่งการปรับแก้งาน เน้นอนว่าด้านหนึ่งมาจากกรรมการและทีมโค้ช ซึ่งส่วนใหญ่เป็นเรื่อง UI UX

“มีแก้งานจากที่พี่เขาคอมเมนต์มา ทั้งเรื่องการปรับโลโก้ไอคอนของแอปฯ ที่ไม่สื่อว่าเป็นแอปฯ อะไร ปรับแก้ปุ่ม Tutorial ให้มีความชัดเจนมากขึ้น หลักๆ มันคือ UI UX ที่พี่เขาบอกว่าเราใช้ปุ่มหลากหลายเกินไป อย่างปุ่ม Next ก็จะมีทั้งแบบเป็นวงกลมและลูกศร เขาก็ให้เราทำให้เป็นรูปแบบเดียวกัน” ปังจังกี้อธิบาย

กับอีกด้านหนึ่งนั้นมาจากกลุ่มผู้ใช้ ซึ่งถือเป็นจุดเด่นของทีมนี้ที่ทำงานร่วมกับผู้ใช้อย่างเหนียวแน่น ทั้งในระดับเชิงกว้างและเชิงลึก

“เรามีกระบวนการทำงานร่วมกับผู้ใช้ตลอด หลังจากไปรีเสิร์ชกับพี่ๆ ศิษย์เก่าจนได้เนื้อหาออกมา ออกแบบแอปฯ เราก็เอาแอปฯ กลับไปให้พี่ๆ กลุ่มเดิมทดสอบอีก รวมถึงกลุ่มผู้ใช้อื่นๆ ด้วย เช่น คุณครูในเครือข่ายเน็ตบอลฯ (คณะภคินีเซนต์ปอล เดอ ชาร์ตร์ แขวงประเทศไทย ปัจจุบันมีโรงเรียนในสังกัด 22 แห่ง) และที่สำคัญคือกลุ่มนักเรียนชั้น ม.4 ซึ่งต้องเรียนวิชาอัลกอริธึมนี้โดยตรง” ครูฝ้ายเล่าพลางหันมองเหม่เหม่ในฐานะที่เป็นกลุ่มผู้ใช้โดยตรง

“หนูใช้โปรแกรมนี้ในการเรียนการสอนจริง ซึ่งมันเข้าใจง่ายขึ้นเยอะมากๆ แล้วหนูยังลองเอาไปให้เพื่อนที่โรงเรียนเก่าลองใช้ด้วย เป็นการขายของนิดนึง (หัวเราะ) เพื่อนก็บอกว่าดี เพราะแต่เดิมเรียนยากมาก ครูให้อ่านหนังสือแล้วสอนแต่เนื้อหาเพียวๆ ไม่มีภาพไม่มีเหตุการณ์ประกอบ เพื่อนเขาก็ไม่เข้าใจ ไม่เข้าใจว่าเรียนไปทำไม และไม่เข้าใจว่าเอาไปใช้อะไรได้บ้าง” เหม่เหม่เล่าด้วยรอยยิ้ม

ไม่เพียงแค่ว่าทดลองใช้เฉยๆ เท่านั้น แต่ทีมยังยกระดับกระบวนการทดลองใช้ไปถึงขั้นใช้เป็นเครื่องมือในการสอบเก็บคะแนนจริง ในวิชาวิทยาการคำนวณเลยทีเดียว

“เราจะคิดโจทย์ขึ้นใหม่เพื่อทดสอบโดยเฉพาะ เพราะเราใช้แอปฯ นี้ในการเรียนและการสอบวัดผลเก็บคะแนนจริงๆ ซึ่งเหมือนเราได้รีเช็กแอปฯ ไปในตัว เช่น ถ้ามีบักช้อหนึ่งก็ต้องรื้อกันใหม่” ครูฝ้ายเล่า

ด้วยความที่ใช้ในการสอนและการสอบจริงกับนักเรียน ม.4 ทั้งโรงเรียน ก็ทำให้ทีมจำเป็นต้องกระตือรือร้นในการรับฟังข้อผิดพลาดและเร่งแก้ไขแอปฯ แบบวันต่อวัน เพราะว่า...

“ทันทีที่เจอบักเราต้องแจ้งหญิงให้แก้เลย เพราะวันรุ่งขึ้นมีสอนอีกห้องหนึ่ง เราไม่อยากให้ห้องต่อๆ ไปเพลถ้าต้องมาเจออีก ดังนั้นเราจะแก้เป็นสัปดาห์ๆ ไปเลย เนื้อหามี 8 บท ก็เทสต์ 8 สัปดาห์ และแก้ตลอด 8 สัปดาห์ ต้องทำอย่างนี้เพราะเจตนาของเราคือต้องการให้มันใช้ได้จริงๆ ทำให้เด็กเข้าใจเนื้อหา

“ไม่อยากจะได้ครูสักว่าเรียนเรื่องนี้แล้วอยาก เรียนเรื่องนี้แล้วไม่ได้ใช้ พอแอปฯ มันใช้ได้จริงเราก็ภูมิใจ” ครูฝ่ายอิมพ์ท้ายประโยค

ด้วยกระบวนการตามติดผลจากผู้ใช้จริงนี่เอง ที่ทำให้ทีมสามารถพัฒนาแอปฯ ได้สมบูรณ์อย่างรวดเร็ว

“วิธีนี้ทำให้การทำงานเร็วขึ้นมาก เพราะต้องแก้แบบวันต่อวัน อย่างสอนวันพฤหัสบดีเสร็จปุ๊บต้องแก้ไขให้เสร็จทันเช้าวันศุกร์เพราะมีเรียนอีกห้อง ช่วงแรกๆ จะเหนื่อยหน่อยเพราะเจอบั๊กเยอะมาก ก็แก้มาจนห้องหลังๆ จะไม่ค่อยเจอบั๊กแล้ว” หลิงเล่าอย่างร่าเริง

ด้วยความทุ่มเทและตั้งใจจริง ผลประโยชน์ก็ตกแก่ผู้ใช้กลุ่มนักเรียน ซึ่งส่วนใหญ่บอกเป็นเสียงเดียวกันว่า แอปฯ นี้ช่วยให้เรียนรู้อย่างสนุกและเข้าใจวิชาวิทยาการคำนวณมากขึ้น

“ด้วยรูปแบบของเกม ทำให้สนุกเพลิดเพลิน ทำให้เราตั้งใจทำมันมากขึ้นและจดจำได้มากขึ้น ถือเป็นการเรียนรู้ไปด้วยระหว่างที่เล่นเกม ช่วยให้เราคิดอย่างมีขั้นตอน รู้จักวางแผน รอบคอบในการลงมือทำอะไร” คือหนึ่งเสียงจากนักเรียน ม.4 โรงเรียนเซนต์ฟรังซิสซาเวียร์คอนแวนต์ ที่ได้ทดลองใช้แอปฯ นี้

และนอกจากกลุ่มนักเรียนแล้ว ยังรวมไปถึงกลุ่มอาจารย์ในเครือเซนต์ปอลฯ ที่ได้นำไปทดลองใช้แล้วให้ฟีดแบ็คกลับมาในทางที่ดีไม่ต่างกัน

“ในการประชุมครูในเครือเซนต์ปอลฯ มีการอบรมเรื่องวิทยาการคำนวณพอดี เราจึงคิดว่าเป็นโอกาสที่ดี จึงนำแอปฯ ไปแนะนำให้คุณครูจาก 22 โรงเรียนได้ลองใช้ พร้อมทั้งแจกซีดีให้คุณครูนำไปลองใช้กับนักเรียนที่โรงเรียนต่อ และขอให้ครูส่งฟีดแบ็ค เช่น อัดคลิปวิดีโอส่งกลับมา ซึ่งฟีดแบ็คส่วนใหญ่ก็บอกว่าดี และให้กำลังใจกันเยอะ” บังจิ่งเล่าอย่างมีความสุข

“เป็นสื่อการเรียนการสอนที่เหมาะสมกับคุณครูที่อยากให้นักเรียนไปศึกษาเองก่อนเบื้องต้นแล้วค่อยกลับมาเรียนในห้องเรียน หรือใช้เป็นสื่อการเรียนการสอนในห้องเรียนโดยใส่ลงไปในคอมพิวเตอร์ทุกเครื่องให้นักเรียนหาความรู้ในห้องเรียน จดใส่สมุด มาทำโปรแกรมเองได้ หรือสมมุติผู้เรียนเรียนในห้องแล้วเขาไม่เข้าใจตัวเนื้อหา บางคนอาจจะไม่ชอบ เด็กบางคนสมาธิสั้น ไม่ถนัดเรื่องการอ่านเนื้อหาที่เป็นหนังสือมากๆ ถ้าได้โปรแกรมนี้อเข้ามาช่วยจะทำให้เขามีสมาธิอยู่กับตรงนี้มากขึ้น” คือหนึ่งในความเห็นจากคุณครูโรงเรียนเซนต์โยเซฟคอนแวนต์ ที่ได้ทดลองใช้

ขณะที่คุณครูจากโรงเรียนอัสสัมชัญคอนแวนต์ ก็ให้คำชื่นชมไม่ต่างกันว่า “ประโยชน์ของแอปพลิเคชันนี้ ดูจากวัตถุประสงค์ของผู้จัดทำ เนื้อหาเป็นการเสนอเกี่ยวกับการแก้ไขปัญหาด้วยกระบวนการเทคโนโลยี การฝึกให้รู้จักลำดับขั้นตอนการคิด หรือที่เราเรียกว่าอัลกอริธึม สิ่งเหล่านี้จะเป็นพื้นฐานของการนำไปพัฒนาโปรแกรมได้ในอนาคต ครูคิดว่าเป็นเรื่องที่ดีเลยทีเดียวในการจัดทำบทเรียนนี้ขึ้นมา”

การทำงานร่วมกับผู้ใช้ โดยเฉพาะในมิติของการเก็บข้อมูลจากผู้ใช้จริงและเร่งปรับแก้ผลงานให้ดีที่สุดเพื่อผู้ใช้กลุ่มต่อไป ไม่อาจปฏิเสธเลยว่านี่คือกระบวนการทำงานที่หนักหนาสาหัสไม่น้อยสำหรับนวัตกรรมหน้าใสวัยเรียนเช่น 3 สาว

อย่างไรก็ตาม ไม่อาจปฏิเสธว่าความทุ่มเทที่ทั้งสามทุ่มลงไปในการพัฒนาผลงาน นอกจากจะทำให้ Algolaxy พัฒนาไปถึงจุดที่ใช้งานได้จริง รอการขยายผลไปสู่ผู้ใช้ในวงกว้าง รวมไปถึงได้รางวัลรองชนะเลิศ อันดับ 1 จากประกวด TICTA (Thailand ICT Awards) และได้สิทธิ์ไปแข่ง APICTA (Asia Pacific ICT Alliance) ต่อที่ประเทศจีนแล้ว ความสำเร็จของผลงานก็หมุนทวนย้อนกลับมาสู่ตัวพวกเธอเอง ในแง่ของพัฒนาการรายบุคคลที่เกิดจากกระบวนการทำงานหนัก

“สำหรับหนูไม่คิดว่าเราต้องมาแก็งานสัปดาห์ต่อสัปดาห์ (หัวเราะ) อย่างที่เข้าต่อกล้าๆ ปีก่อนๆ จะทำอาทิตย์ละครั้งช่วงเสาร์อาทิตย์ ทำเรื่อยๆ ไม่ได้กำหนดเป้าหมาย แต่งานนี้ครูฝ้ายสอนเสร็จปั๊บหนูต้องมาแก้ให้ทันวันถัดไป มันก็ค่อนข้างตื่นเต้น ต้องทำให้ทัน เพราะมันคือการนำไปใช้งานจริงกับผู้ใช้ ได้ฝึกความรับผิดชอบอย่างเข้มข้นมาก (หัวเราะ) และได้เห็นพีตแบ็คจริงๆ เวลาที่ครูนำแพอฯ ไปสอนน้องๆ มันมีคำชมกลับมาที่เราโดยตรง ก็รู้สึกภูมิใจ” หญิงเล่าอย่างอารมณ์ดี

ก่อนที่ปิงจิงจะเล่าในส่วนของตัวเองว่า

“การทำแพอฯ นี้สอนหนูหลายอย่าง ที่เด่นที่สุดน่าจะเป็นเรื่องกระบวนการทำงาน แพอฯ อื่นๆ หนูจะเร่งทำรวดเดียวตอนใกล้ส่ง (ยิ้ม) เวลาเกิดปัญหาก็จะไม่ทัน เหมือนเรือแล่นเร็วๆ แล้วเจอพายุก็ล่มง่าย แต่งานนี้เหมือนเรือที่แล่นสม่ำเสมอ ปลอดภัยกว่า”

ตัวอย่างสรุปจากแบบจำลอง

ทำให้สนุกเพลิดเพลิน แล้วหนูยังลองเอาไปให้เพื่อนที่โรงเรียนเก่าลองใช้ด้วย ซึ่งพีตแบ็คส่วนใหญ่ก็บอกว่าดี และให้กำลังใจกันเยอะ” ปิงจิงเล่าอย่างมีความสุข

“เป็นสื่อการเรียนการสอนที่เหมาะสมกับคุณครูที่ยากให้นักเรียนไปศึกษาเองก่อนเบื้องต้นแล้วค่อยกลับมาเรียนในห้องเรียน หรือใช้เป็นสื่อการเรียนการสอนในห้องเรียนโดยใส่ลงในคอมพิวเตอร์ทุกเครื่องให้นักเรียนหาความรู้ในห้องเรียน หากแต่สำหรับผลงาน Algolaxy แอพพลิเคชั่นสื่อการเรียนการสอนเรื่องอัลกอริธึม (Algorithm) ในครั้งนี้ เป็นความต้องการจากผู้ใช้ที่ไม่ใช่ นักเรียน หน้าที่ยังอยู่ใกล้ตัวพวกเธอมาก จนคาดไม่ถึง

“เราจะทำผลงานส่งประกวด NSC ตั้งใจว่าอยากทำโปรแกรมที่เป็นสื่อการเรียนรู้ แต่ยังไม่รู้ว่าจะทำเกี่ยวกับวิชาอะไรดี แพอฯ อื่นๆ เขาจะได้อัลกอริธึมมันอยู่ในชีวิตเรา แต่เราไม่รู้ว่าถ้าใช้มันแล้วจะทำให้ชีวิตดีขึ้น เช่น การคิดเป็นลำดับ พอแพอฯ ปลอดภัยกว่า”

เพราะวันรุ่งขึ้นมีสอนอีกห้องหนึ่ง ขั้นตอนการทำงาน ครูเองก็ไม่เคยสอน ต้องทำให้ทัน และแก้ตลอด 8 สัปดาห์ ดูจากวัตถุประสงค์ของผู้จัดทำ เด็กบางคนสมาธิสั้น ช่วยให้เราคิดอย่างมีขั้นตอน รู้จักวางแผน รอบคอบในการลงมือทำอะไร” คือหนึ่งเสียงจากนักเรียน ม.4 โรงเรียนเซนต์ฟรังซิสซาเวียร์คอนแวนต์ ที่ได้ทดลองใช้แพอฯ นี้

และนอกจากกลุ่มนักเรียนแล้ว ไม่มีภาพไม่มีเหตุการณ์ประกอบ มันก็ค่อนข้างตื่นเต้น ไม่เข้าใจว่าเรียนไปทำไม และไม่เข้าใจว่าเอาไปใช้อะไรได้บ้าง” เหม่เหม่เล่าด้วยรอยยิ้ม

ไม่เพียงแค่ทดลองใช้เฉยๆ เท่านั้น แต่ทีมยังยกระดับกระบวนการทดลองใช้ไปถึงขั้นใช้เป็นเครื่องมือในการสอบเก็บคะแนนจริง ในวิชาวิทยาการคำนวณเลยทีเดียว

“เราจะคิดโจทย์ขึ้นใหม่เพื่อการสอบโดยเฉพาะ เพราะเราใช้แอปฯ นี้ในการเรียนและการสอบวัดผลเก็บคะแนนจริงๆ ผลประโยชน์ก็ตกแก่ผู้ใช้กลุ่มนักเรียน เลยมองว่าถ้าทำเป็นสื่อการเรียนการสอน น่าจะเป็นประโยชน์มาก คิดคิดว่าเป็นเรื่องที่ดีเลยทีเดียวในการจัดทำบทเรียนนี้ขึ้นมา”

การทำงานร่วมกับผู้ใช้ โดยเฉพาะในมิติของการเก็บข้อมูลจากผู้ใช้งานจริงและเร่งปรับแก้ผลงานให้ดีที่สุดเพื่อผู้ใช้กลุ่มต่อไป ไม่อาจปฏิเสธเลยว่านี่คือกระบวนการทำงานที่หนักหนาสาหัสไม่น้อยสำหรับนวัตกรรมหน้าใสวัยเรียนเช่น 3 สาว

อย่างไรก็ตาม ไม่อาจปฏิเสธว่าความทุ่มเทที่ทั้งสามทุ่มลงไปในการพัฒนาผลงาน เราจึงไปขอความรู้จากอาจารย์มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี และรุ่นพี่ศิษย์เก่าเซนต์ฟรังฯ ที่ไปเรียนมหาวิทยาลัยในสาขาที่เกี่ยวข้อง ให้ช่วยมาทดสอบเนื้อหา ทั้งลำดับความรู้และความถูกต้อง” ปังจิงอธิบายกระบวนการค้นคว้าและตรวจสอบข้อมูลของทีม

โจทย์อีกข้อของทีมในการพัฒนาผลงานเพื่อตอบสนองผู้ใช้กลุ่มนักเรียน ก็คือการทำเนื้อหาที่ยากให้ง่ายขึ้น ซึ่งมันเข้าใจง่ายขึ้นเยอะมากๆ มันมีหลายวิธีมากมาย นักพัฒนาที่สามารถพัฒนานวัตกรรมและตอบโจทย์ผู้ใช้ ในฐานะนวัตกรรมที่ยังเป็นนักเรียน บางคนอาจจะไม่ชอบ ไม่อยากให้เด็กรู้สึกว่ายากแล้วยาก ทำให้เด็กเข้าใจเนื้อหา เพื่อนเขาก็ไม่เข้าใจ ด้วยการขอความช่วยเหลือจากผู้รู้

“ด้วยความที่เป็นวิชาใหม่ โดยแรกทีเดียวนั้นหญิงรับหน้าที่เขียนโค้ด ครูฝ่ายบอกว่ามีวิชาใหม่คือวิชาวิทยาการคำนวณ เพราะแต่เดิมเรียนยากมาก รอการขยายผลไปสู่ผู้ใช้ในวงกว้าง เนื้อหามี 8 บท ก็ทดสอบ 8 สัปดาห์ รวมไปถึงได้รางวัลรองชนะเลิศอันดับ 1 จากประกวด TICTA (Thailand ICT Awards) และได้สิทธิ์ไปแข่ง APICTA (Asia Pacific ICT Alliance) ต่อที่ประเทศจีนแล้ว อย่างปุ่น Next ก็จะมีทั้งแบบเป็นวงกลมและลูกศร

ตัวอย่างที่ 3

ตัวอย่างบทความ

สืบเนื่องจากกรณีน้องเปาเปาโดนแฟนคลับหยิก (ด้วยความรัก-ตามการชี้แจง) ซึ่งสังคมให้ความเห็นอย่างมากมายหลากหลาย ทีมงาน Potential ต่อสายตรงถึง ผู้ช่วยศาสตราจารย์ ดร.ปนัดดา ธนเศรษฐกร ผู้ช่วยผู้อำนวยการด้านศูนย์พัฒนาเด็กปฐมวัย สถาบันแห่งชาติเพื่อการพัฒนาเด็กและครอบครัว มหาวิทยาลัยมหิดล เจ้าของงานวิจัย “1015 เทคนิคการสร้างวินัยเชิงบวก&พัฒนาการสมอง จิตใจ และพฤติกรรม” ถามความเห็นต่อประเด็นดังกล่าว

สิ่งที่ต้องทำความเข้าใจร่วมกันในการโพสต์รูปลูกๆ ลงโซเชียลมีเดีย เพราะภายใต้ความน่ารักน่าเอ็นดูที่คุณพ่อแม่อยากส่งต่อ มันคือข้อมูลส่วนตัวพื้นฐาน ความปลอดภัย และมีผลต่อจิตใจของเด็กในอนาคต

ปัจจุบัน โลกโซเชียลได้เข้ามาเป็นส่วนหนึ่งในชีวิตประจำวันของหลายๆ คนไปแล้ว หรืออาจเรียกได้ว่า โลกโซเชียลได้กลายเป็นโลกแห่งความเป็นจริง และโลกแห่งความเป็นจริงกลายเป็นโลกโซเชียลจากเมื่อก่อนขึ้นมา จิบกาแฟ อ่านหนังสือพิมพ์ พูดคุยกับครอบครัว อาบน้ำ แต่งตัว ไปทำงาน แต่ปัจจุบันโซเชียลมีเดียแทบจะเป็นอย่างแรกที่เราเช็คเมื่อตื่นนอน และเป็นสิ่งสุดท้ายที่จะทำก่อนเข้านอน โซเชียลมีเดียทำให้การเข้าถึงสมาชิกครอบครัว ญาติ เพื่อน หรือคนรู้จัก ได้ง่ายขึ้น ติดต่อกันได้มากขึ้น โกลขึ้นและเร็วขึ้น ส่งผลให้พฤติกรรมใหม่ๆ ของเราเกิดขึ้นมากมาย

Sharenting คือ พฤติกรรมของพ่อแม่ที่โพสต์รูปลูกลงโซเชียลมีเดีย ซึ่งเป็นหนึ่งในตัวอย่างของพฤติกรรมใหม่พ่อแม่ในยุคสมัยนี้ แน่หน่อนว่าการลงรูปของลูกที่ยังเป็น Baby หรือ ยังเป็นเด็กเล็กอยู่ เป็นสิทธิ์ของพ่อแม่ สามารถทำได้ แต่เมื่อลูกเริ่มมีพัฒนาการด้านตัวตน มีความรู้สึกนึกคิดที่ชัดเจน สามารถบอกความรู้สึก และแสดงความคิดเห็นได้ (ประมาณ 3 ขวบเป็นต้นไป) ก่อนลงรูป พ่อแม่จำเป็นต้องขออนุญาตจากลูกก่อน

ถึงแม้ว่า การโพสต์รูปของลูกในโซเชียลมีเดีย มีข้อดีมากมาย เช่น เพื่อส่งข่าวสารให้ครอบครัวคนรู้จัก หรือเพื่อเก็บเป็นความทรงจำดี ๆ แต่การโพสต์รูปลูกลงโซเชียลมีเดียแบบรู้เท่าไม่ถึงการณ์ก็เป็นโทษเช่นกัน

พ่อแม่อาจจะเข้าข่ายละเมิดสิทธิมนุษยชนได้ เช่น สิทธิในการถูกคุ้มครอง ซึ่งเป็นหนึ่งในสิทธิขั้นพื้นฐานที่เด็กจะต้องได้รับจากผู้เลี้ยงดู หรือการละเมิดสิทธิความเป็นส่วนตัวของเด็ก

ดังนั้นก่อนโพสต์รูปของลูก พ่อแม่จำเป็นจะต้องคำนึงถึงสิทธิของลูกที่พ่อแม่ต้องคุ้มครองด้วย

สิ่งที่พ่อแม่ควรทำก่อนตัดสินใจโพสต์รูปของลูกลง Social Media:

ทุกรูปที่โพสต์ลงโซเชียลมีเดียไปแล้วจะไม่สามารถเรียกคืน หรือควบคุมได้ ไม่ว่าพ่อแม่จะลบออกหน้าโซเชียลมีเดียของตัวเอง ก็ไม่อาจแน่ใจได้ว่าคนอื่นจะไม่ได้รูปของลูกเราไป

ดังนั้น พ่อแม่จะต้องคิดพิจารณาให้ดี ว่ารูปนี้จะเกิดโทษกับลูกในอนาคตหรือไม่ เช่น การโพสต์รูปของลูกที่อาจส่งผลให้ลูกอับอายในอนาคต หรือ คนอื่นนำรูปลูกเราไปโพสต์ต่อ ด้วยข้อความที่เป็นเท็จ และไม่เหมาะสม เป็นต้น

ที่จริงแล้ว ทุกคน ทุกวิชาชีพ มีหน้าที่ปกป้องสิทธิส่วนตัว และเคารพสิทธิของผู้อื่น ไม่เว้นแม้แต่โลกโซเชียลที่เราก็ต้องรู้จักปกป้องสิทธิของตนเอง และเคารพสิทธิของผู้อื่นด้วย ก่อนจะโพสต์อะไร หรือจะคอมเมนต์ใคร ก็ต้องคิดพิจารณาก่อนว่า เมื่อแชร์ลงโซเชียลไปแล้ว จะกลายเป็นเป้าให้ผู้อื่นเข้ามาล่วงละเมิดสิทธิเราได้หรือไม่

เพราะเมื่อแชร์ลงไปแล้ว ก็เหมือนเราวางคุกก็ไว้บนโต๊ะ เปิดช่องทางให้ผู้อื่นเข้ามาหยิบชิมฟรีๆ ได้ ชื่นชมก็ได้ ตำหนิก็ได้ แกรมเอากลับบ้านได้อีกด้วย ดังนั้นหากสิ่งที่จะโพสต์เป็นเรื่องส่วนตัว ที่เราไม่พร้อมจะให้คนอื่นมาละเมิดสิทธิเรา เราก็ต้องไม่โพสต์ ไม่แชร์ ปกป้องสิทธิของตัวเองไว้ หากในโลกแห่งความเป็นจริง เราคอยระแวงระวัง ไม่ทิ้งกระเป๋าตังค์ หรือของมีค่าไว้ในที่สาธารณะอย่างไร บนโลกโซเชียลเราก็ต้องระวังอย่างนั้น

แต่สิ่งสำคัญที่เพิ่มเติมสำหรับผู้ใหญ่ทุกคนในสังคม ก็คือ ช่วยกันเคารพและคุ้มครองสิทธิของเด็ก จนกว่าพวกเขาจะสามารถปกป้องสิทธิของตัวเองได้ โดยเฉพาะผู้ที่ทำงานกับเด็กปฐมวัย นอกจากจะต้อง

คำนึงถึงการเคารพ และปกป้องสิทธิของเด็กแล้ว ยังจะต้องคอยเป็นกระบอกเสียงให้เด็ก ๆ คอยพิทักษ์สิทธิเด็ก คอยสื่อสารกับผู้ปกครอง และสังคมให้มีความเข้าใจและสามารถปฏิบัติเรื่องคุ้มครองสิทธิเด็กได้อีกด้วย

ตามหลักการพัฒนาการมนุษย์ พัฒนาการที่สำคัญของเด็กเล็กวัย 0-3 ปี คือ พัฒนาการด้านการรับรู้ตัวตน หรือ sense of self ซึ่งเป็นพัฒนาการพื้นฐานที่มีความสำคัญมาก และส่งผลต่อพัฒนาการ 4 ด้าน และพัฒนาการทักษะสมอง EF ด้วย การรับรู้ตัวตน คือการที่เด็กรู้ว่า ตนเองมีความรู้สึกนึกคิดเป็นของตัวเอง แยกความคิด ความต้องการของตนเองออกจากแม่ และผู้อื่นได้ เด็กรับรู้ว่า คำพูด และการกระทำของตนเองที่แสดงออกมา มีผลกับผู้อื่น และคำพูดและการกระทำของผู้อื่น ก็มีผลกับตัวเองเช่นกัน ซึ่งประสบการณ์ที่สะสมในช่วง 3 ปีแรกนี้เอง จะเป็นต้นกำเนิดของ ความรู้สึกนึกคิดที่มีต่อตนเอง ความชอบ แรงบันดาลใจ การตีความโลกใบนี้ และเป้าหมายในชีวิต

ปัจจัยที่สำคัญที่สุดของการส่งเสริมพัฒนาการด้านการรับรู้ตัวตน คือ การทำที่บ้านให้เป็นฐานที่มั่นคงทางใจให้กับลูก เพื่อให้ลูกโตมาพร้อมกับความพึงพอใจในความสามารถ และความฝันของตนเอง หากเด็กช่วงนี้ได้รับประสบการณ์ที่มีคุณภาพ ก็จะมีความรู้สึกนึกคิดที่ดีต่อตนเอง ตีความโลกใบนี้ในทางที่ดี มีความภาคภูมิใจในตัวเอง รู้ว่าตนเองมีค่ากับพ่อแม่ มีแรงบันดาลใจในการตามฝัน และมีความหวังในการใช้ชีวิต ประสบการณ์เดิมที่มีคุณภาพในช่วงวัย 0-3 คือ ประสบการณ์ความรัก ความอบอุ่น และความผูกพันที่ได้รับจากพ่อแม่ และผู้เลี้ยงดู นอกจากนี้ ประสบการณ์เดิมที่มีคุณภาพ ที่สำคัญและขาดไม่ได้อีกประการ คือ การให้โอกาสเด็กได้ใช้ทักษะสมอง EF ในการคิด ตัดสินใจ แก้ไขปัญหา ลองผิดลองถูก ใช้ความพยายามจนเกิดความสำเร็จตามเป้าหมาย ซึ่งการสะสมประสบการณ์แห่งการทำอะไรให้สำเร็จด้วยตนเองนี้ เป็นกุญแจสำคัญของการส่งเสริมพัฒนาการของเด็กทุกคนให้ก้าวผ่าน พัฒนาการด้านการรับรู้ตนเองไปให้ได้ อย่างมีคุณภาพ เนื่องจากส่งผลให้เด็กรู้ว่าตนเองเป็นคนมีความสามารถ มีค่า และมีความภาคภูมิใจในตนเอง ไม่จำเป็นต้องหาความสุข ความพึงพอใจจากความนิยมที่คนอื่นมอบให้

ข้อดี คือ การได้รับโอกาสที่ได้พบปะผู้คนมากมาย การได้ไปสถานที่หลากหลาย ได้สัมผัส และเรียนรู้โลกใบนี้ในแบบที่แตกต่าง ซึ่งทำให้น้องได้เรียนรู้ ฝึกฝนการปรับตัว ปรับอารมณ์ ให้เข้ากับสถานการณ์ในทุกๆ วัน ซึ่งหากมาพร้อมกับความรักความปลอดภัย และความผูกพันของพ่อแม่ด้วยแล้ว จะกลายเป็นการสะสมประสบการณ์เดิมที่มีคุณภาพมากยิ่งขึ้น

ตัวอย่างสรุปจากแบบจำลอง

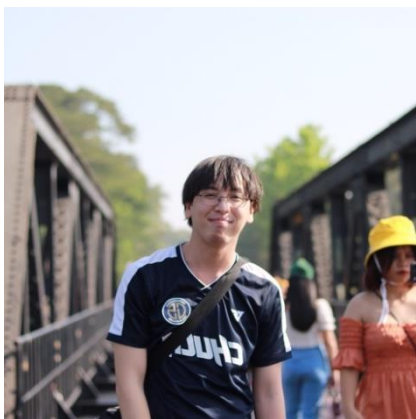
มีความรู้สึกนึกคิดที่ชัดเจน ไม่ทิ้งกระเป๋าดังค์ หรือของมีค่าไว้ในที่สาธารณะอย่างไร บนโลกโซเชียลเราก็ต้องระวังอย่างนั้น

แต่สิ่งสำคัญที่เพิ่มเติมสำหรับผู้ใหญ่ทุกคนในสังคม ก็คือ ช่วยกันเคารพและคุ้มครองสิทธิของเด็ก จนกว่าพวกเขาจะสามารถปกป้องสิทธิของตัวเองได้ ได้ง่ายขึ้น และเร็วขึ้น และเคารพสิทธิของผู้อื่นด้วย ติดต่อผู้คนได้มากขึ้น โกลขึ้น ทักษะอาชีพ ปกป้องสิทธิของตัวเองไว้ ไม่แชร์ มีความภาคภูมิใจในตัวเอง เด็กรับรู้ คำพูด และการกระทำของตนเองที่แสดงออกมา มีผลกับผู้อื่น หากในโลกแห่งความเป็นจริง และผู้อื่นได้ เป็นต้น

ที่จริงแล้ว ทุกคน ตีความโลกใบนี้ในทางที่ดี พ่อแม่จำเป็นต้องขออนุญาตจากลูกก่อน ถึงแม้ว่า การโพสต์รูปของลูกในโซเชียลมีเดีย มีข้อดีมากมาย เช่น เพื่อส่งข่าวสารให้ครอบครัวคนรู้จัก หรือเพื่อเก็บเป็นความทรงจำดี ๆ แต่การโพสต์รูปลูกลงโซเชียลมีเดียแบบรู้เท่าไม่ถึงการณ์ก็เป็นโทษเช่นกัน

พ่อแม่อาจจะเข้าข่ายละเมิดสิทธิมนุษยชนได้ เช่น สิทธิในการถูกคุ้มครอง เราคอยระแวดระวัง

ประวัติผู้เขียน



ชื่อ นายธนะเดช รุ่งศรีพัฒนพร

การศึกษา สาขาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์
และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย

ที่อยู่ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

อีเมล tanadate407@gmail.com



ชื่อ นายชลิท ปฐมกสิวัฒนา

การศึกษา สาขาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์
และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย

ที่อยู่ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

อีเมล chalit101@hotmail.com