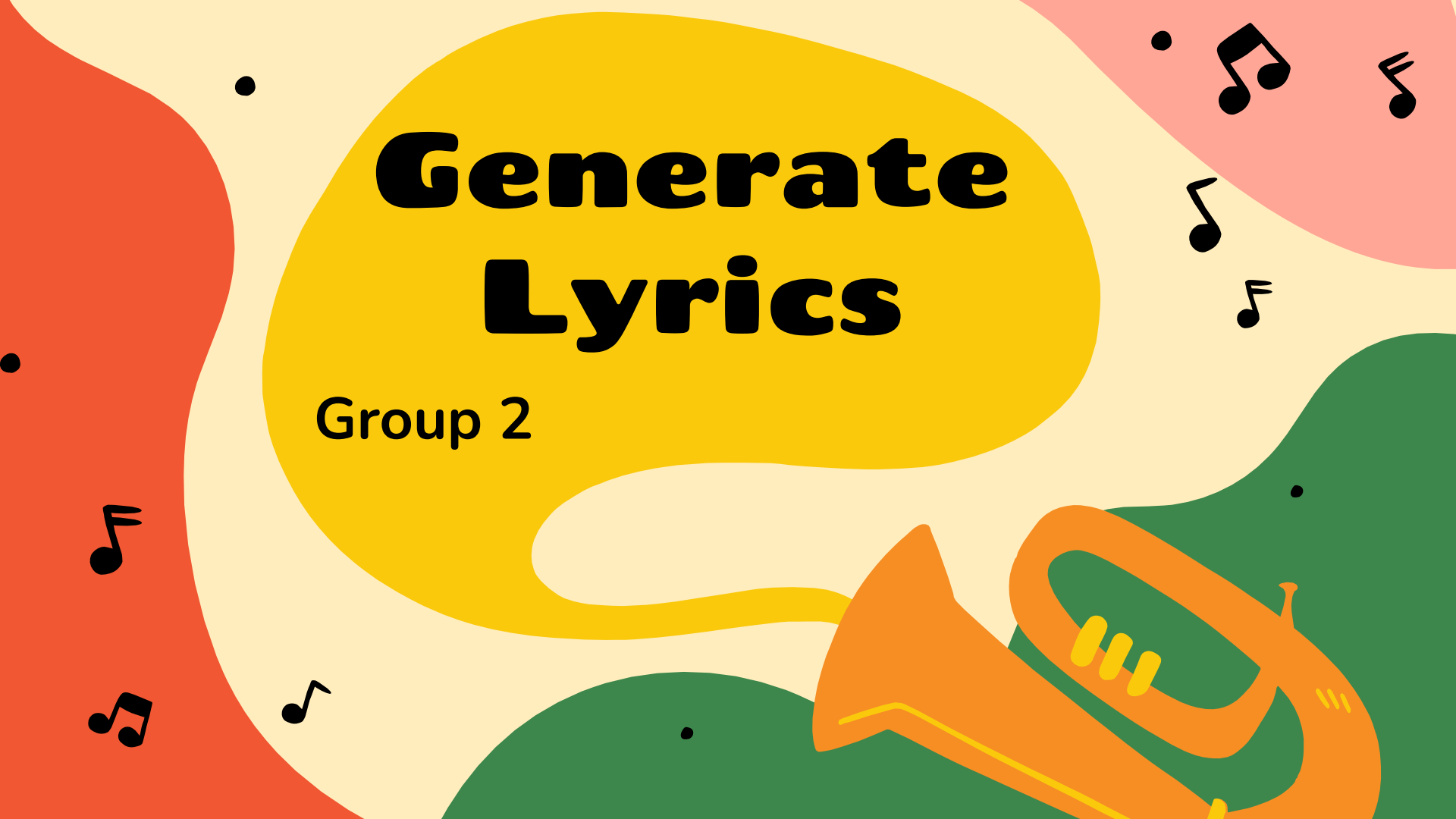


# **Generate Lyrics**

Group 2





# TABLE OF CONTENTS

**1**

## **INTRODUCTION**

- Goal
- Dataset
- Challenge

**2**

## **DATA PREPARATION**

- Clean data
- Tokenizing text

**3**

## **MODEL**

- GRN
- Result

**4**

## **Result**

# INTRODUCTION

## Goal

แต่งเพลงใหม่ โดยใช้ Recurrent Neural Network (RNN), GRU

## Dataset

เนื้อเพลงภาษาไทย 1400 เพลง โดยรวบรวมจากศิลปินไทยที่ขึ้น  
ชอบ 44 วง  
จาก เว็บไซต์ <https://xn--72c9bva0i.meemodel.com/>

## Challenge

สร้างเพลงใหม่จากภาษาไทย



# Data Preparation



## Web Scraping

Collect lyrics from 44 artists in  
[https://xn--72c9bva0i.meemodel.co  
m/](https://xn--72c9bva0i.meemodel.com/)



## Cleaning data

Remove marks,  
numbers, and irrelevant  
words.



## Tokenize Text

word\_tokenize  
by use pythainlp

# Web Scraping

## Python Package and Library

- BeautifulSoup
- Requests
- Pandas

เพลง			
เนื้อเพลง :: ศิลปิน >> 25 Hours			
ค้นหาเพลง/ศิลปิน			
ค้นหา			
รายชื่อเพลงของ 25 Hours			
f แชรส์ t แชรส์			
📌 เรียงตามความนิยม 📌 เรียงตามตัวอักษร			
ชื่อเพลง	ศิลปิน	อัลบั้ม	Tube
Amazing 🎧	25 Hours		
Come on 🎧	25 Hours	Color In White	
Have you here (happy new year)	25 Hours		
Midnight serenade 🎧	25 Hours	Color In White	
Wow!	25 Hours	เพลงประกอบโฆษณา Honda	
กลับมาเหมือนเดิม 🎧	25 Hours		
คนข้างข้าง 🎧	25 Hours	Color In White	
คำตอบ 🎧	25 Hours	25 Hours	
คำตอบที่ใช่ 🎧	25 Hours		
คิดเหมือนกันหรือเปล่า 🎧	25 Hours	Color In White	
คืนวางคืน เพลงวางเพลง	25 Hours		

▼	name	▼	link	▼	lyrics
0	Amazing		https://xn--72c9bva0i.meemodel.com/เนื้อเพลง/Amazing_25%20Hours		วันนี้เรื่องราวที่เคยได้เจอกันในความฝัน จะเป็นจริงทุกอย่าง เธอจะมีทุกอย่าง ดอกไม้กำลังร้องเพลงเพื่อมอบให้เธอกับฉัน คือเสียงเพลงที่เราเก็บมาน
					* it's amazing it's amazing โลกแห่งความฝันอยู่ใกล้ ( มีทุกสิ่งพร้อมจะมอบให้เธอ ) it's amazing it's amazing แต่เธอพร้อมเดินไปด้วยกัน สิ่งที่เราฝันนั้นกำลังรอ เป็นของเธอ
					ตรงนี้มีแสงงดงามเหมือนธรรมชาติสร้างสรรค์ มีแต่ความพิเศษ รอให้เธอพิสูจน์ เมื่อสายลมที่พัดไหวชวนให้ฉันไม่เดินล่า อากาศให้เธอได้เดินไปด้วยกัน
					it's amazing it's amazing โลกแห่งความฝันอยู่ใกล้ ( มีทุกสิ่งพร้อมจะมอบให้เธอ ) it's amazing it's amazing แต่เธอพร้อมเดินไปด้วยกัน สิ่งที่เราฝันนั้นกำลังรอ
					จะพาเธอไปยังที่แห่งนั้น สิ่งที่เราฝันจะไม่เป็นแค่ฝัน
					it's amazing it's amazing โลกแห่งความฝันอยู่ใกล้ ( มีทุกสิ่งพร้อมจะมอบให้เธอ ) it's amazing it's amazing แต่เธอพร้อมเดินไปด้วยกัน
					( ข้า * )

# Cleaning data

## ตัวอย่าง

Remove special characters  
(such as ',', '(', ')', '[', ']' etc)

เมื่อฉันและเธอนั้นได้รักกัน  
เธอเป็นคนดี เป็นคนที่ดี  
และฉันโชคดีจริงๆ  
ถ้าวันหนึ่งเธอนั้นแก่ (เธอนั้นแก่)  
ส่งมือเหี่ยวๆ มาให้ฉัน (มาให้ฉัน)

\*\*ก็เพราะฉันยังหยุดใจไม่ไหว หยุดคงเหมือนตาย



เมื่อฉันและเธอนั้นได้รักกัน  
เธอเป็นคนดี เป็นคนที่ดี  
และฉันโชคดีจริงๆ  
ถ้าวันหนึ่งเธอนั้นแก่ เธอนั้นแก่  
ส่งมือเหี่ยวๆ มาให้ฉัน มาให้ฉัน

ก็เพราะฉันยังหยุดใจไม่ไหว หยุดคงเหมือนตาย

# Word Tokenization

## Python Package and Library

```
import pythainlp  
from pythainlp import word_tokenize
```

## ตัวอย่าง

วันนี้เรื่องราวที่เคยได้เจอกันในความฝัน  
จะเป็นจริงทุกอย่าง เธอจะมีทุกอย่าง

word\_tokenize



['วันนี้', 'เรื่องราว', 'ที่', 'เคย', 'ได้', 'เจอกัน', 'ใน', 'ความฝัน',  
,  
'จะ', 'เป็นจริง', 'ทุกอย่าง', 'เธอ', 'จะ', 'มี', 'ทุกอย่าง']

# Train Model

Create Training and Targets



Create Model



Train Model

```
# The maximum length sentence we want for a single input in words
seqLength = 8
examples_per_epoch = len(corpusList)//(seqLength + 1) # number of seqLength+1 sequences in the corp

# Create training / targets batches
wordDataset = tf.data.Dataset.from_tensor_slices(word_as_int)
print(wordDataset)
sequencesOfWords = wordDataset.batch(seqLength + 1, drop_remainder=True) # generating batches of 8
print(sequencesOfWords)

def split_input_target(chunk): # This is where right shift happens
    input_text = chunk[:-1]
    print(input_text)
    target_text = chunk[1:]
    print(target_text)
    return input_text, target_text # returns training and target sequence for each batch

dataset = sequencesOfWords.map(split_input_target) # dataset now contains a training and a target :
```



# Train Model

Create Training and Targets batch



Create Model



Train Model

```
#define BATCH_SIZE and BUFFER_SIZE
BATCH_SIZE = 64 # each batch contains 64 sequences. Each sequence contains 8 words (seqLength)
BUFFER_SIZE = 100 # Number of batches that will be processed concurrently
dataset = dataset.shuffle(BUFFER_SIZE).batch(BATCH_SIZE, drop_remainder=True)
```

```
# Length of the vocabulary in words
vocab_size = len(vocab)
# The embedding dimension
embedding_dim = 256
# Number of GRU units
rnn_units = 1024
```

```
def createModel(vocab_size, embedding_dim, rnn_units, batch_size):
    model = tf.keras.Sequential([
        tf.keras.layers.Embedding(vocab_size, embedding_dim,
                                   batch_input_shape=[batch_size, None]),
        tf.keras.layers.GRU(rnn_units,
                             return_sequences=True,
                             stateful=True,
                             recurrent_initializer='glorot_uniform'),
        tf.keras.layers.Dense(vocab_size)
    ])
    return model

model = createModel(vocab_size = len(vocab), embedding_dim=embedding_dim, rnn_units=rnn_units, batch_size=BATCH_SIZE)
```

# Train Model

Create Training and Targets batch



Create Model



Train Model

```
#Train Model
EPOCHS = 20
history = model.fit(dataset, epochs=EPOCHS, callbacks=[checkpoint_callback])
tf.train.latest_checkpoint(checkpoint_dir)
model = createModel(len(vocab), embedding_dim, rnn_units, batch_size=1)

model.load_weights(tf.train.latest_checkpoint(checkpoint_dir))

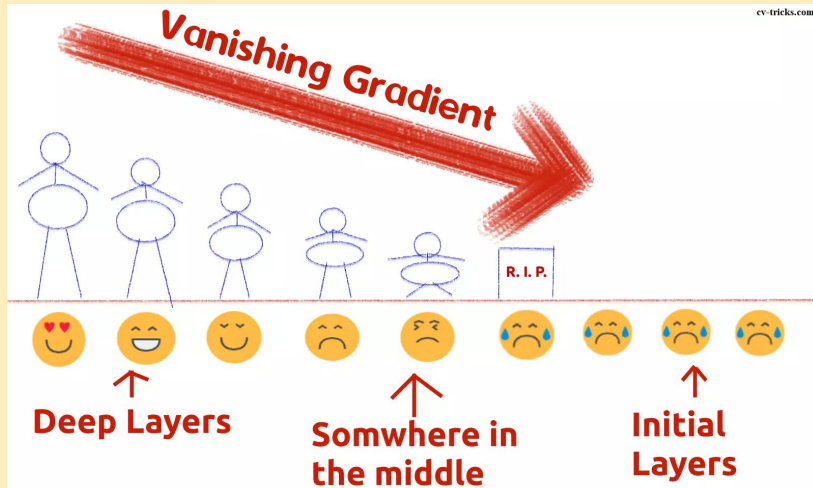
model.build(tf.TensorShape([1, None]))
model.summary()
```



**3**

**Model**

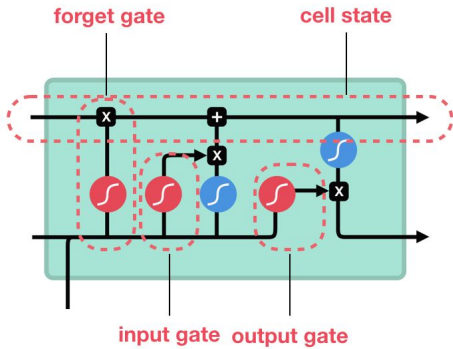
# Why Gated Recurrent Units(GRU) ?



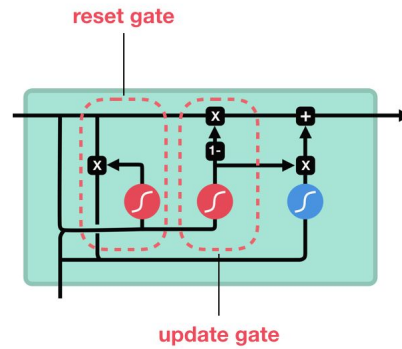
- Solve the vanishing/Exploding gradient problem of RNNs

# Why Gated Recurrent Units(GRU) ?

LSTM



GRU



sigmoid



tanh



pointwise  
multiplication



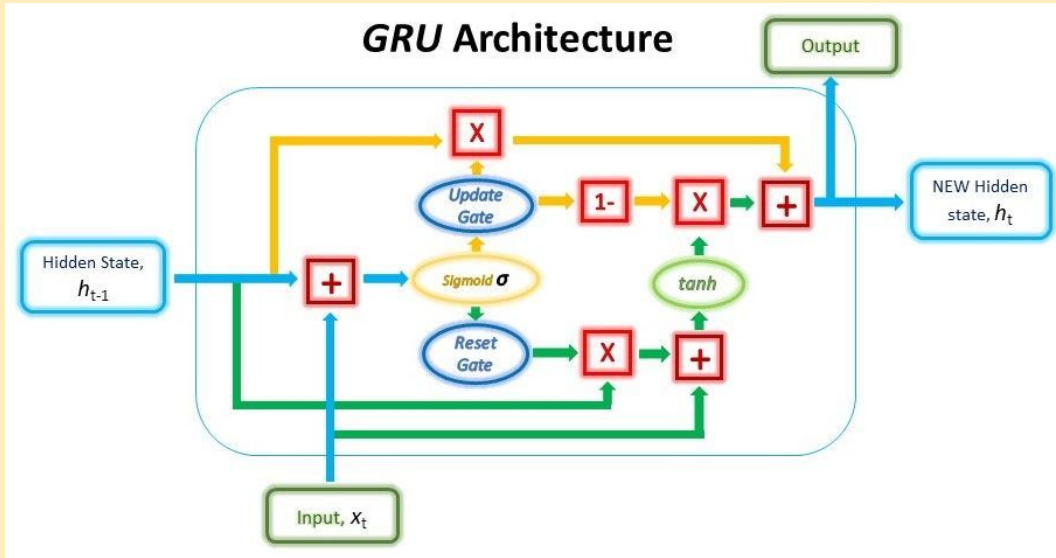
pointwise  
addition



vector  
concatenation

- GRU does not possess any internal memory, they don't have an output gate that is present in LSTM
- GRU's has fewer tensor operations ,they are a little speedier to train then LSTM

# Gated Recurrent Unit (GRU)

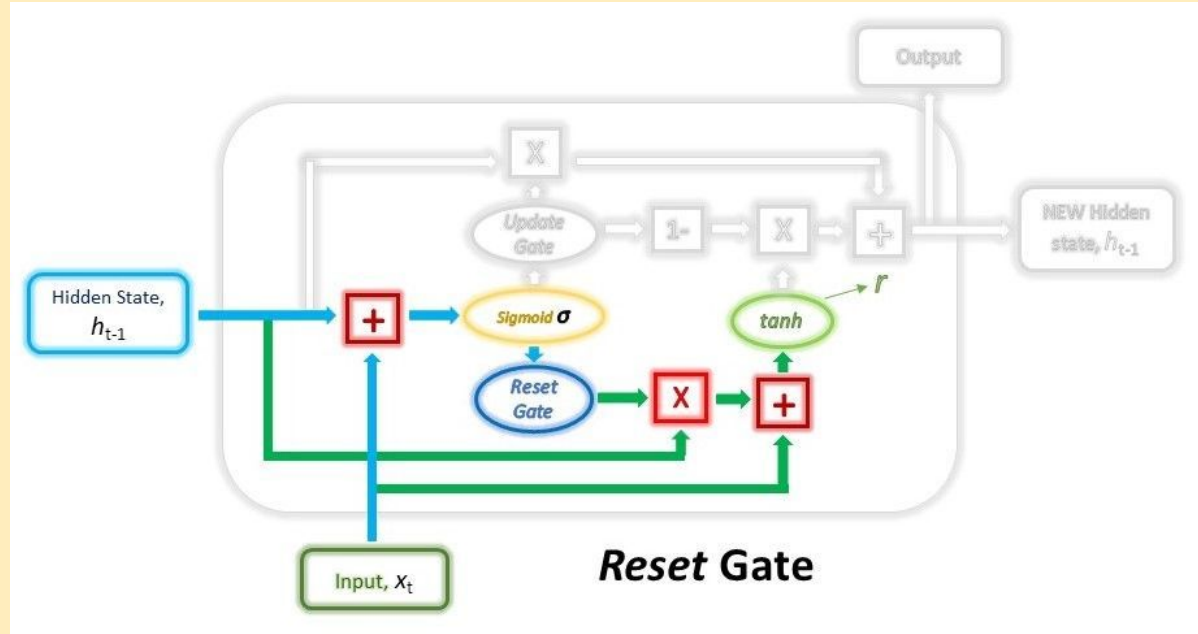


The GRU cell contains only two gates:

- The Update gate
- The Reset gate

# Gated Recurrent Unit (GRU)

The Reset gate

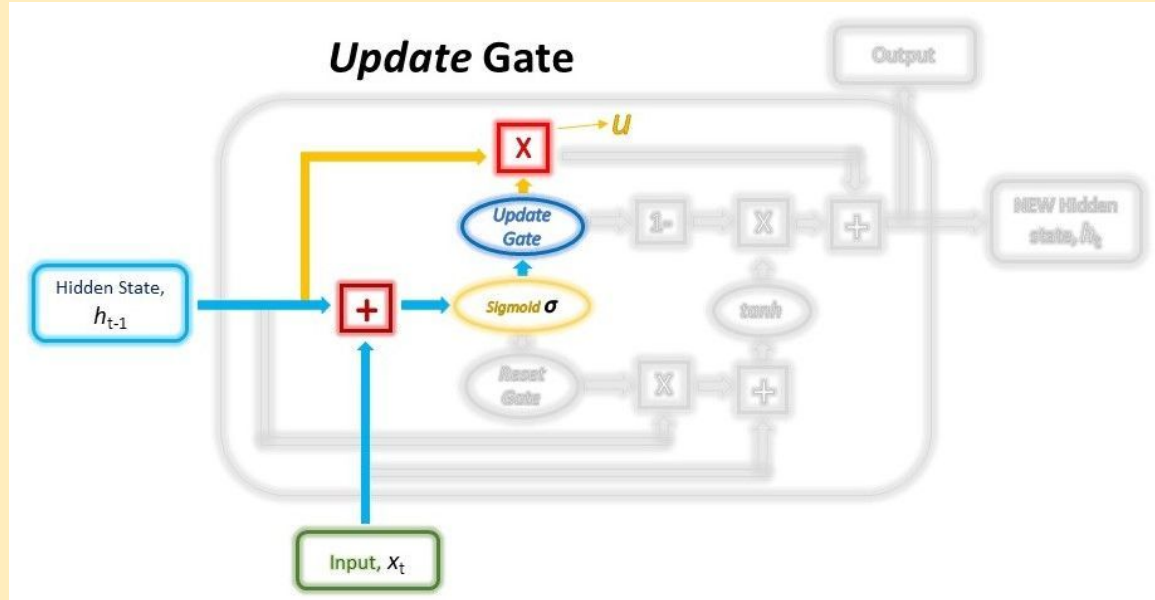


$$gate_{reset} = \sigma(W_{input_{reset}} \cdot x_t + W_{hidden_{reset}} \cdot h_{t-1})$$

$$r = \tanh(gate_{reset} \odot (W_{h_1} \cdot h_{t-1}) + W_{x_1} \cdot x_t)$$

# Gated Recurrent Unit (GRU)

The Update gate



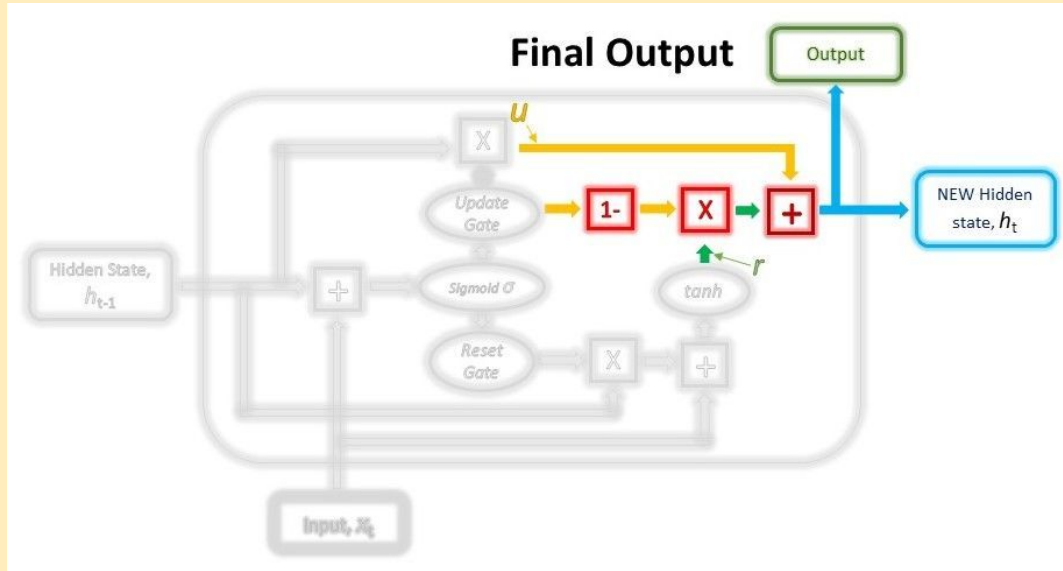
$$gate_{update} = \sigma(W_{input_{update}} \cdot x_t + W_{hidden_{update}} \cdot h_{t-1})$$

$$u = gate_{update} \odot h_{t-1}$$



# Gated Recurrent Unit (GRU)

Combining the output



# Result

คำเริ่มต้น : มอง

มอง ตา กัน ไป มอง ตา กัน มา อีก นาน  
อีก นาน เท่าไหร่ ต้อง จม อยู่ กับ ดวงตา  
ที่ น้ำตา ยังคง ริน ไหล เหตุใด เธอ ต้อง  
ไป แต่ ฉัน ต้อง ทำ

มองตากันไป มองตากันมา

อีกนาน อีกนาน เท่าไหร่ต้องจมอยู่

กับดวงตาที่น้ำตายังคงรินไหล

เหตุใดเธอต้องไปแต่ฉันต้องทำ



# Result

คำเริ่มต้น : มือ

มือ ไม่ ลั่น ไม่ ว่า รัก นี้ จะ หยุด รัก เธอ  
เท่ากับ ใจ ร้อง ด้วย รัก เรา ได้ แต่ เธอ  
อยู่ กับ ฉัน จมอยู่กับ ชีวิต อยู่ กับ ความ  
ทรงจำ ที่ คอย ตอกย้ำ

มือไม่ลั่น ไม่ว่ารักนี้

จะหยุดรักเธอ เท่ากับใจร้องด้วยรักเรา

ได้แต่เธออยู่กับฉัน จมอยู่กับชีวิต

อยู่กับความทรงจำ ที่คอยตอกย้ำ



# Result

คำเริ่มต้น : ใคร

ใครอื่นใด ยากนัก ใซ้ใหม่ ให้ รัก ฉัน เพียง คนเดียว ที่ ใจ  
ฉัน ตามหา จาก วันนี้ ยังอีก ไกล มัน คง ไม่ เหมือน เก่า ที่  
เรา คอย กัน เหมือน เป็น เพราะฉัน ไม่ มีสิทธิ์ ลิขิต ดี ชั่ว ต่ำ  
สูง เลว ทราม งดงาม แล้วแต่ ใจ

คิด ชีวิต ของ เรา ไม่ ใซ้ ของ เขา โปรด ฟัง อีกครั้ง เรา ก็ ยัง  
รัก กัน จวบ จนถึง วัน ตาย คำ ว่า รัก มัน เป็น อะไร ไม่ ใซ้  
เธอ เป็น อะไร ไม่ ใซ้ เลย อยาก จะ ไป เรื่อยเปื่อย ฉัน  
เหนื่อย ที่จะ เข้าใจ ที่ เธอ ไม่ นำ มา พบ มา เจอ มา ช้า ต้อง  
พบ และ แม่ จะ นาน

ใครอื่นใด ยากนัก ใซ้ใหม่ ให้ รัก ฉัน เพียง คนเดียว

ที่ ใจ ฉัน ตามหา จาก วันนี้ ยังอีก ไกล

มัน คง ไม่ เหมือน เก่า ที่ เรา คอย กัน เหมือน เป็น

เพราะฉัน ไม่ มีสิทธิ์ ลิขิต ดี ชั่ว

ต่ำ สูง เลว ทราม งดงาม แล้วแต่ ใจ คิด

ชีวิต ของ เรา ไม่ ใซ้ ของ เขา โปรด ฟัง อีกครั้ง

เราก็ ยัง รัก กัน จวบ จนถึง วัน ตาย คำ ว่า รัก มัน เป็น  
อะไร

ไม่ ใซ้ เธอ เป็น อะไร ไม่ ใซ้ เลย อยาก จะ ไป เรื่อยเปื่อย

ฉัน เหนื่อย ที่จะ เข้าใจ ที่ เธอ ไม่ นำ มา พบ มา เจอ

มา ช้า ต้อง พบ และ แม่ จะ นาน




# Result

คำเริ่มต้น : โลก กลม

โลก กลม ๆ เธอ กับ ฉัน พบกัน ที่ไหน ถ้าหาก ความ  
รัก เป็น เช่นไร โว้วว โฮ่ว ฉัน ต้อง เสีย ใจมา แค่  
เพียง วูบ เดียว มา วูบ เดียว แค่ จาก ผ่าน เข้ามา  
า แล้ว ทำให้ น้ำตา ริน ไหล ไม่ รู้ว่า นาน เท่าไร ที่  
จะ หลุดพ้น กับ ความรัก ที่ ทำให้ หมองหม่น  
เหลือเกิน แต่ ผ่าน มา เพียง

โลกกลม ๆ เธอ กับ ฉัน พบกัน ที่ไหน  
ถ้าหากความรักเป็นเช่นไร โว้วว โฮ่ว  
ฉันต้องเสียใจมาแค่เพียงวูบเดียว  
มาวูบเดียวแค่จากผ่านเข้ามา  
แล้วทำให้น้ำตารินไหล  
ไม่รู้ว่านานเท่าไรที่จะหลุดพ้น  
กับความรักที่ทำให้หมองหม่นเหลือเกิน  
แต่ผ่านมาเพียง



# Result

คำเริ่มต้น : คอย กับ เธอ

คอย กับ เธอ ก็ ไม่ กลัว เชื่อใจ ว่า เรา จะมี แค่  
เพียง ใจถึง เธอ จะเป็น กำลังใจ ให้ ใจ เธอ  
ได้ รู้ ว่า เป็น เรื่องใหญ่ มี ฉัน อยู่ใน ใจ ฉัน  
ยังมี เธอ ที่ ยังคง งดงาม ฉัน ยังคง รัก เธอ  
แม้ รู้ ว่า เธอ นั้น อยู่ ตรงนี้ อยู่ กับ ความทรง  
จำ ที่ คอย

คอยกับเธอก็ไม่กลัว

เชื่อใจว่าเราจะมีแค่เพียงใจถึงเธอ

จะเป็นกำลังใจให้ใจเธอ ได้รู้ว่าเป็นเรื่องใหญ่

มีฉันอยู่ในใจ ฉันยังมีเธอที่ยังคงงดงาม

ฉันยังคงรักเธอแม้รู้ว่า

เธอนั้นอยู่ตรงนี้อยู่กับความทรงจำที่คอย



Thank  
you

## สมาชิกกลุ่ม

1. กชกร เลื่อนสุขสันต์ รหัสนักศึกษา 6310412001
2. สลิลวสุ เทียงธรรม รหัสนักศึกษา 6310412017
3. สุพัตรา ตั้งสกุลระหง รหัสนักศึกษา 6310412028