Deep Learning Final Project

108753206李鈺祥、109753106王中伶

CONTENTS

01 Introduction

02 Method

03 Result

Introduction

People or Machine

- Daily Meditation Tranquility
- Daily Meditation: Light Up Your Life
- Daily Meditations, Prayers and Mindfulness Exercises
- Daily Meditation: The Temple

Written by Human

02

Method

Dataset of Generation

HuffPost News

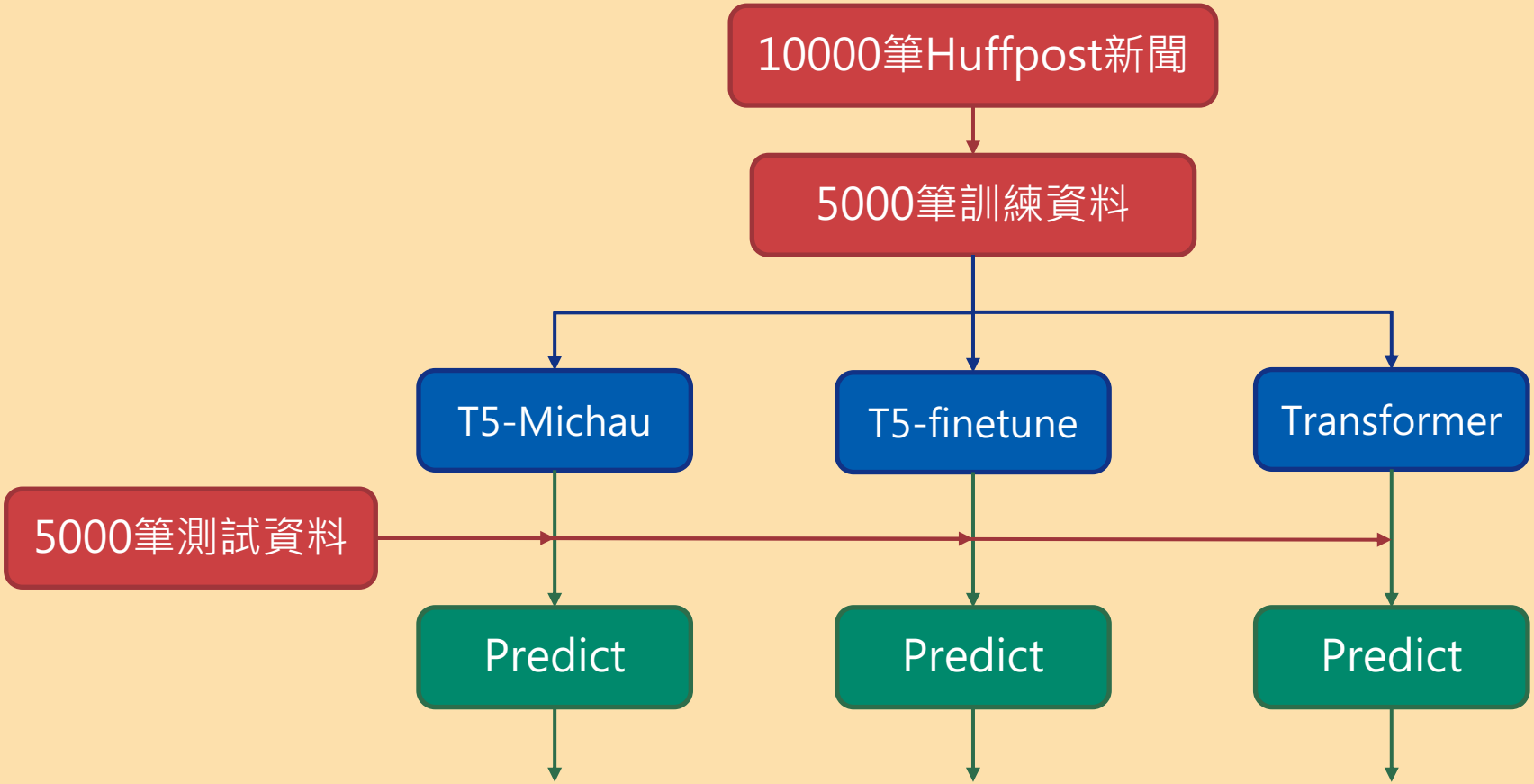
Training Data

- 5000筆155字以內之新聞內文
- 5000筆上述內文之對應標題

Testing Data

- 5000筆155字以內之新聞內文

Framework of Generation



Dataset of Classification

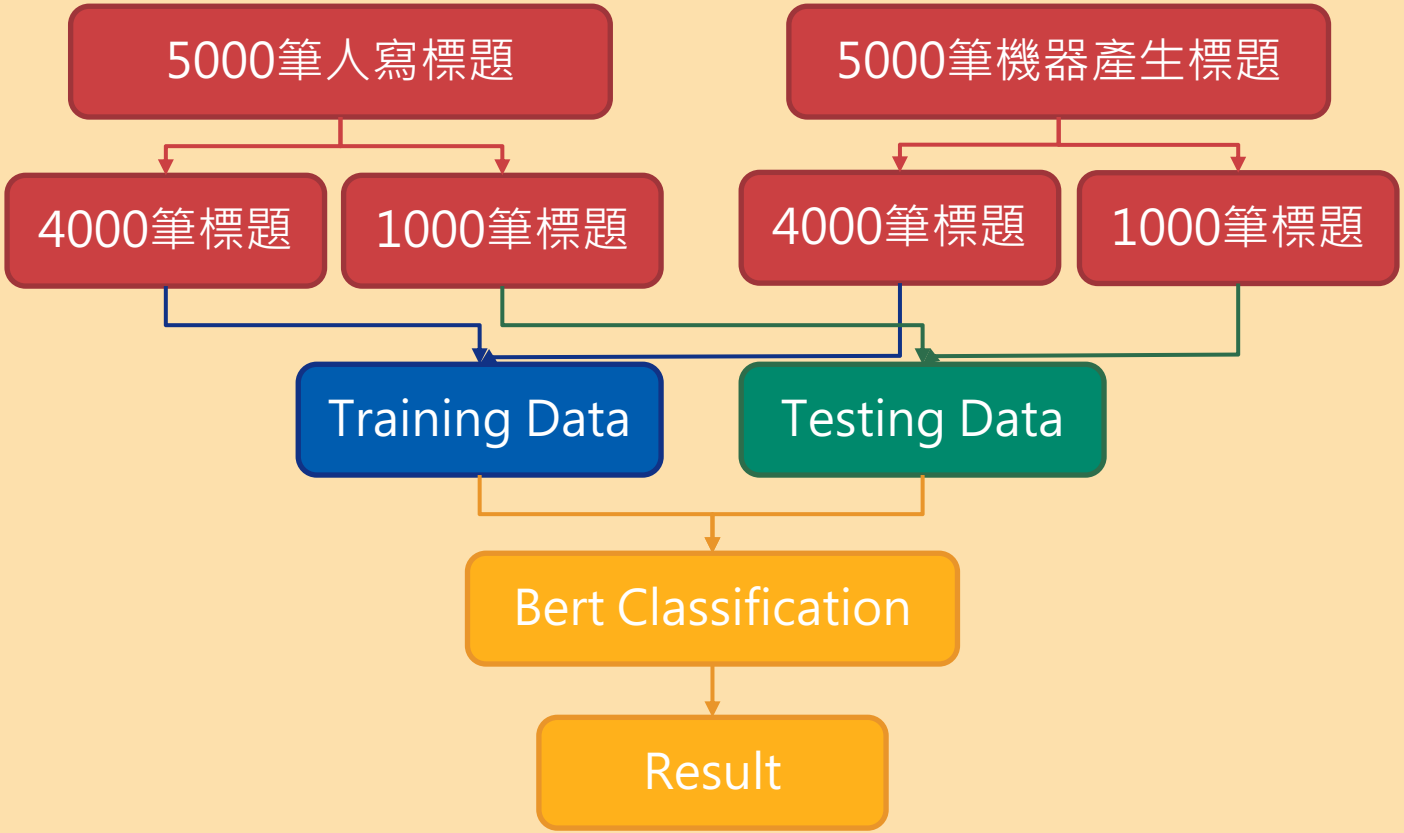
Training Data

- 4000 筆人所寫之標題
- 4000 筆機器產生之標題

Testing Data

- 1000 筆人所寫之標題
- 1000 筆機器產生之標題

Framework of Classification



T5

- T5是Google於2019年釋出的預訓練模型，比2017年釋出的BERT
  - https://arxiv.org/abs/1910.10683
  - 擁有更多單詞量與參數量
  - 更彈性的自然語言形式的輸入輸出框架
  - 對於生成類的任務相當強悍
  - 依據任務前綴詞，可使用同個模型，支援不同任務
  - 對於各種自然語言處理的相關任務，幾乎都可以勝任

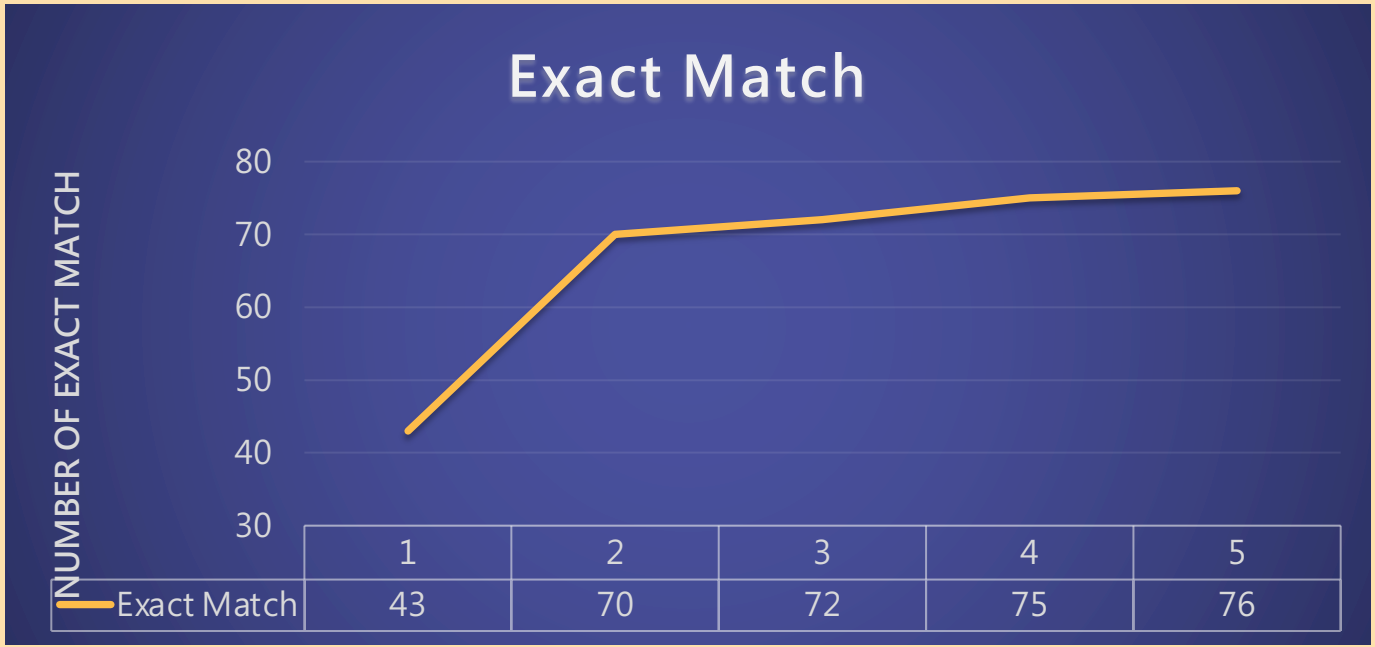
T5-Michau

- **Michau/t5-base-en-generate-headline**
  - https://huggingface.co/Michau/t5-base-en-generate-headline

T5-Fine-tune

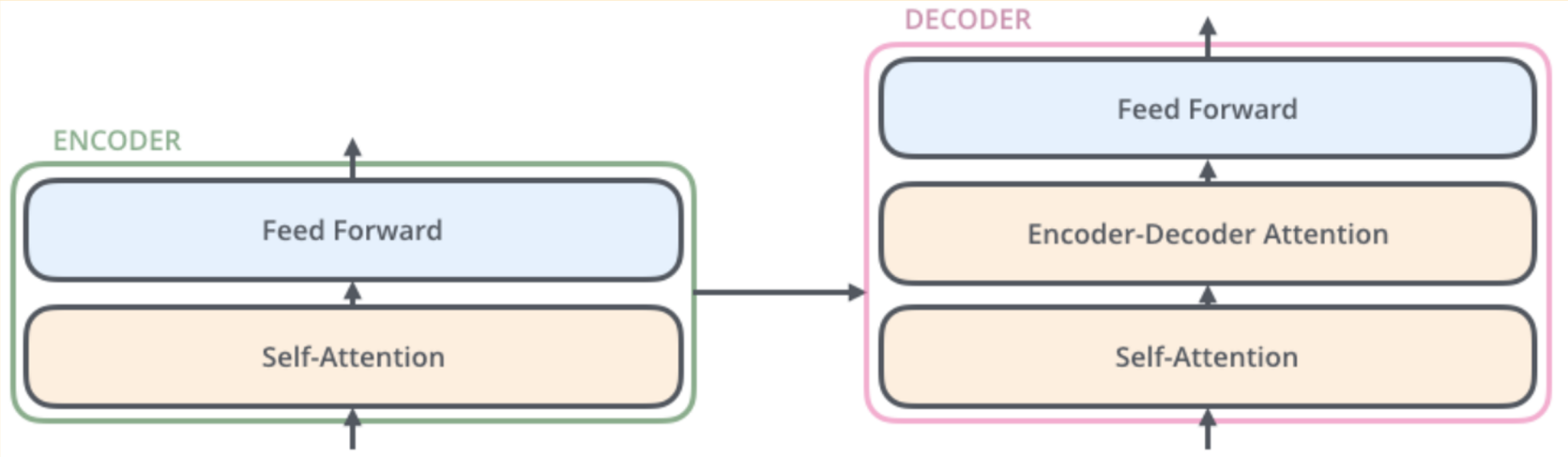
- SQuAD（Stanford Question Answering Dataset）任務
- Exact Match常為評估方式之一
- 評估5個Epoch的Exact Match

Exact Match on Different Epochs



Transformer Summarizer

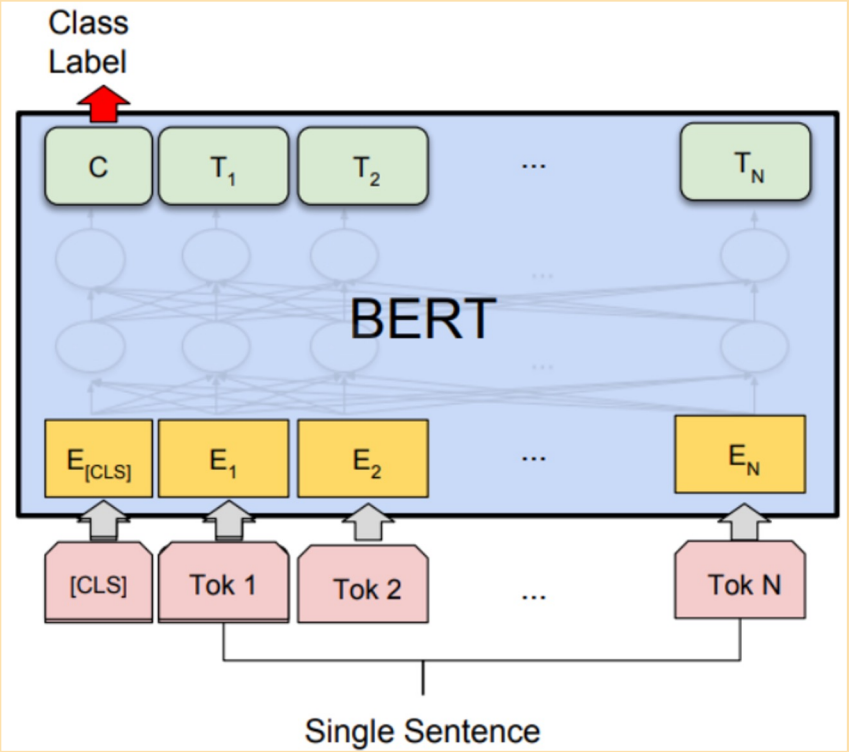
- 資料格式：Tuple→Iterator
- Seq2Seq + Attention
- Epoch: 5



圖片來源：http://jalammar.github.io/illustrated-transformer/

Bert Classification

- Pre-trained: bert-large-uncased
- 0為人寫標題；1為機器產生標題



圖片來源：https://www.geeksforgeeks.org/sentiment-classification-using-bert/

03

Result

On race relations, President Obama is feeling optimistic. At least, that's how he comes across in an interview with NPR's Steve Inskeep, who asks if "the United States is more racially divided than it was" when he took office. "No," Obama says, "I actually think that it's probably in its day-to-day interactions less racially divided."

Generation

Trump White House Leadership	Transformer
Obama Says The United States Is 'Probably Less Racially Divided'	T5-fine-tune
President Obama: Is America More Racially Divided Than It Was When He Takes Office?	T5-Michau
Why I Am Optimistic About The Future Of Race Relations In America	Written by Human

Classification

F1-score

85.67%

T5-Michau

63.99%

T5-Fine-tune

87.99%

Transformer

Classification

False Negative

10.40%

T5-Michau

35.50%

T5-Fine-tune

13.20%

Transformer

Reference

- Headliner.

  https://as-ideas.github.io/headliner/
- T5-Michau. https://huggingface.co/Michau/t5-base-en-generate-headline

Thanks for listening