Machine Learning assisted Cantopop Lyric Composition

LYU2101

Tang Ka Lok, 1155125745

01

Nature of Cantonese and Cantopop

Cantonese is a tone language

- A specific tone representing a unique pitch associated to a Chinese Character
- The pitch is used to distinguish between different words and meanings

Tonal System of Cantonese

MARS MERCURY VENUS JUPITER

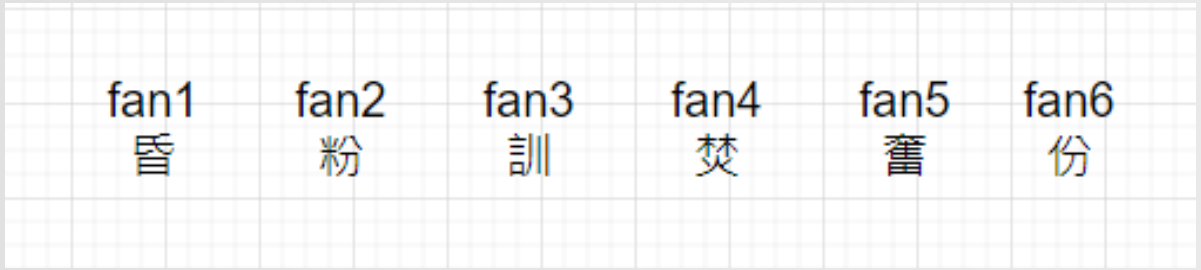


Figure 1.1: 6 characters having same syllable with 6 different tones

- Each Chinese Character is combining the syllable ‘fan’ and tone from 1 to 6 (Jyutping representation).

- 昏: dizzy
- 粉: pink
- 訓: train
- 焚: burn
- 奮: exert
- 份: portion

Tone-melody Matching Mechanism in Cantopop

- Melody contains the sequence of music notes with different pitch height.
- Matching the tone of lyrics to the melody of the song.
- The song is basically not understandable if the mechanism is not followed

Tone-melody Matching Mechanism

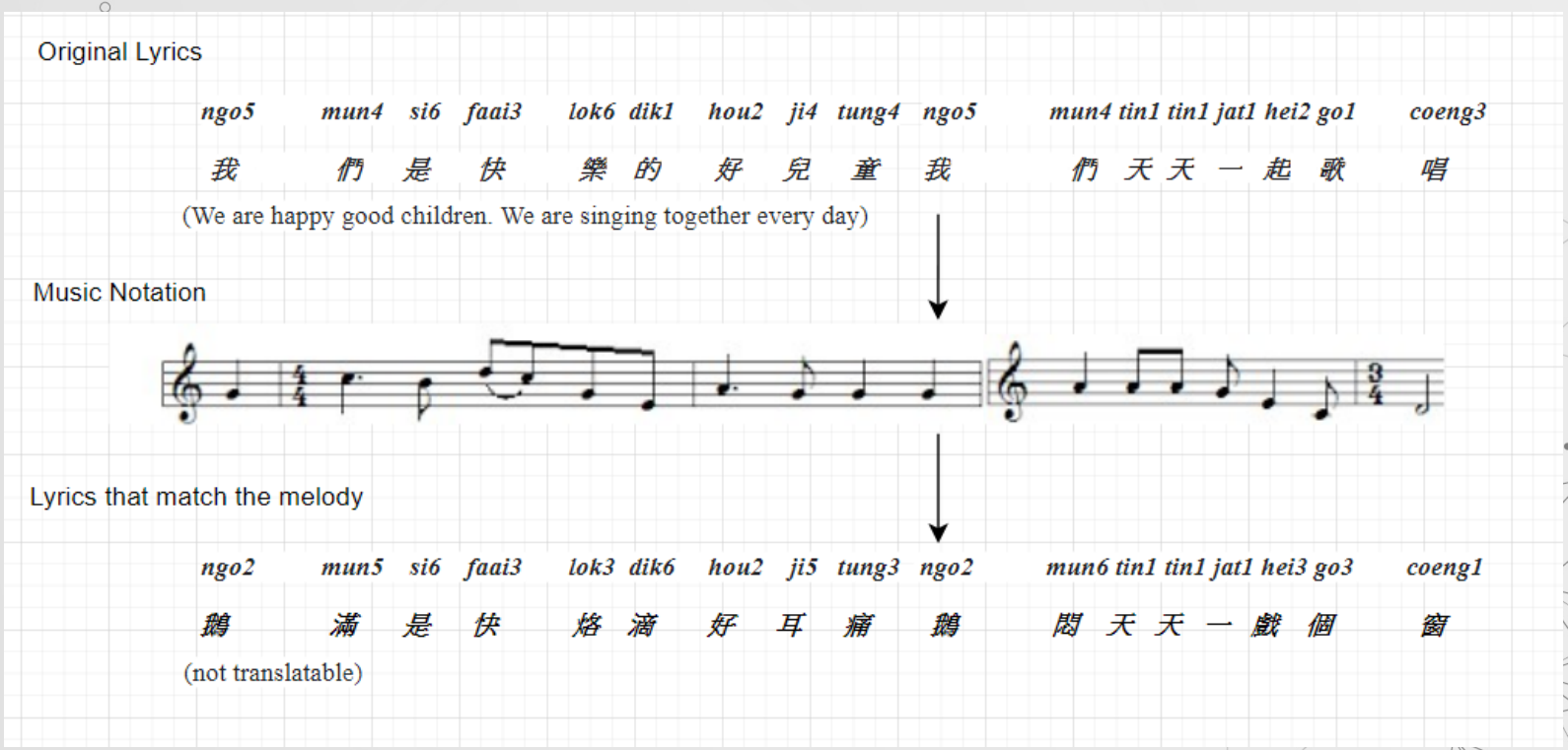


Figure 1.2 Example of fitting lyrics with unmatched tone into the melody

Tone-melody Matching Mechanism

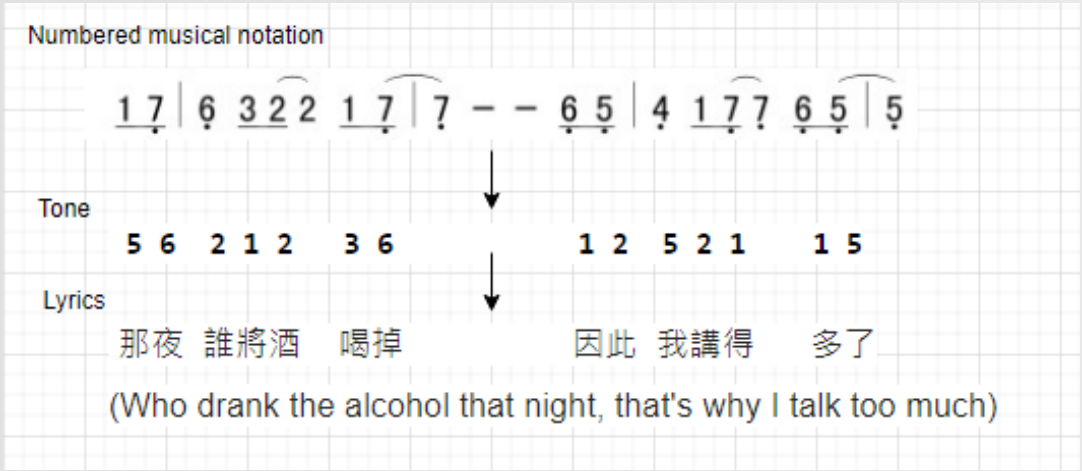


Figure 1.3 Example of the steps to write Cantopop lyrics

Steps to write Cantopop Lyrics

1. Convert the melody to the tones of Cantonese based on the pitch of the music notes.

2. Fill in the lyrics that match the tones.

02

Model Description

Transformer: Self-attention Mechanism

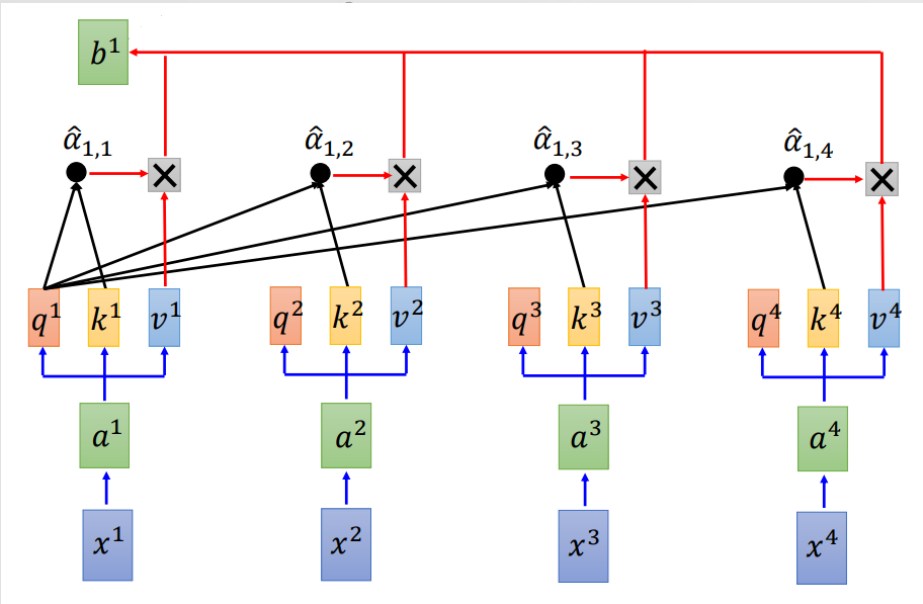


Figure 2.1 Self-Attention mechanism[1]

- With an input sequence passed into self-attention layer, the attention score between each input is calculated.
- With the attention calculated, the contextual information between inputs are captured and learned by the model.

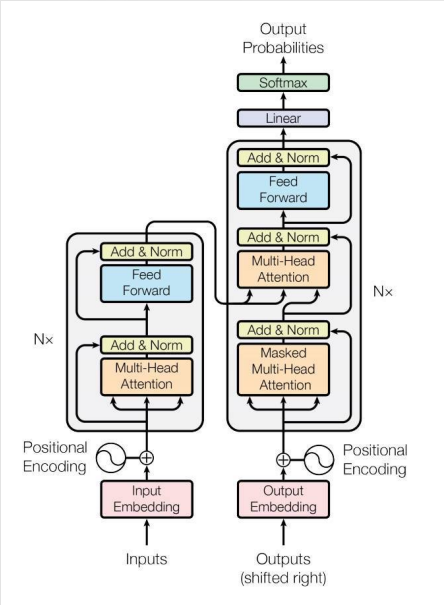


Figure 2.2 Transformer Architecture [2]

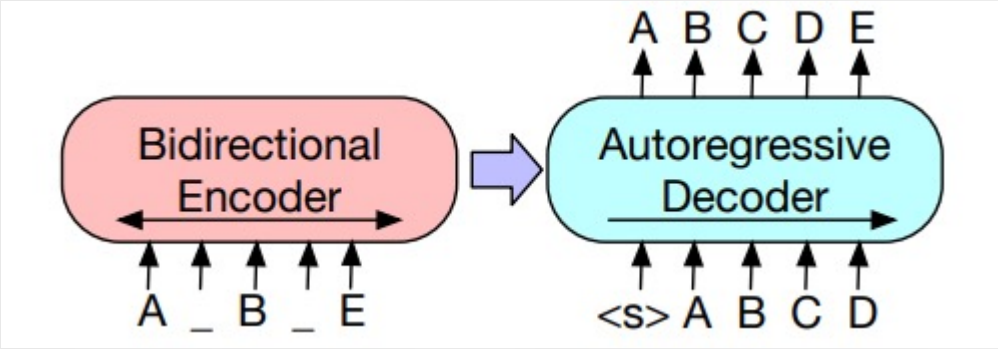


Figure 2.4 Bart [3]

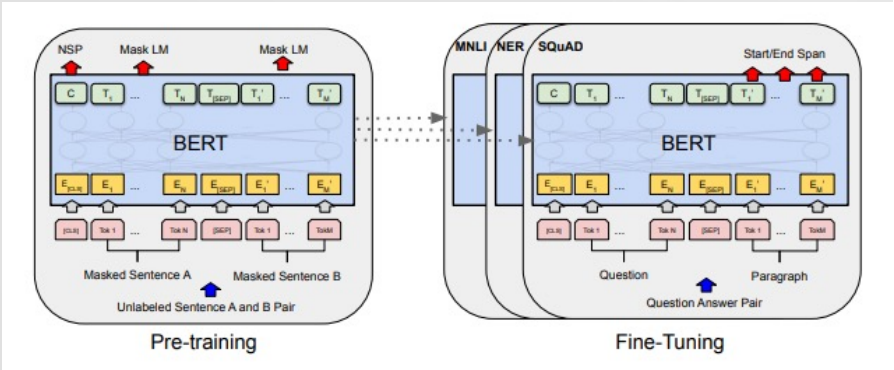


Figure 2.3 Bert [4]

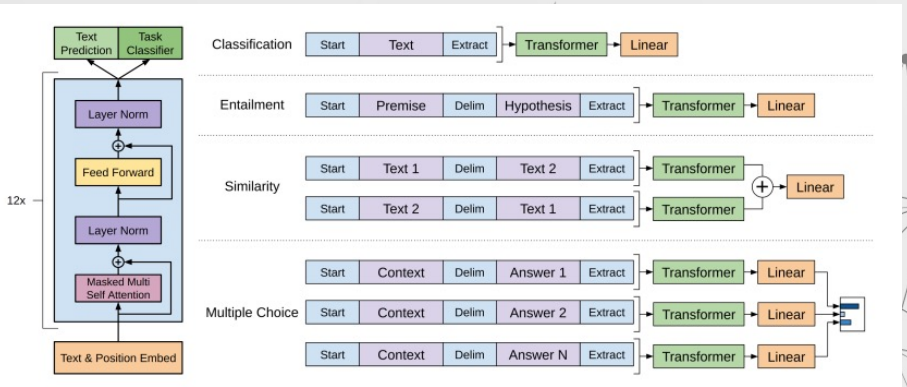


Figure 2.4 GPT [5]

03

Methodology

Related Works

Given the previous words input, do the next word prediction to generate lyrics

Word-prediction-based Lyrics Generation

Takes the melody as input and generate corresponding lyrics as output. It captures the relationship between melody and lyrics.

Melody-based Lyrics Generation

Recap: Cantopop Lyrics Composition

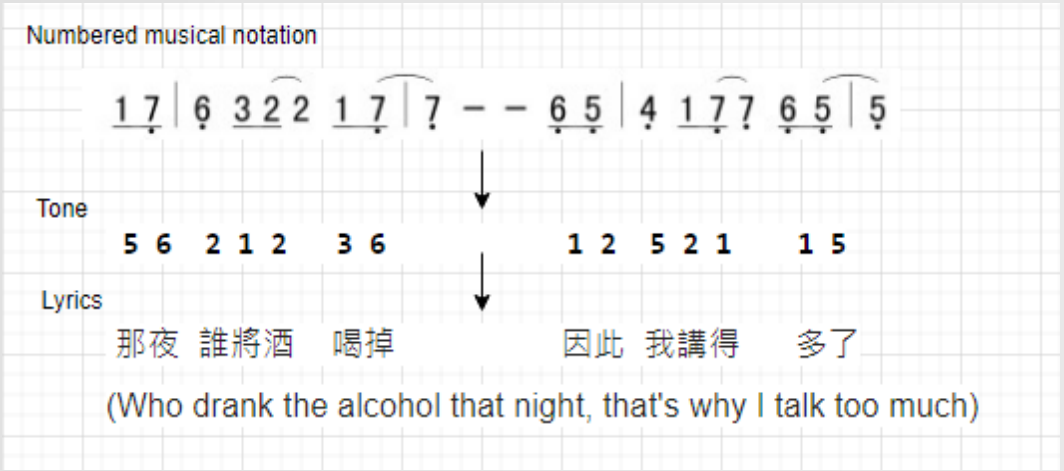


Figure 3.1 Example of the steps to write Cantopop lyrics

Steps to write Cantopop Lyrics

1. Convert the melody to the tones of Cantonese based on the pitch of the music notes. (Easy)

2. Fill in the lyrics that match the tones. (Extremely Hard)

Tone-based Lyrics Generation

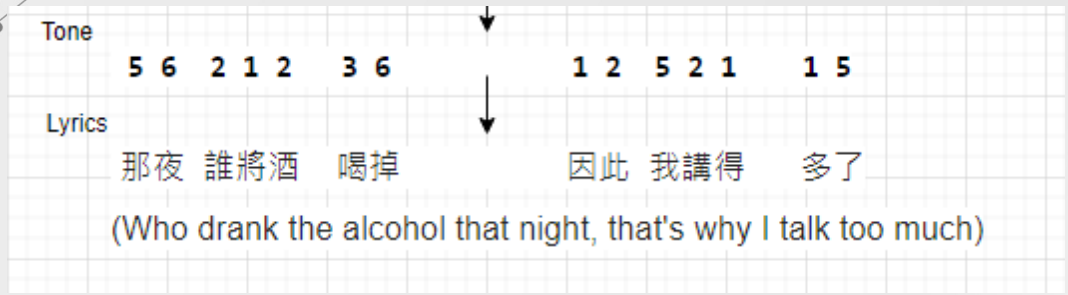


Figure 3.2 Illustrating Tone-based Lyrics Generation

- Skip the first step
- Build a model takes tones as input and generate lyrics that match the tones

Advantages

Easier Data Preparation

- The tone of word can be extracted from the lyrics.
- Difficulties on getting data are reduced while maintaining the nature of Cantopop lyrics composition.

Expand Size of Dataset

- Dataset can be expanded from Cantopop lyrics data to all Chinese pop lyrics data.
- Melody and lyrics are indirectly related to each other
- Greatly increase the diversity of the model.

Training Approaches

Directly train the model using the target data

Training from Scratch

First pre-train a model using in-domain data and then we fine-tune it to our specific task using the target data

Pre-training and Fine-tuning

04

Workflows

Data Preparation

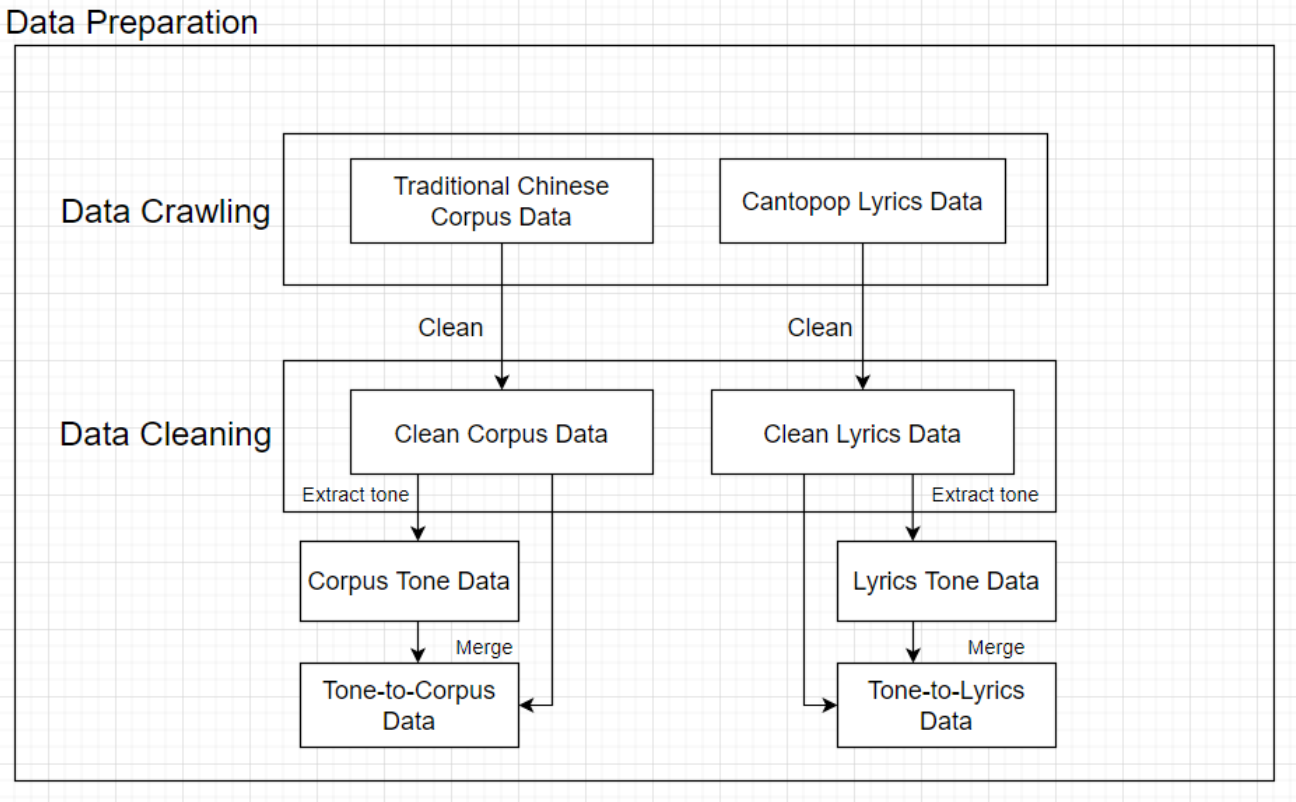


Figure 4.1 Flow chart of data preparation phase

Dataset Overview

10 GB — Traditional Chinese Corpus Data

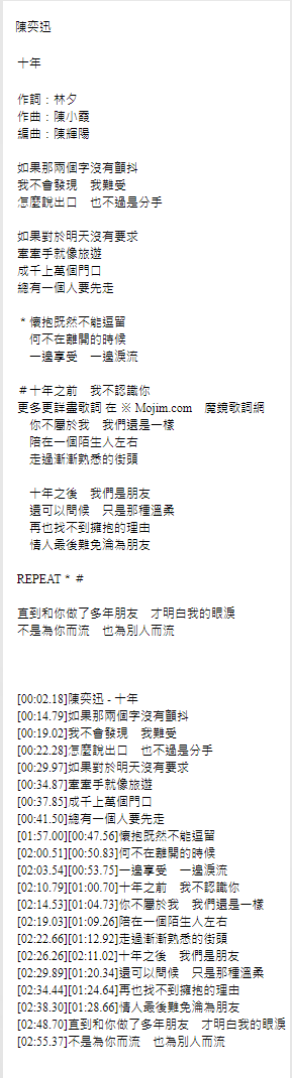
200 MB — Tone-to-Corpus Data

15000 — Lyrics / Tone-to-Lyrics Data

MARS MERCURY VENUS

Example of Data Cleaning

Before cleaning:



陳奕迅

十年

作詞：林夕
作曲：陳小霞
編曲：陳輝陽

如果那兩個字沒有顫抖
我不會發現　我難受
怎麼說出口　也不過是分手

如果對於明天沒有要求
牽牽手就像旅遊
成千上萬個門口
總有一個人要先走

* 懷抱既然不能逗留
何不在離開的時候
一邊享受　一邊淚流

十年之前　我不認識你
更多更詳盡歌詞 在 ※ Mojim.com　魔鏡歌詞網
你不屬於我　我們還是一樣
陪在一個陌生人左右
走過漸漸熟悉的街頭

十年之後　我們是朋友
還可以問候　只是那種溫柔
再也找不到擁抱的理由
情人最後難免淪為朋友

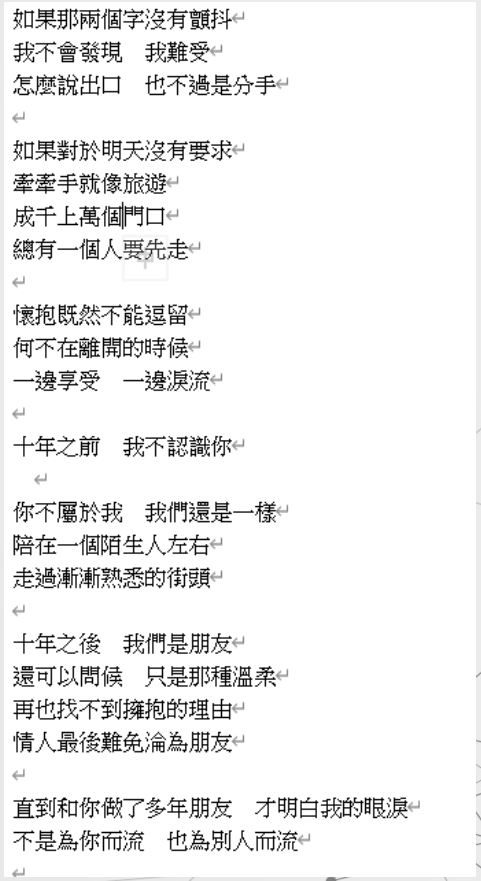
REPEAT * #

直到和你做了多年朋友　才明白我的眼淚
不是為你而流　也為別人而流

[00:02.18]陳奕迅 - 十年
[00:14.79]如果那兩個字沒有顫抖
[00:19.02]我不會發現　我難受
[00:22.28]怎麼說出口　也不過是分手
[00:29.97]如果對於明天沒有要求
[00:34.87]牽牽手就像旅遊
[00:37.85]成千上萬個門口
[00:41.50]總有一個人要先走
[01:57.00][00:47.56]懷抱既然不能逗留
[02:00.51][00:50.83]何不在離開的時候
[02:03.54][00:53.75]一邊享受　一邊淚流
[02:10.79][01:00.70]十年之前　我不認識你
[02:14.53][01:04.73]你不屬於我　我們還是一樣
[02:19.03][01:09.26]陪在一個陌生人左右
[02:22.66][01:12.92]走過漸漸熟悉的街頭
[02:26.26][02:11.02]十年之後　我們是朋友
[02:29.89][01:20.34]還可以問候　只是那種溫柔
[02:34.44][01:24.64]再也找不到擁抱的理由
[02:38.30][01:28.66]情人最後難免淪為朋友
[02:48.70]直到和你做了多年朋友　才明白我的眼淚
[02:55.37]不是為你而流　也為別人而流

Figure 4.2 Lyrics before cleaning

After cleaning:



如果那兩個字沒有顫抖
我不會發現　我難受
怎麼說出口　也不過是分手

如果對於明天沒有要求
牽牽手就像旅遊
成千上萬個門口
總有一個人要先走

懷抱既然不能逗留
何不在離開的時候
一邊享受　一邊淚流

十年之前　我不認識你

你不屬於我　我們還是一樣
陪在一個陌生人左右
走過漸漸熟悉的街頭

十年之後　我們是朋友
還可以問候　只是那種溫柔
再也找不到擁抱的理由
情人最後難免淪為朋友

直到和你做了多年朋友　才明白我的眼淚
不是為你而流　也為別人而流

Figure 4.3 Lyrics After cleaning

Example of Tone-to-Text Data

Original Lyrics:

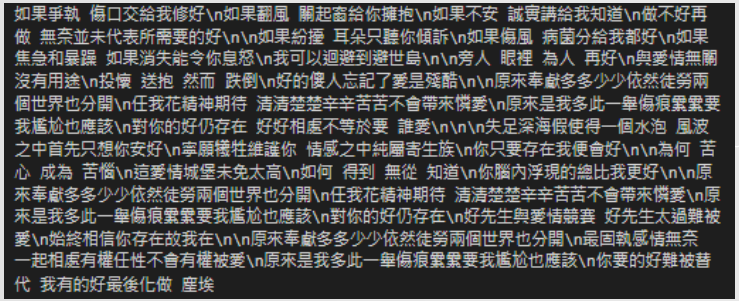


Figure 4.4 Original Lyrics Data

Extracted Tones:

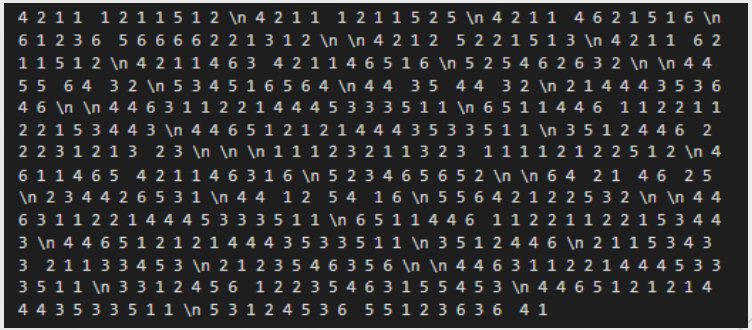


Figure 4.5 Extracted Tone Data

Tone-to-Text Data:

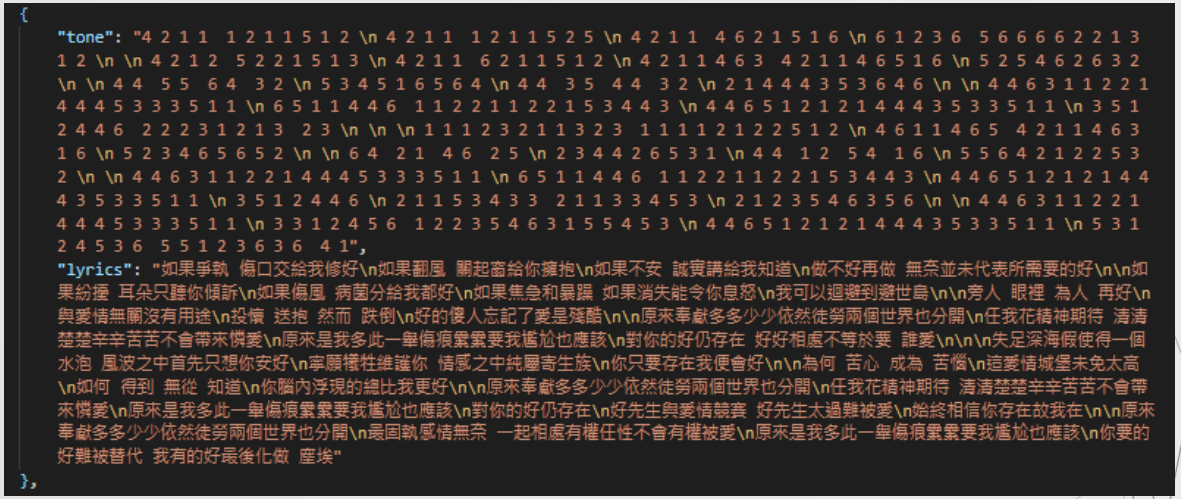


Figure 4.6 Tone-to-Text Data

First Stage: GPT-2 (Language Model)

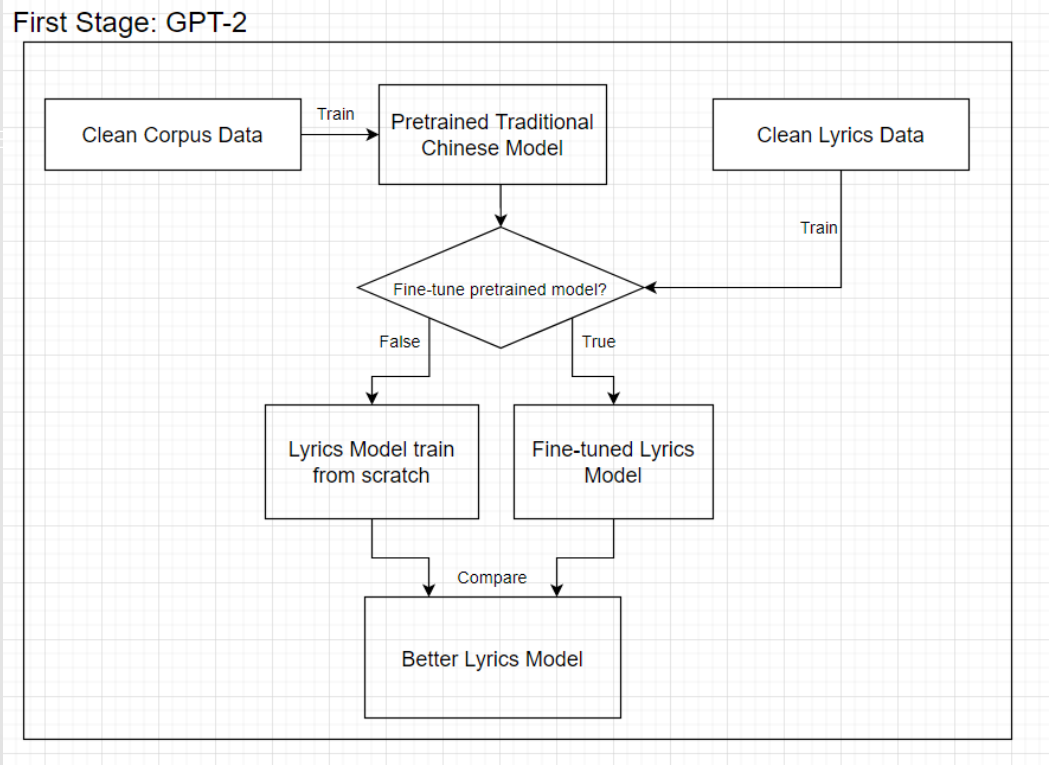


Figure 4.7 Flow chart of GPT-2 model phase

GPT-2 Evaluation

	GPT2 trained from scratch	Fine-tuned GPT2
BLEU	0.2	4.8
Perplexity	343.5803	235.6772

Second Stage:
Bart (Seq2Seq Model)

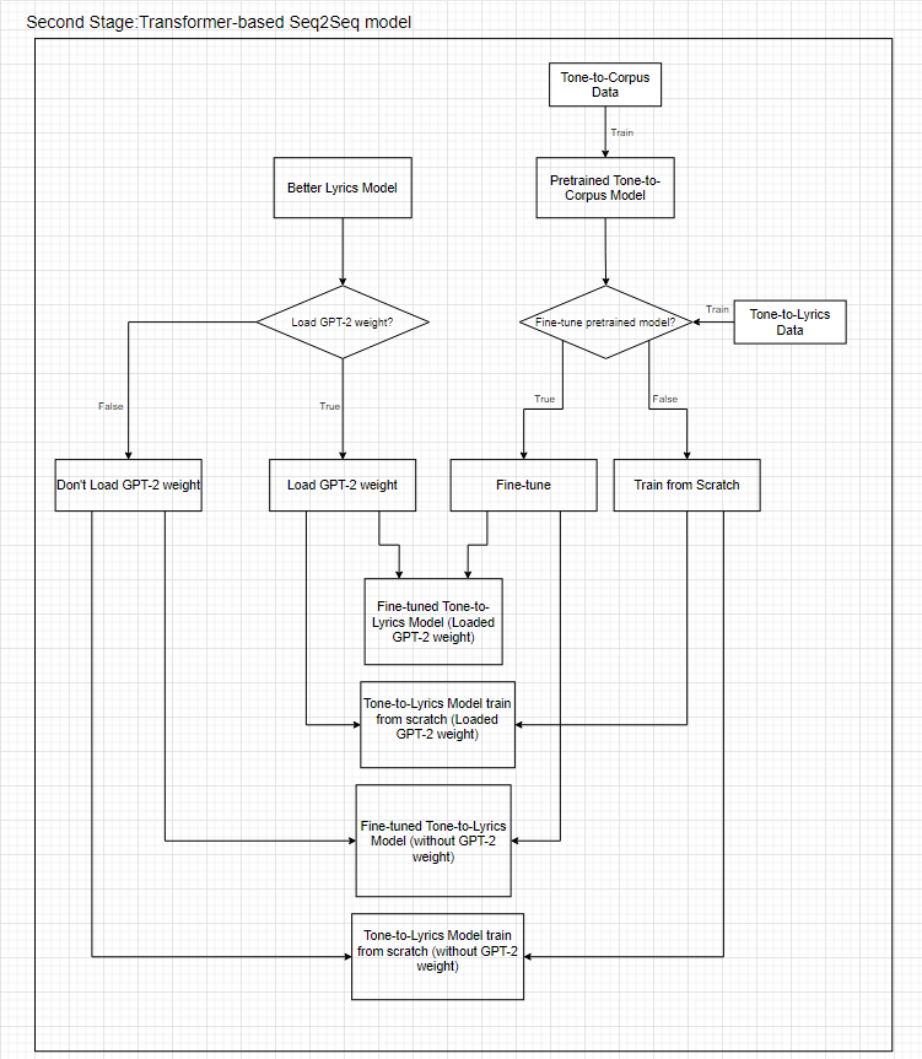


Figure 4.8 Flow chart of Bart model phase

Second Stage: Bart

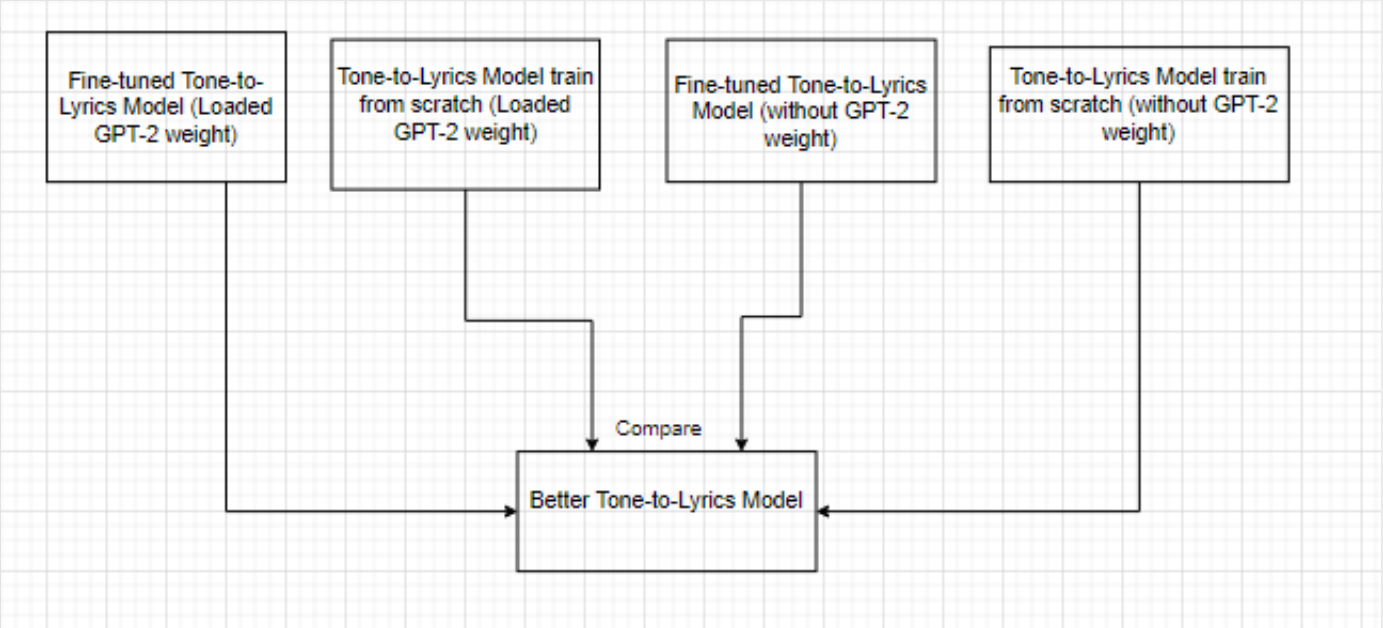


Figure 4.9 Flow chart of Bart models comparison

Second Stage: Bart

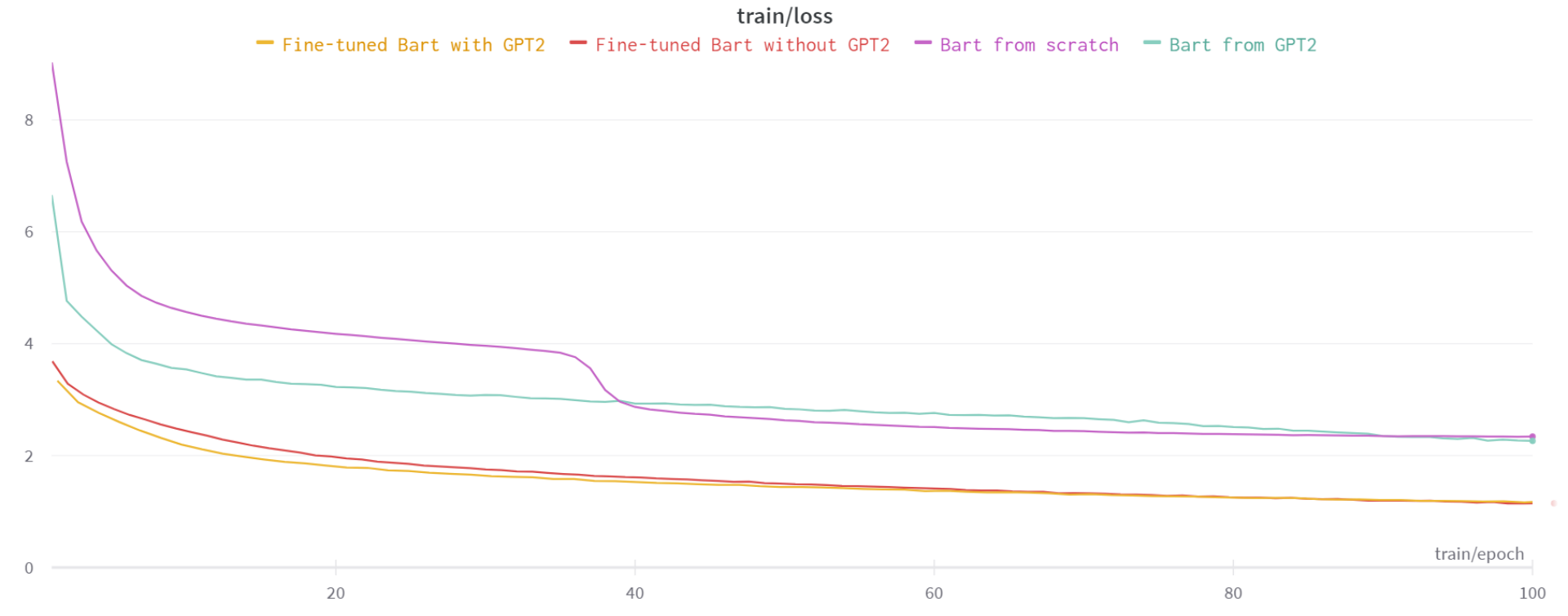


Figure 4.10 Training Loss

Second Stage: Bart

Figure 4.11 Training Final Loss

Second Stage: Bart

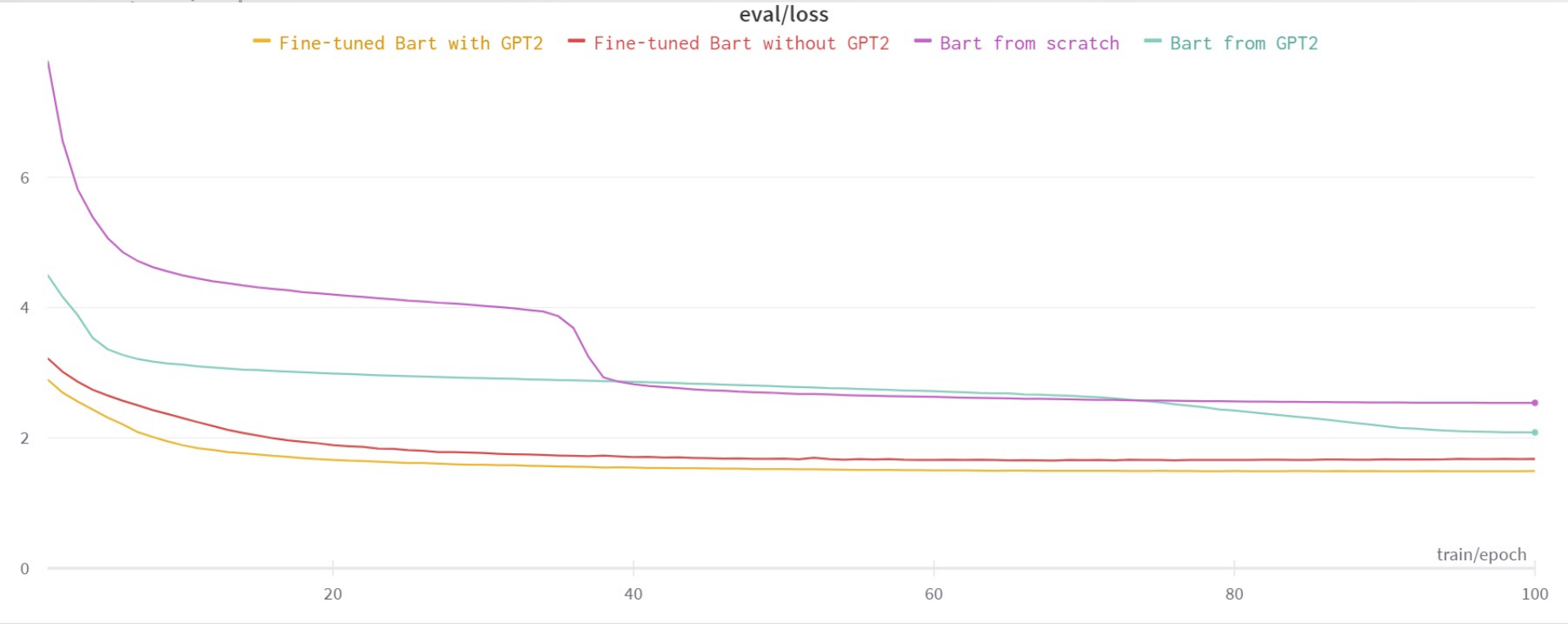


Figure 4.12 Validation Loss

Second Stage: Bart

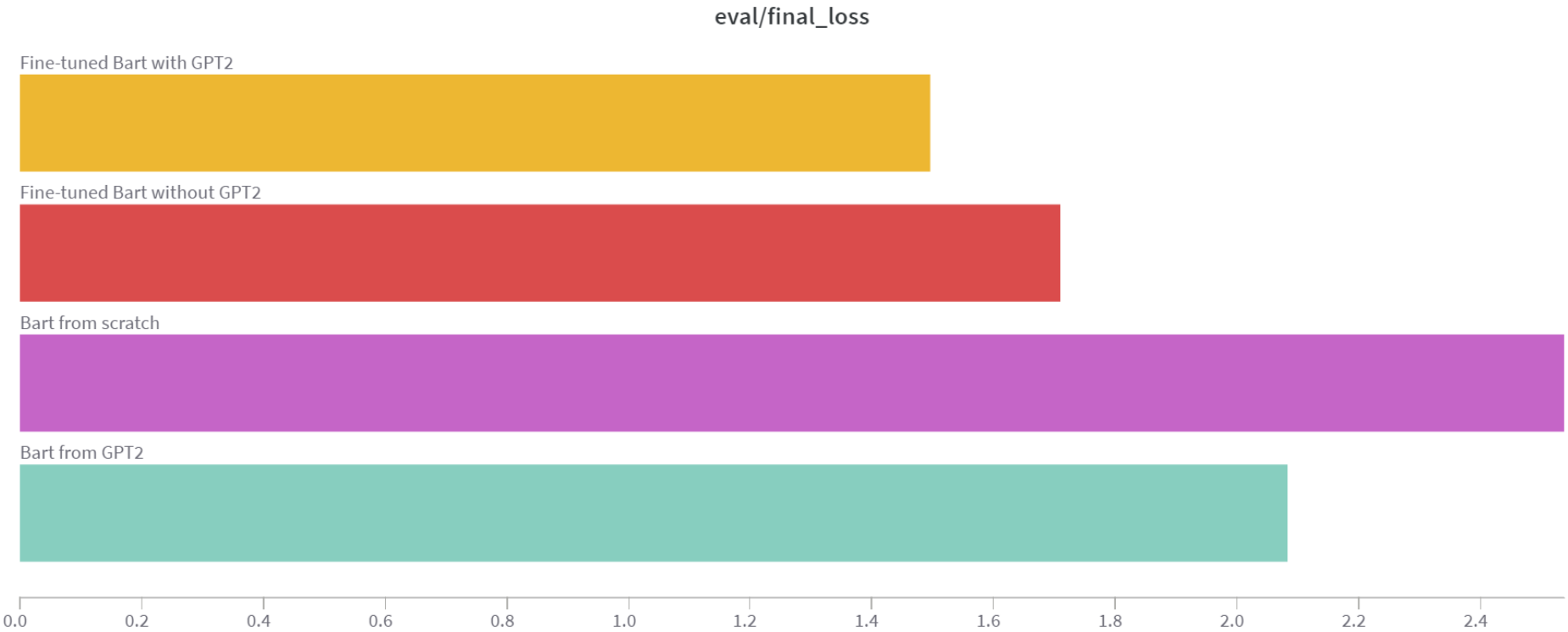
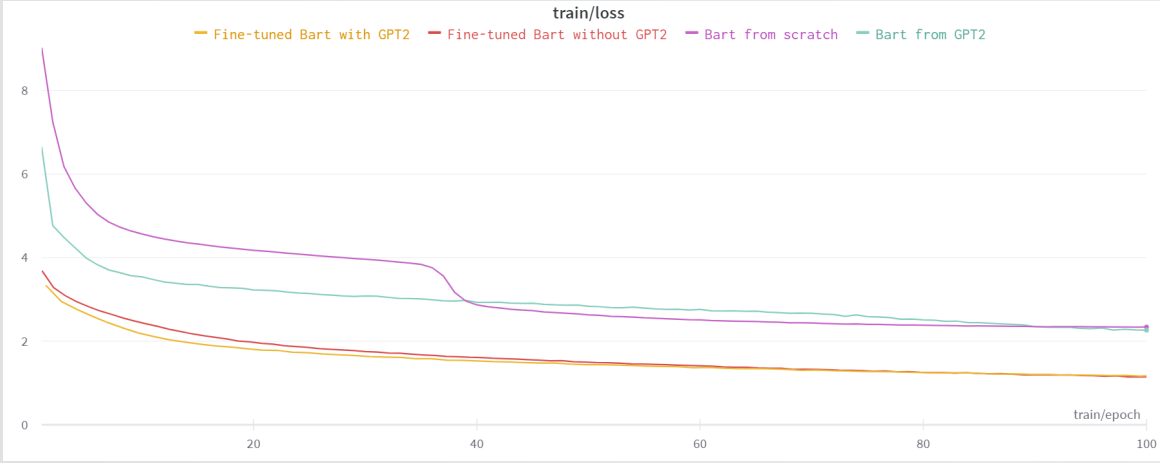
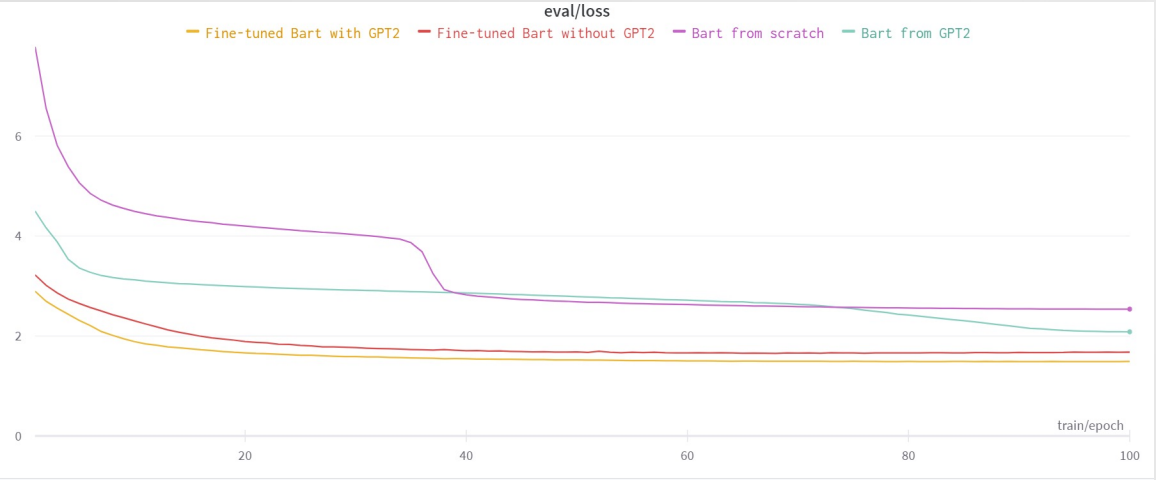
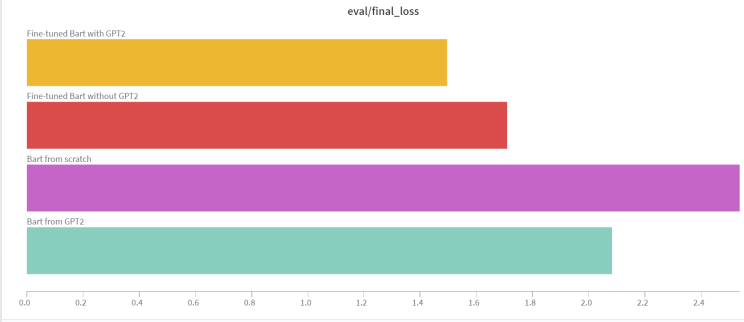


Figure 4.13 Validation Final Loss


train/loss
Fine-tuned Bart with GPT2
Fine-tuned Bart without GPT2
Bart from scratch
Bart from GPT2
8
6
4
2
0
20
40
60
80
100
train/epoch


train/final_loss
Fine-tuned Bart with GPT2
Fine-tuned Bart without GPT2
Bart from scratch
Bart from GPT2
0.0
0.5
1.0
1.5
2.0
2.5
3.0


eval/loss
Fine-tuned Bart with GPT2
Fine-tuned Bart without GPT2
Bart from scratch
Bart from GPT2
6
4
2
0
20
40
60
80
100
train/epoch


eval/final_loss
Fine-tuned Bart with GPT2
Fine-tuned Bart without GPT2
Bart from scratch
Bart from GPT2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4

Bart Evaluation

	Bart trained from scratch	Bart trained from a model loaded GPT-2 weight	Bart fine-tuned from a pretrained model	Bart fine-tuned from a pretrained model loaded GPT-2 weight
BLEU	0.3	0.3	0.7	0.8
Perplexity	674.8038	635.6714	344.2669	337.1880
Tone Accuracy	0.98975	0.9877	0.99655	0.9938

Samples

Original Lyrics

我不如過街老鼠

人人亦是個判官審判我壞處

令你羞恥有餘

如果抱住誰都叫住

Input

5 1 4 3 1 5 2

4 4 6 6 3 3 1 2 3 5 6 3

6 5 1 2 5 4

4 2 5 6 2 1 3 6

Sample Output

Bart trained from scratch without GPT-2 weight

Sample 1

我身從變的吻現

唯求命運氣轉一顆愛你沒變

願你一起你能

而戀你亦始終接近

Sample 2

無聲其實不會相

其情是別要怪不可算無事過

恨與深相距離

情可會做可不變動

Bart trained from a model loaded GPT-2 weight

Sample 1

我的情節分秒到

情形就像要跳得到跳上就快

為了擔險了迷

眉苦了就想不要亂

Sample 2

我的頭看不下這

情人沒問過去多久卻也沒怨

像我不想已從

還好你沒怎麼過沒

Bart fine-tuned from a pretrained model without GPT-2 weight

Sample 1

冷風流過的雨水

原來就是帶著一起向我蕩去

夜裡奔走旅程

回到你在這一個字

Sample 2

我的微笑不渺小

原來命運要看多少次你重看

沒有知己哪尋

而這已是誰的寄望

Bart fine-tuned from a pretrained model loaded GPT-2 weight

Sample 1

曙光融化心裡透

遙遙望外雪卻不懂放下內疚

在你的口裡留

留到你在此刻照舊

Sample 2

無辜來世將永久

流離樹絕世欠生死去了又見

沒有花果也甜

回首也沒感恩怨恨

05

Tone2Cantopop

web tool for public to generate lyrics

Lyrics Generation

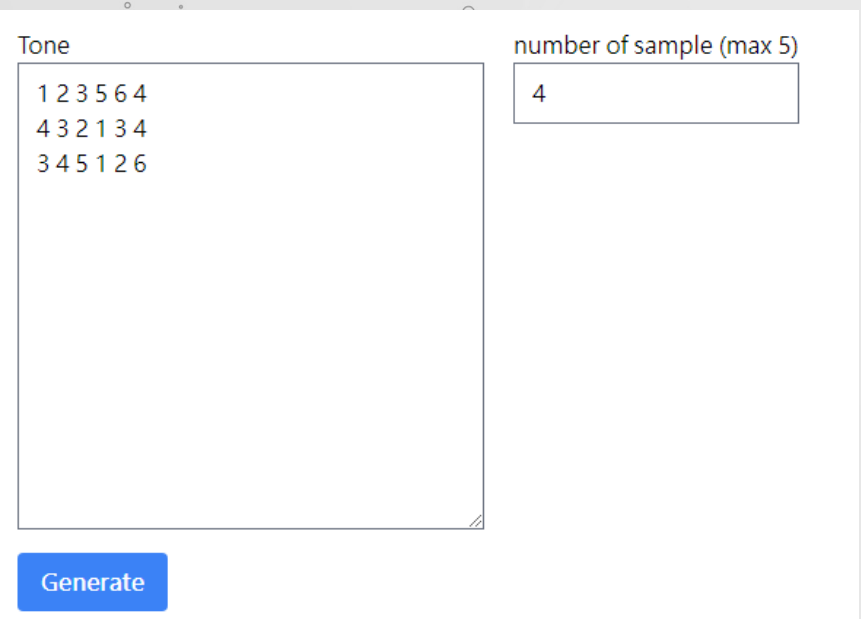
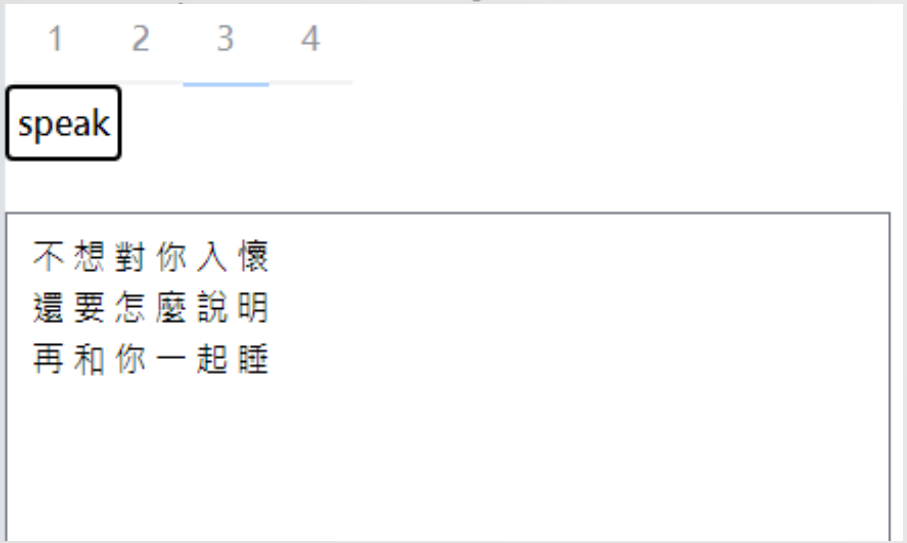


Figure 5.1 Input interface for lyrics generation

- Input the tones converted from the melody
- Input the number of sample to be generated

Lyrics Generation



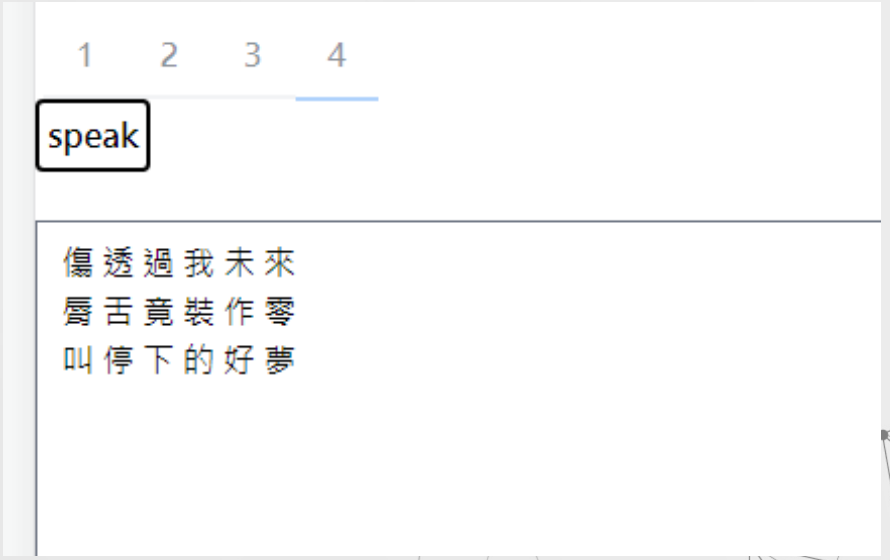


Figure 5.2 Tabs to toggle between generated samples

Text-to-speech

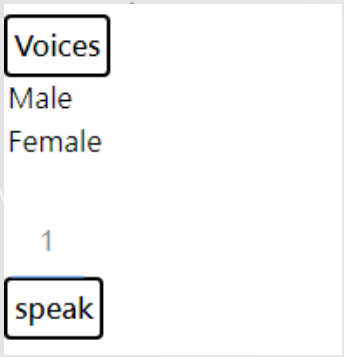


Figure 5.3 Text-to-speech interface

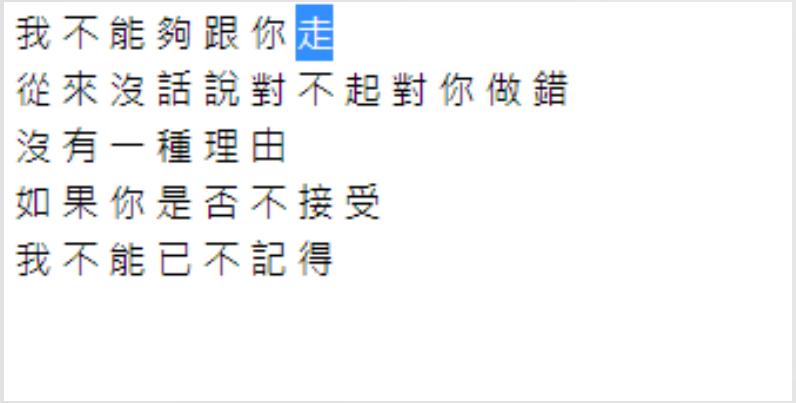


Figure 5.4 Spoken word highlight feature

- Two voices are provided
- Spoken word highlighting feature is provided

Tone Comparison

Input Tone	Output Tone	Tone Accuracy
1 2 3 5 6 4	1 2 3 5 6 4	88.89%
4 3 2 1 3 4	4 3 2 1 3 4	
3 4 5 1 2 6	3 4 5 2 1 6	

Figure 5.5 Tone comparison feature

- Wrong output tone is highlighted in red
- Overall Tone Accuracy is calculated

Live Demonstration

06

Limitation

Size of dataset

10 GB — Traditional Chinese Corpus Data

200 MB — Tone-to-Corpus Data

15000 — Lyrics / Tone-to-Lyrics Data

Multiple possible tones of a Chinese character

中

部首: | [2]

大五碼: A4A4

音節 (香港語言學學會)	粵音	根據	同音字	相關音節	
zung1		黃(p.47) 周(p.1) 李(p.44) 何(p.342)	鐘, 盅, 腫 [46..]	--選擇--	中土, 中心, 中天[32..]
zung3		黃(p.47) 周(p.1) 李(p.44) 何(p.343)	堹, 倲, 瘇 [11..]	--選擇--	中邪, 中肯, 中風[7..]

Figure 6.1 Example of a Chinese character with multiple tones [6]

- Example: 中 (middle)
  - 中心 (center) and 中風 (stroke)
- Character-based Tokenization is used
- Word level / Sentence level Tokenization may solve the problem

07

Possible Future Development

Possible Future Development

- Theme/Title based lyrics generation
- Lyrics generation with partly finished lyrics
- Model Configuration Adjustment

Model improvement

- Account system
- Rating system

App Improvement

Conclusion

- Tone-based lyrics generation
  - A tailor made approach for Cantopop lyrics generation.
- Sequence-to-Sequence Models (Bart)
  - Trained with Training from scratch, Pre-training and Fine-tuning
- Tone2Cantopop
  - A Web tool open to public

THANKS

Reference

[1] Hung-yi Lee, *Transformer*, (Jun. 01, 2019). Accessed: Nov. 23, 2021. [Online Video]. Available: https://www.youtube.com/watch?v=ugWDIIOHtPA

[2] A. Vaswani *et al.*, 'Attention Is All You Need', *arXiv:1706.03762 [cs]*, Dec. 2017, Accessed: Nov. 15, 2021. [Online]. Available: http://arxiv.org/abs/1706.03762

[3] M. Lewis *et al.*, 'BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension', *arXiv:1910.13461 [cs, stat]*, Oct. 2019, Accessed: Nov. 15, 2021. [Online]. Available: http://arxiv.org/abs/1910.13461

[4] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, 'BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding', *arXiv:1810.04805 [cs]*, May 2019, Accessed: Nov. 15, 2021. [Online]. Available: http://arxiv.org/abs/1810.04805

[5] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, 'Improving Language Understanding by Generative Pre-Training', p. 12.

[6] '粵語審音配詞字庫' . https://humanum.arts.cuhk.edu.hk/Lexis/lexi-can/ (accessed Dec. 01, 2021).