



**ข้อเสนอโครงการวิศวกรรมคอมพิวเตอร์**  
**วิชา 01076014 การเตรียมโครงการวิศวกรรมคอมพิวเตอร์**  
**ภาคเรียนที่ 2 ปีการศึกษา 2562**

1. ชื่อหัวข้อโครงการ (ไทย) ..... แต่งวรรณกรรมประเภทเรื่องสั้นด้วยการเรียนรู้เชิงลึก .....
2. ชื่อหัวข้อโครงการ (อังกฤษ) ..... Create Short Fictional Story using Deep Learning .....
3. คำสำคัญ 3 คำ (3 keywords) ..... Deep Learning, Fictional, Neural Networks .....
4. รายชื่อผู้ทำโครงการ
  - 4.1. นาย/นางสาว ..... เขมชาติ ศิลาวรรณ ..... รหัส ..... 60010106 .....
  - 4.2. นาย/นางสาว ..... ธน นิมถาวร ..... รหัส ..... 60010379 .....
5. อาจารย์ที่ปรึกษา
  - 5.1. อาจารย์ที่ปรึกษาหลัก ..... รศ. ดร. เกียรติกุล เจียรนัยธนะกิจ .....
  - 5.2. อาจารย์ที่ปรึกษาร่วม .....

## 1. ที่มาและความสำคัญของปัญหา (Motivation)

เรื่องสั้นเป็นเรื่องที่แต่งขึ้นมาในรูปแบบร้อยแก้วคล้ายคลึงกับนวนิยายแต่สั้นกว่า ได้รับอิทธิพลมาจากตะวันตก เรื่องสั้นมีหลากหลายประเภทหลากหลายแนว ตัวอย่างเช่น เรื่องสั้นแนวแฟนตาซี แนวไซไฟ แนวลึกลับ มีเหตุการณ์ในเรื่อง ตัวละครน้อย มักจะจบแบบเหนือความคาดหมายและอาจจะทิ้งท้ายให้คิดตามแบบฉบับของผู้แต่ง การอ่านเรื่องสั้นสามารถให้ความสนุกสนาน และอาจก่อให้เกิดคำถามจากความคิดที่ได้หลังอ่านเรื่องสั้นจบสำหรับเด็ก การอ่านเรื่องสั้นสามารถพัฒนาทักษะการสื่อสารได้เป็นอย่างดี

ในปัจจุบันเทคโนโลยีปัญญาประดิษฐ์ได้พัฒนาไปอย่างมาก และเข้ามามีบทบาทสำคัญในการดำเนินชีวิต มีการนำปัญญาประดิษฐ์ การเรียนรู้ของเครื่อง และการเรียนรู้เชิงลึกมาใช้แก้ปัญหาที่มีความซับซ้อน ทั้งยังสามารถทำการทำนายอนาคตจากข้อมูลที่มีอยู่ได้อีกด้วย แต่ว่างานบางประเภทนั้นการใช้ปัญญาประดิษฐ์จะทำได้ไม่ดีเท่ากับที่มนุษย์ทำ เช่น งานที่ต้องใช้ความคิดและอารมณ์ การแต่งเพลง การวาดภาพ การเขียนนิยายหรือแต่งเรื่องราวขึ้นมา การเอาปัญญาประดิษฐ์มาใช้กับงานเหล่านี้มันผลลัพธ์ที่ได้จะไม่มีความเป็นธรรมชาติเท่าที่ควร

โครงการ “แต่งวรรณกรรมประเภทเรื่องสั้นด้วยการเรียนรู้เชิงลึก (Create Short Fictional Story using Deep Learning)” นี้จึงถูกจัดทำขึ้น เพื่อการศึกษาวิธีใช้การเรียนรู้เชิงลึกในการเขียนวรรณกรรมประเภทเรื่องสั้นขึ้นมา ด้วยการศึกษาโครงสร้างจากงานเขียนฝีมือมนุษย์และทำการสร้างงานเขียนใหม่ ทั้งนี้เพื่อให้งานเขียนที่ได้มีความเป็นเหตุเป็นผลและมีเรื่องราวที่สัมพันธ์กัน จึงหาเทคนิคและวิธีการต่าง ๆ มาใช้ควบคู่กันไปด้วยเพื่อให้งานเขียนมีความเป็นธรรมชาติมากที่สุด เรื่องราวที่จะได้จะสามารถดำเนินต่อไปได้จากตัวแปรต่าง ๆ ที่กำหนด เป็นเหตุเป็นผล อ่านแล้วเข้าใจ โดยผลสำเร็จจะเป็นโมเดลที่สามารถสร้างวรรณกรรมประเภทเรื่องสั้นขึ้นมาด้วยตัวเองตามประเภทของเรื่องราวที่ป้อนเข้าไป เช่น แฟนตาซี ไซไฟพร้อมกับชุดข้อมูลที่เป็นตัวอย่างเรื่องสั้นสำหรับการประมวลผลในการเรียนรู้เชิงลึก และออกมาเป็นงานซึ่งมีความสั้นไหล อ่านแล้วสามารถเข้าใจได้

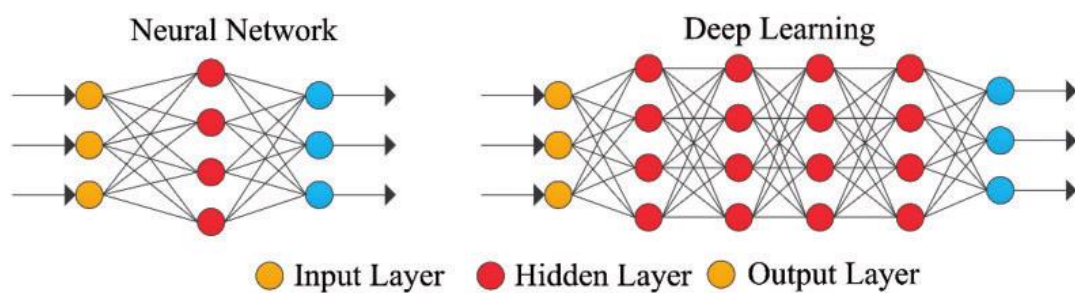
## 2. วัตถุประสงค์ (Objectives)

เพื่อศึกษาวิธีใช้เทคนิคต่าง ๆ ควบคู่กับการเรียนรู้เชิงลึกในการสร้างงานวรรณกรรมประเภทเรื่องสั้นขึ้นมา จากชุดข้อมูลที่เป็นตัวอย่างเรื่องสั้น มีเนื้อหาตรงกับหัวข้อที่ป้อนเข้าไป และสามารถดำเนินเรื่องได้

### 3. ทฤษฎีที่เกี่ยวข้อง (Theoretical Background)

#### 3.1 Artificial Neural Network

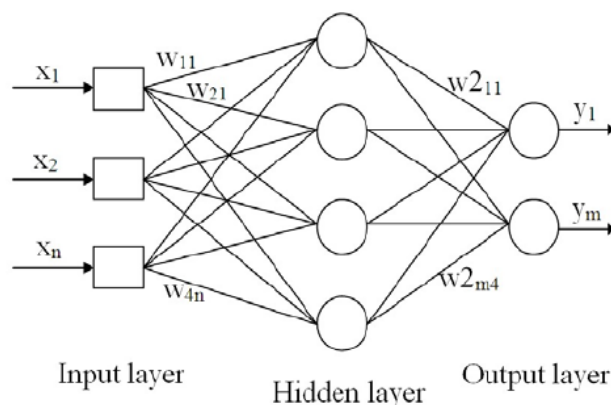
โครงข่ายประสาทเทียม เป็นแบบจำลองทางคณิตศาสตร์หรือโมเดลทางคอมพิวเตอร์ ที่จำลองเซลล์ประสาท (Neurons) ในสมองออกมา แบบจำลองโครงข่ายประสาทเทียมจะประกอบไปด้วย Input layer, Neurons Hidden layer และ Output layer ถ้าหากว่ามีจำนวน Hidden Layer ซ้อนกันเยอะจะเรียกว่า Deep Learning [1]



รูป 3.1 โครงสร้างโครงข่ายประสาทเทียม<sup>1</sup>

#### 3.2 Multi-Layer Perceptron

เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่มีโครงสร้างป้อนไปข้างหน้าเป็นแบบหลายๆชั้น โดยข้อมูลชั้นขาเข้าจะเข้าสู่เซลล์ของชั้นขาออก และเกิดการคำนวณให้เกิดผลลัพธ์ [2]



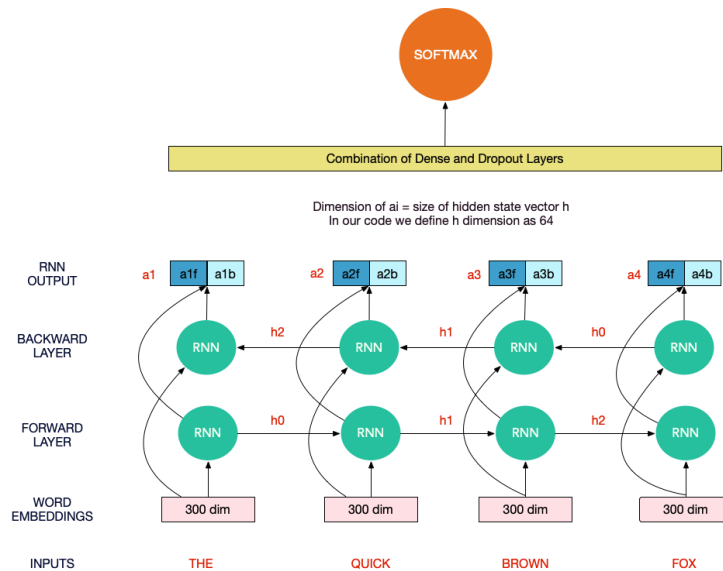
รูป 3.2 โครงสร้าง Multi-Layer Perceptron <sup>2</sup>

<sup>1</sup> <https://mc.ai/สรุป-deep-learning-ที่เรียนมาจากคลา/>

<sup>2</sup> [https://www.researchgate.net/figure/A-schematic-diagram-of-a-Multi-Layer-Perceptron-MLP-neural-network\\_fig3\\_257071174](https://www.researchgate.net/figure/A-schematic-diagram-of-a-Multi-Layer-Perceptron-MLP-neural-network_fig3_257071174)

### 3.3 Natural Language Processing

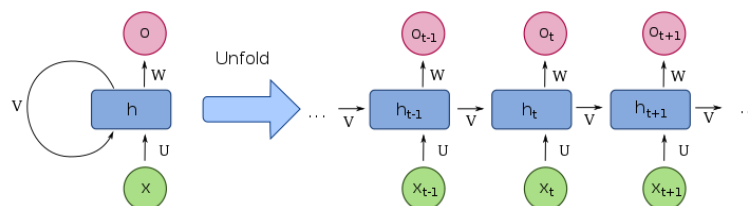
เป็นวิธีการประมวลภาษาธรรมชาติ สาขาย่อยของปัญญาประดิษฐ์และภาษาศาสตร์ ที่ศึกษาเพื่อให้คอมพิวเตอร์สามารถเข้าใจภาษาธรรมชาติหรือภาษามนุษย์ได้ [3] ประกอบไปด้วยวิธีและเทคนิค หลายแบบ ตัวอย่างเช่น Stemming, Tokenization, Sequential Model [4]



รูป 3.3 Bidirectional recurrent neural networks ที่ใช้ในการทำ NLP <sup>3</sup>

### 3.4 Recurrent Neural Network

เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมชนิดหนึ่ง ที่ออกแบบมาเพื่อใช้งานกับข้อมูลที่เป็นลำดับ แนวคิดคือ ทำให้โมเดลสามารถจำสถานะก่อนหน้าได้ โดยการนำเอาพหุการคำนวณที่ได้มาไปใช้เป็นอินพุตในการคำนวณครั้งถัดไปพร้อมกับอินพุตในครั้งนั้น [5]



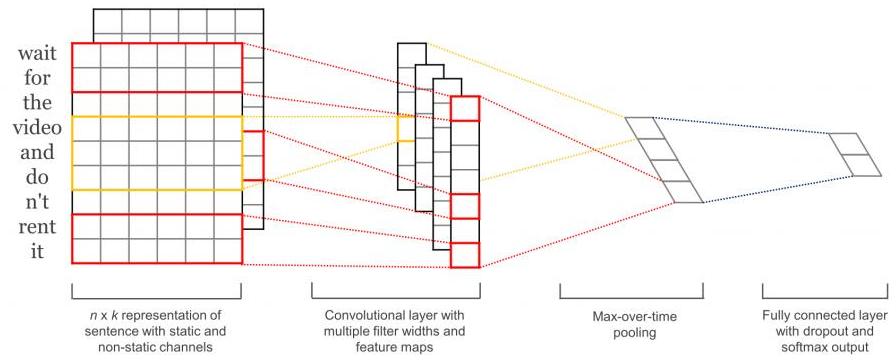
รูป 3.4 โครงสร้าง Recurrent Neural Network <sup>4</sup>

<sup>3</sup> <https://thaikeras.com/2019/workshop-nlp-quora/>

<sup>4</sup> <https://www.bualabs.com/archives/3103/what-is-rnn-recurrent-neural-network-what-is-gru-gated-recurrent-unit-teach-how-to-build-rnn-gru-with-python-nlp-ep-9/>

### 3.5 Convolutional Neural Network

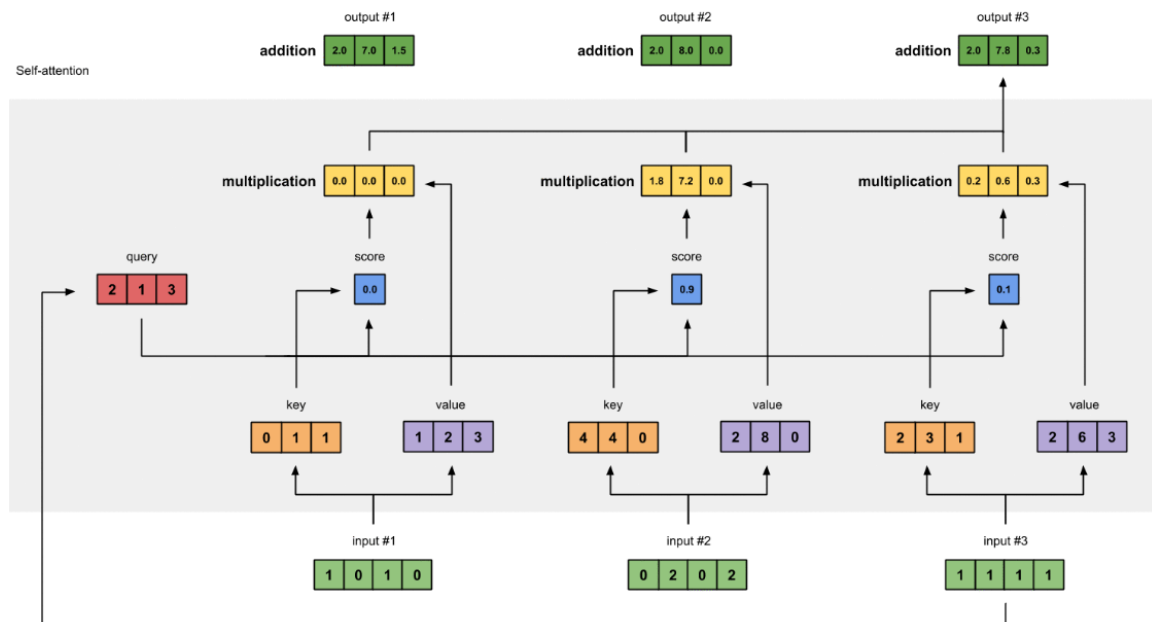
เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมชนิดหนึ่ง ประกอบไปด้วย Convolutional layer และ Pooling layer สามารถใช้คำนวณประโยคหลายๆ ประโยค พร้อมกันเป็นแบบขนานได้ [6]



รูป 3.5 ใช้ โครงสร้าง Recurrent Neural Network <sup>5</sup>

### 3.6 Self Attention Mechanism

เป็นกระบวนการที่จะใช้ Input จากส่วนอื่นเข้ามาปรับค่าของตัวเองได้ และทำการพิจารณาว่าโมเดลควรจะให้ความสำคัญกับค่าไหน [7]



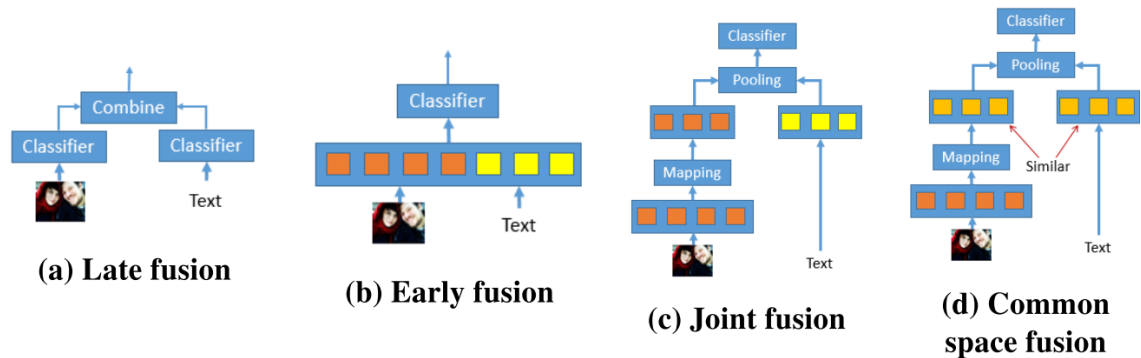
รูป 3.6 กระบวนการปรับค่าด้วย Self Attention <sup>6</sup>

<sup>5</sup> <http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/>

<sup>6</sup> <https://towardsdatascience.com/illustrated-self-attention-2d627e33b20a>

### 3.7 Model Fusion

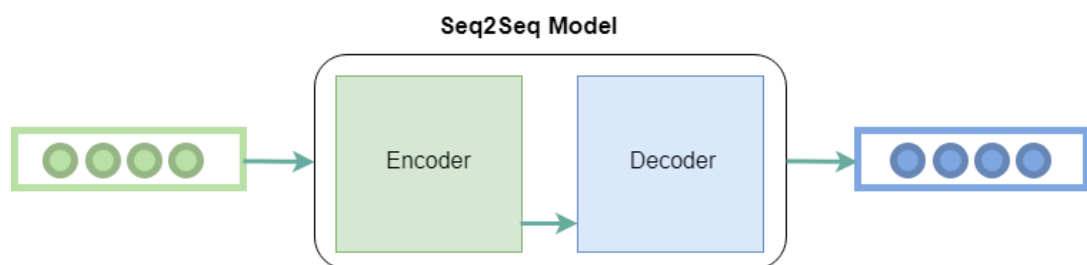
โมเดลฟิวชัน เป็นการรวมเอาโมเดลที่มีรูปแบบต่างมารวมกันเพื่อใช้ประมวลผลด้วยการเรียนรู้เชิงลึก จะทำให้ Feature ของทั้งสองรูปแบบมีขนาดเท่ากัน [8]



รูป 3.7 เทคนิคต่าง ๆ ในการ Fusion โมเดล <sup>7</sup>

### 3.8 Sequence to Sequence Model

เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมชนิดหนึ่ง ที่จะรับประโยคเข้าไปเป็นอินพุต แล้วเอาพหุออกมาเป็นประโยคถัดไป ทำโดยใช้ RNN , LSTM หรือ GRU ประกอบไปด้วย ส่วน Encoder และส่วน Decoder ส่วน Encoder จะแปลงแต่ละประโยคเป็นเวกเตอร์ที่เก็บประโยคและบริบท ส่วน Decoder จะทำกระบวนการย้อนกลับ คือเปลี่ยนเวกเตอร์ที่ได้มาเป็นประโยคเอาท์พุท และใช้เอาท์พุทของการคำนวณครั้งก่อนมาเป็นอินพุตบริบท [9]



รูป 3.8 Sequence to Sequence Model <sup>8</sup>

<sup>7</sup> <https://medium.com/dair-ai/detecting-emotions-with-cnn-fusion-models-b066944969c8>

<sup>8</sup> <https://www.geeksforgeeks.org/seq2seq-model-in-machine-learning/>

## 4. งานวิจัยที่เกี่ยวข้อง (Related Works)

### 4.1 Hierarchical Neural Story Generation [10]

งานวิจัยเกี่ยวกับการสร้างเรื่องราวที่มีความต่อเนื่องและสั้นไหลด้วยการประยุกต์ใช้ Hierarchical Neural Network โดยมีชุดข้อมูลที่ใช้ในการวิจัยเป็นงานเขียนโดยมนุษย์จำนวน 300,000 เรื่องจากเว็บไซต์ Hierarchical โมเดลในการสร้างเรื่องราวจะทำการสร้างหัวเรื่องก่อน

แล้วนำหัวเรื่องที่ได้นำไปสร้างเรื่องราวที่มีความเกี่ยวข้องกับหัวเรื่อง

งานวิจัยนี้ทำให้เรื่องราวที่สร้างมีความต่อเนื่องโดยการใช้พัวชนโมเดล และเพิ่ม gated multi-scale self-attention mechanism เพื่อให้สร้างเนื้อเรื่องที่ยาวได้ ใช้ทั้งมนุษย์และโปรแกรมที่สร้างมาวัดผลเรื่องราวที่ได้ ผลการวิจัยมนุษย์ที่เป็นผู้ตัดสินชอบเรื่องราวที่สร้างโดย Hierarchical โมเดล มากกว่าเรื่องราวที่สร้างโดยโมเดลที่เป็น Non Hierarchical

### 4.2 Convolutional Sequence to Sequence Learning [11]

งานวิจัยนี้เกี่ยวกับการทำ Neural machine translation โดยใช้ CNN แทนที่ RNN ด้วยข้อเสียของการใช้ RNN คือ การที่จะทำกระบวนการ NMT ได้ จะต้องนำประโยคมาใส่ในเวกเตอร์ที่กำหนดขนาดไว้ตายตัว แม้ว่าขนาดของประโยคจะสั้น-ยาวแค่ไหนขนาดของเวกเตอร์ก็ไม่มีการเปลี่ยนแปลง และ RNN มีรูปแบบการทำงานเป็นลำดับ การจะเปลี่ยนรูปแบบการทำงานให้เป็นแบบขนานนั้นทำได้โดยยาก งานวิจัยนี้ได้นำโมเดลรูปแบบ CNN มาใช้ในกระบวนการ NMT โดยใช้งานคู่กับ Self Attention และ Gated Linear Units

### 4.3 Attention Is All You Need [12]

โมเดลที่เป็นที่นิยมดั้งเดิมในการทำ Sequence Modeling คือ recurrent neural network หรือ Convolutional ที่ซับซ้อน และภายในมี encoder และ decoder โมเดลที่มีประสิทธิภาพมากที่สุดก็เป็นโมเดลที่ต่อ encoder และ decoder ผ่าน Attention mechanisms งานวิจัยนี้ได้นำเสนอสถาปัตยกรรม neural network แบบใหม่ Transformer Model ที่เป็น Attention mechanisms เพียงอย่างเดียว

### 4.4 Language Modeling with Gated Convolutional Networks [13]

งานวิจัยนี้เป็นการทำ Language Modeling โดยใช้วิธีที่เรียกว่า Gated linear unit โดยจะทำการแบ่ง Convolutional Layer ออกเป็นสองส่วน จากนั้นครึ่งหนึ่งจะถูกส่งเข้า Sigmoid Function เพื่อใช้กำหนดว่า ส่วนไหนบ้างในอีกครึ่งหนึ่งจะถูกเลือกไปใช้ หากข้อมูลในครั้งที่เข้า Sigmoid Function มีค่ามาก การกำหนดส่วนของอีกครึ่งก็จะมากตาม ขั้นตอนเหล่านี้นำมาใช้ในการเลือกข้อมูลที่เป็นส่วนสำคัญ

## 5. ขอบเขตของโครงการ (Scope)

- 5.1 สร้างโมเดลการเรียนรู้เชิงลึกที่สามารถสร้างเรื่องสั้นออกมาเป็นข้อความในภาษาอังกฤษจาก  
เซตข้อมูลเรื่องสั้นที่มีผู้แต่งไว้แล้ว มีเรื่องราวที่ต่อเนื่อง และเกี่ยวข้องกับหัวข้อเรื่องที่ระบุไว้ โดย  
มนุษย์สามารถอ่านแล้วทำความเข้าใจได้
- 5.2 เรื่องสั้นที่จะเอามาใส่ในเซตข้อมูลควรมีจำนวนตัวอักษรอย่างน้อย 30 คำ ในหนึ่งเรื่อง
- 5.3 เนื่องจากการใช้การเรียนรู้เชิงลึกในการสร้างงานเขียนออกมา  
จำเป็นจะต้องใช้ชุดข้อมูลที่เป็นเรื่องสั้น หรือเรื่องแต่งโดยมนุษย์  
ในปริมาณที่เพียงพอเพื่อให้ได้งานเขียนที่สามารถอ่านแล้วเข้าใจได้
- 5.4 ผลลัพธ์ที่ได้จะอยู่ในรูปของโมเดล ไม่ใช่ซอฟต์แวร์  
ทั้งนี้เพื่อให้งานมีประเด็นอยู่ในเรื่องการเรียนรู้เชิงลึก และการใช้เทคนิคต่างๆเพื่อปรับปรุงผลลัพธ์  
เป็นหลัก

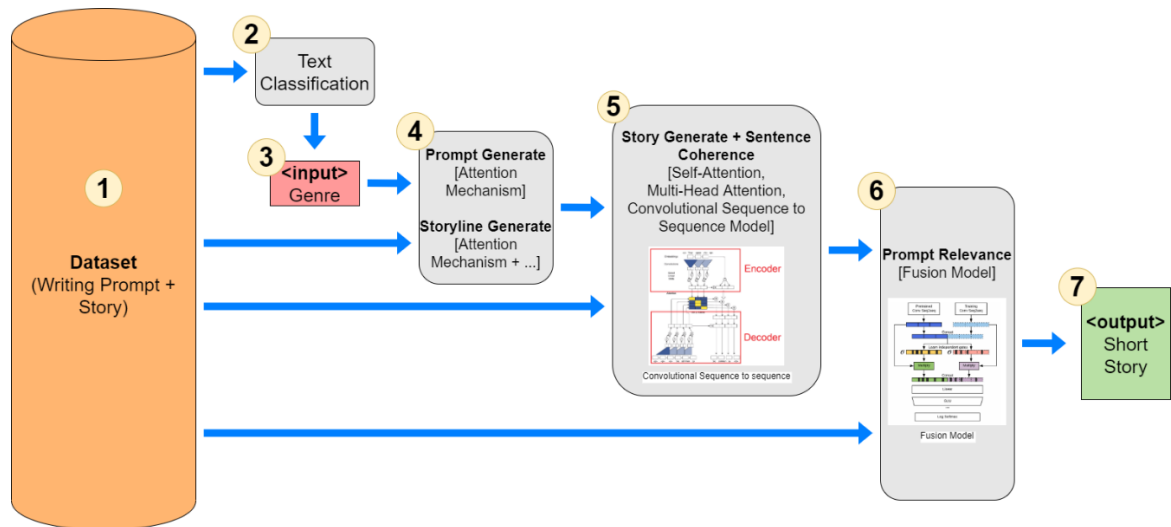
## 6. การพัฒนาโครงการ (Project Development)

### 6.1 ขั้นตอนการพัฒนา (Methodology)

- 6.1.1. ตั้งคำถามที่สนใจ กำหนดเป้าหมาย กำหนดขอบเขตของโครงการ และกำหนดวัตถุประสงค์
- 6.1.2. วางแผนการดำเนินงาน
- 6.1.3. แบ่งการดำเนินงานออกเป็น 7 ส่วนหลักๆ ดังนี้
  - 6.1.3.1. Dataset
  - 6.1.3.2. Text Classification
  - 6.1.3.3. Input
  - 6.1.3.4. Prompt Generate + Storyline Generate
  - 6.1.3.5. Story Generate + Sentence Coherence
  - 6.1.3.6. Prompt Relevance
  - 6.1.3.7. Output
- 6.1.4. ศึกษาและพัฒนาในแต่ละส่วน
- 6.1.5. ทดสอบการทำงานของระบบ



## 6.2 การออกแบบ (Design)



รูป 6.2 Concept Learning Structure

### 6.2.1. Dataset

เป็น text ตัวอย่าง Writing Prompt และ Story จำนวนมาก

### 6.2.2. Text Classification

ใช้ Machine learning หา Genre จาก Dataset ทำโดยการ ตัดคำศัพท์ในแต่ละบทความทำเป็น bag-of-word แล้วหา TF-IDF TF(Term Frequency) คือ ค่าความถี่ของคำที่ปรากฏในบทความนั้น ๆ แล้วนำ TF-IDF ไปทำ Classification โดยใช้ Machine Learning Algorithm เช่น Stochastic gradient descent, Multinomial Naive Bayes

### 6.2.3. Input

เป็น Genre รูปแบบของ Genre จะเป็นคำ ๆเดียว ใช้สำหรับกำหนดประเภทของเรื่องราวที่จะ Generate ออกมา โดยจะใช้ในการ Generate Prompt เพื่อนำไปใช้ต่อไป

### 6.2.4. Prompt Generate + Storyline Generate

Prompt Generate ใช้ Attention Mechanism สร้าง Prompt จาก Writing Prompt ใน Dataset และจาก Genre ที่รับมา รูปแบบของ Prompt จะเป็นข้อความสั้น บอกถึงบทที่จะนำไปใช้สร้างเรื่องราวต่อไป

Storyline Generate ใช้ Attention Mechanism ร่วมกับเทคนิคอื่น ๆ เพื่อสร้างข้อความที่มีรูปแบบเหมือนกับ Prompt แต่มีการเปลี่ยนไปของเรื่องราว เพื่อให้เกิดการดำเนินเรื่อง จะนำไปใช้ใน Story Generate

### 6.2.5. Story Generate + Sentence Coherence

Story Generate ทำหน้าที่ในการสร้างประโยคที่เป็นข้อความตามรูปแบบการแต่งเรื่องสั้น

Sentence Coherence ทำหน้าที่ควบคุมข้อความที่จะสร้างขึ้น ให้มีความเกี่ยวข้องกับข้อความก่อนหน้า สร้างไว้แล้ว ในที่นี้จะใช้ Convolutional Sequence to sequence เป็นวิธีการที่ผสมผสานหลายวิธีเข้าด้วยกัน และถูกใช้ในการทำ Story Generate ด้วย

### 6.2.6. Prompt Relevance

ทำหน้าที่ควบคุมการสร้างประโยคให้เป็นไปตาม Genre ที่ป้อนเข้ามา ถ้าหากไม่ควบคุมความสัมพันธ์ระหว่างประโยคที่จะสร้าง กับ Prompt ที่รับเข้ามา จะทำให้ผลลัพธ์ที่ได้ เป็นกลุ่มของประโยคที่ไม่มีความสัมพันธ์กัน การควบคุมจะใช้ Pretrained Model รูปแบบคล้ายคลึงกันสร้างจาก Dataset แต่ไม่สนใจ Prompt มาเทียบกัน และสร้าง Learning Gate มาเรียนรู้ความแตกต่างเพื่อเลือกส่วนสำคัญที่มี Prompt มาแสดงผล

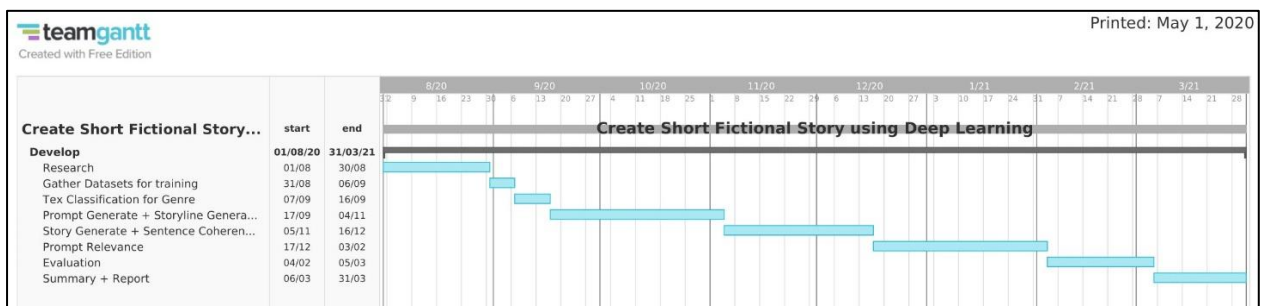
### 6.2.7. Output

เป็น text เรื่องสั้นที่มี Genre ตามที่กรอกเข้าไป เป็นเรื่องที่สามารถอ่านเข้าใจ และมีการดำเนินเรื่องเป็นเรื่องราว

## 6.3 แนวทางการทดสอบและการวัดประสิทธิภาพ (Test and Performance Evaluation Approaches )

แนวทางการทดสอบและวัดผลประสิทธิภาพ จะทำได้โดยใช้ Metrix วัดผลเป็น Perplexity และวัดผลด้วยการทดสอบจากคนจริง ๆ เพื่อให้ได้ความคิดเห็นแล้วนำไปปรับปรุงผลลัพธ์ให้ดีขึ้น

## 7. แผนการดำเนินโครงการ (Timeline)



รูป 7 แผนการดำเนินโครงการ

## 8. ประโยชน์ที่คาดว่าจะได้รับ (Expected Benefits)

ในปัจจุบันเทคโนโลยีปัญญาประดิษฐ์ การเรียนรู้ของเครื่อง และการเรียนรู้เชิงลึก ถูกใช้ในหลากหลายวัตถุประสงค์ ส่วนใหญ่เป็นการนำมาใช้แก้ปัญหาที่มีความซับซ้อนเป็นอย่างมากเช่นปัญหาจราจร และการพยากรณ์สภาพอากาศ การนำเทคโนโลยีเหล่านี้มาใช้กับงานที่อาศัยจินตนาการและความคิดสร้างสรรค์ ก็ไม่สามารถออกมาสู่งานที่มนุษย์ทำจริงๆได้ ผลงานที่ออกมาจะขาดความเป็นธรรมชาติและมีข้อจำกัดที่มากมายเช่น ไม่มีความสัมพันธ์กันในงาน

โครงการนี้จะจัดทำขึ้นเพื่อมุ่งเน้นที่จะศึกษาวิธีการเรียนรู้เชิงลึกควบคู่กับเทคนิคต่างๆมาสร้างงานเขียนทางวรรณกรรมประเภทเรื่องสั้น เพื่อให้ชิ้นงานที่ได้มีความสัมพันธ์กัน มีการดำเนินเรื่องอ่านแล้วเข้าใจได้ หลังจากที่ได้โครงการนี้ได้จัดทำตามแผนงานจนเสร็จสมบูรณ์ จะสามารถนำไปใช้พัฒนาต่อยอดในด้านอื่นๆได้อีก และสามารถใช้เป็นแนวทางในการศึกษาวิจัยสำหรับผู้สนใจหาความรู้ทางด้านนี้

## 9. เอกสารอ้างอิง (Reference)

[1] โครงข่ายประสาทเทียม. Wikipedia.,

<https://th.wikipedia.org/wiki/โครงข่ายประสาทเทียม>

[2] ทำความรู้จักกับโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึก. Phyblas.hinaboshi Blog.,

<https://phyblas.hinaboshi.com/umaki00>

[3] การประมวลผลภาษาธรรมชาติ. Wikipedia.,

<https://th.wikipedia.org/wiki/การประมวลผลภาษาธรรมชาติ>

[4] The Neural Engineer. Workshop : ทำโปรเจกต์ Deep NLP. ThAIKeras.com.,

<https://thaikeras.com/2019/workshop-nlp-quora/>

[5] Keng Surapong. สอนสร้าง RNN ถึง GRU ด้วยภาษา Python – NLP ep.9. BUA Labs.,

<https://www.bualabs.com/archives/3103/what-is-rnn-recurrent-neural-network-what-is-gru-gated-recurrent-unit-teach-how-to-build-rnn-gru-with-python-nlp-ep-9/>

[6][10] Angela Fan, Mike Lewis, Yann Dauphin.(2018). Hierarchical Neural Story Generation. Association for Computational Linguistics.,

<https://www.aclweb.org/anthology/P18-1082/>

[7] Raimi Karim. Illustrated: Self-Attention. Towards Data Science.,

<https://towardsdatascience.com/illustrated-self-attention-2d627e33b20a>

[8] Elvis. Detecting Emotions with CNN Fusion Models. Medium.,

<https://medium.com/dair-ai/detecting-emotions-with-cnn-fusion-models-b066944969c8>

[9] Seq2seq. Wikipedia.,

<https://en.wikipedia.org/wiki/Seq2seq>

[11] Jonas Gehring, Michael Auli, David Grangier, Denis Yarats, Yann N. Dauphin. (2017). Convolutional Sequence to Sequence Learning. Proceedings of the 34th International Conference on Machine Learning.,

<https://arxiv.org/pdf/1705.03122.pdf>

[12] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. (2017). Attention Is All You Need. Cornell University.

<https://arxiv.org/abs/1706.03762>

[13] Yann N. Dauphin, Angela Fan, Michael Auli, David Grangier. (2016). Language Modeling with Gated Convolutional Networks. Cornell University.

<https://arxiv.org/abs/1612.08083>

[14] Jonas Gehring, Michael Auli, David Grangier, Denis Yarats, Yann N. Dauphin. (2017). Convolutional Sequence to Sequence Learning. Cornell University.

<https://arxiv.org/abs/1705.03122>