GPT2_Chinese

The advancement of AI technology has made it possible to automatically generate many things with AI. The language model of the GPT series can already generate articles, answer questions, and translate. Here is a GPT-2 Chinese article generation APP for you, which can automatically generate various articles.

Version 20230223

Applications

- Using the GPT2 algorithm can be applied to automatically generate lyrics, automatically generate ancient prose, automatically generate poems, automatically generate couplets, automatically generate articles, automatically generate novels, etc.

How to use

- The main process is:

Prepare article samples ->
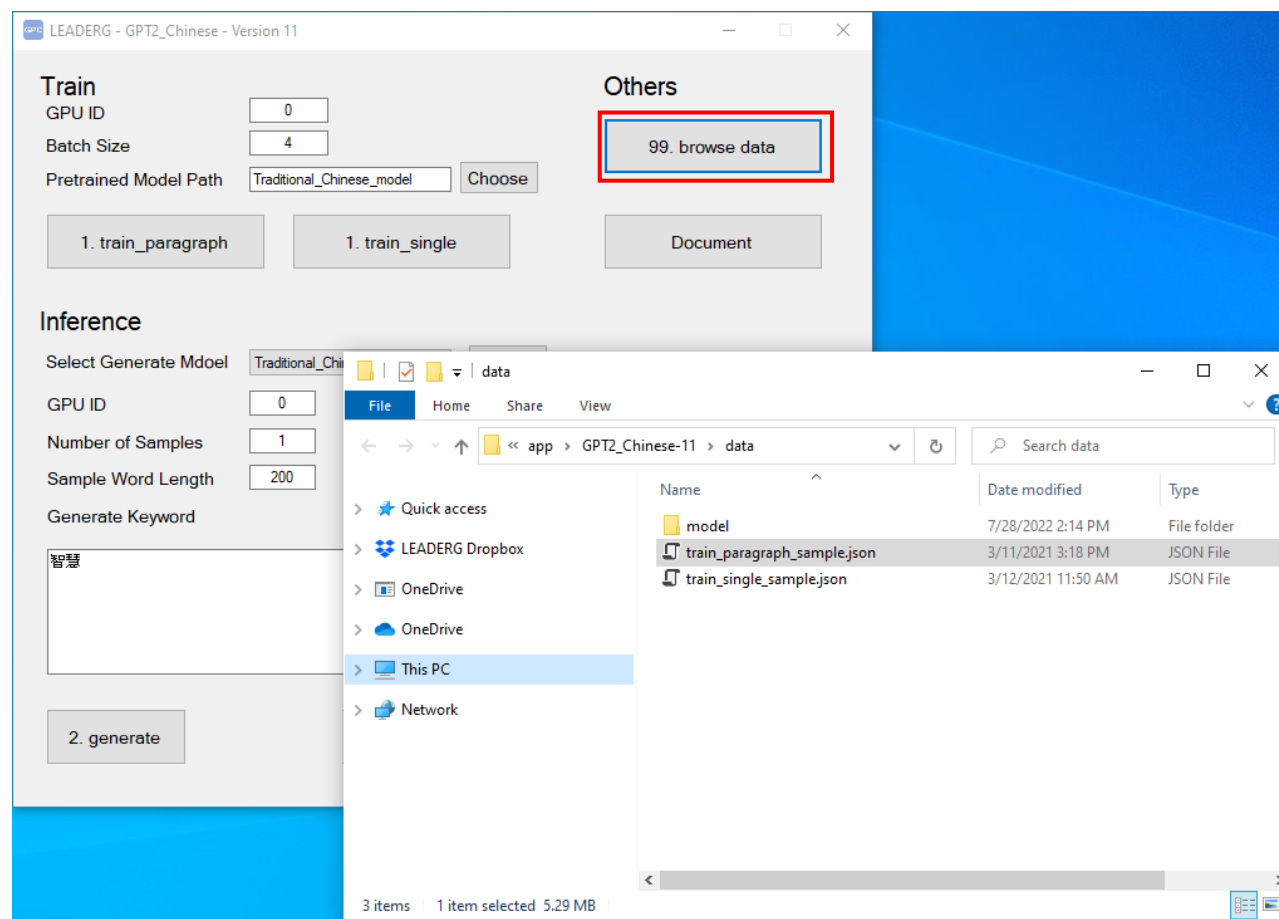
train -> generate articles

Prepare the article

After pressing [99. browse data], open the data folder, you will see train_paragraph_sample.json and train_single_sample.json, one is a text format composed of many paragraph articles, and the other is a text format with a single article content.

Prepare training samples of paragraph type

If the prepared sample is a paragraph type article, such as lyrics, ancient texts, news headlines, etc., the format of this book can refer to "train_paragraph_sample.json".
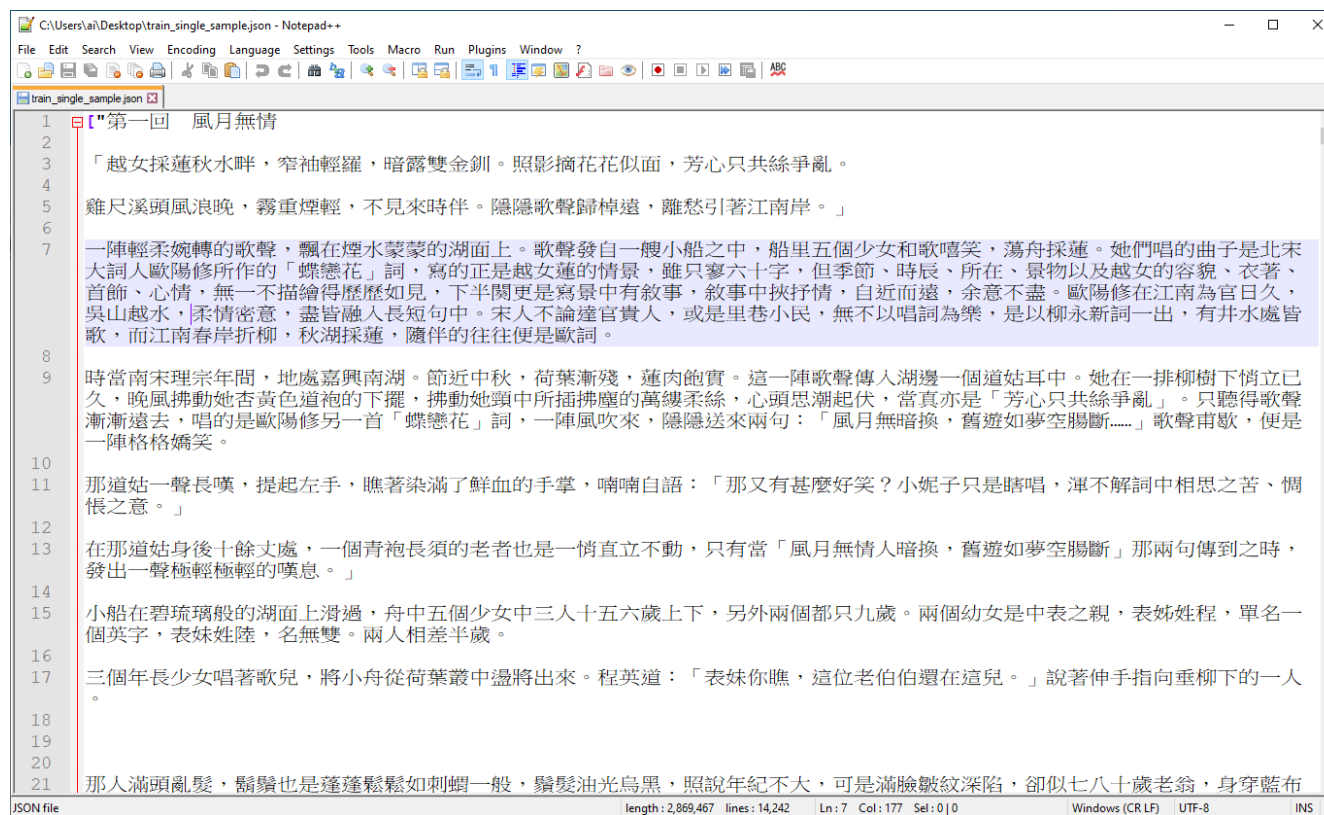
The format of the sample is ["paragraph 1", "paragraph 2", "…", "paragraph n"]，, taking the example of APP as an example, it is like ["哈通保卡山……", "帕哈林保卡山……", "夏侯獻……", …… , "任天堂Labo …… 截然不同。"]。

Prepare a training sample for an entire article

If the prepared sample is of the type of an entire article, such as the content of a novel prepared by the APP, please refer to "train_single_sample.json" for the format of the sample. The format of the sample is ["an entire article content"], taking the example of the APP as an example, it is like ["第一回　風月無情 ….. 難為情。」"]。

Model introduction

GPT2_Chinese-11 > data > model

The APP provides 8 models:

- chinese_lyrics_model: Simplified Chinese lyrics model.
- classical_chinese_model: Simplified Chinese classical Chinese model.
- couplet_model: Simplified Chinese couplet model.
- general_chinese_model: Simplified Chinese article model.
- general_chinese_small_model: Simplified Chinese short article model.
- JinYong_model: Traditional Chinese Jin Yong novel model.
- poetry_model: Simplified Chinese ancient poetry model.
- Traditional_Chinese_model: Traditional Chinese article model.

Remark:

- final_model is the resulting model location after training.
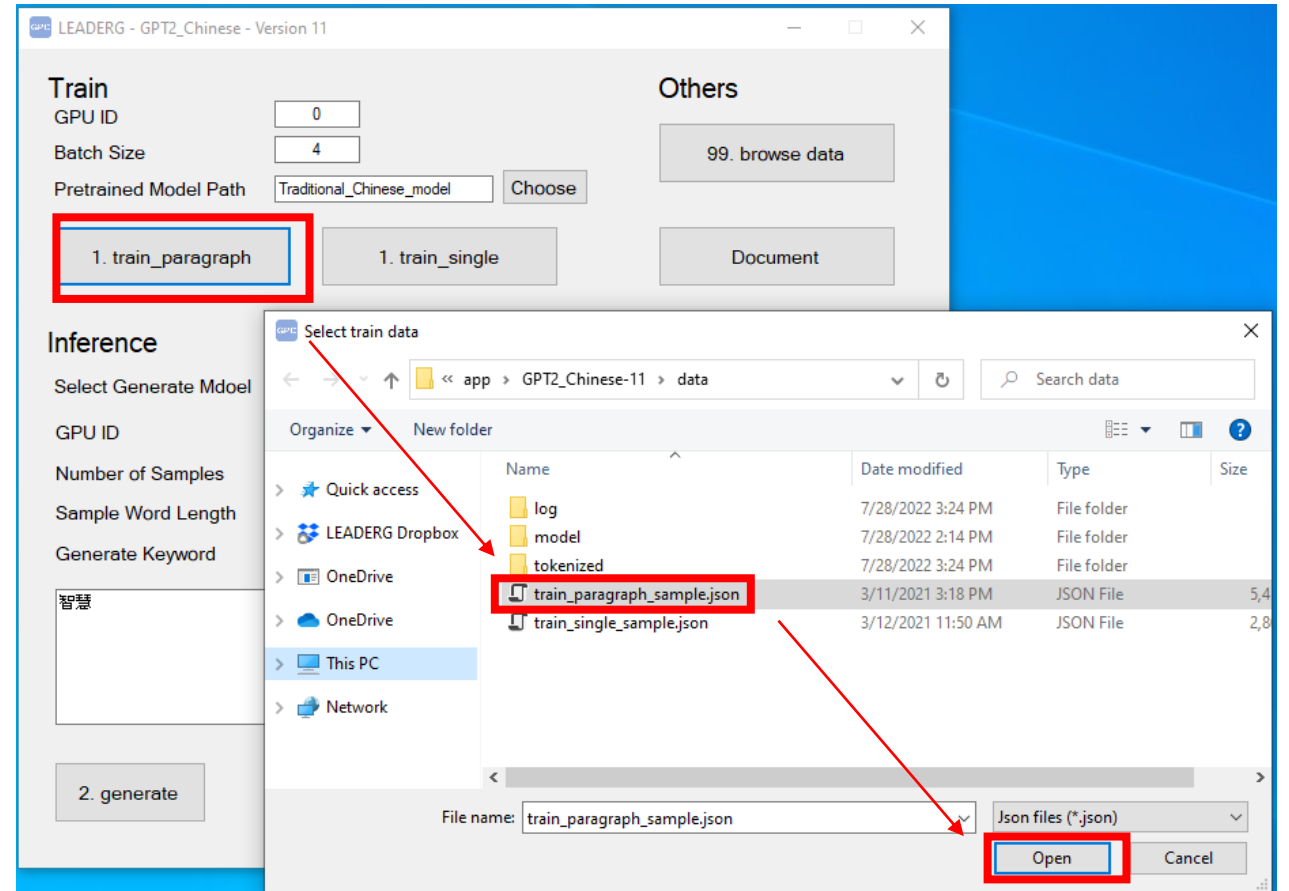
Name

- chinese_lyrics_model
- classical_chinese_model
- couplet_model
- final_model
- general_chinese_model
- general_chinese_small_model
- JinYong_model
- poetry_model
- Traditional_Chinese_model

1. train_paragraph

If the prepared text type is paragraph type, you can press [1. train_paragraph] to select your own samples for training. The default Traditional_Chinese_model (Traditional Chinese model) is a pre-trained model.
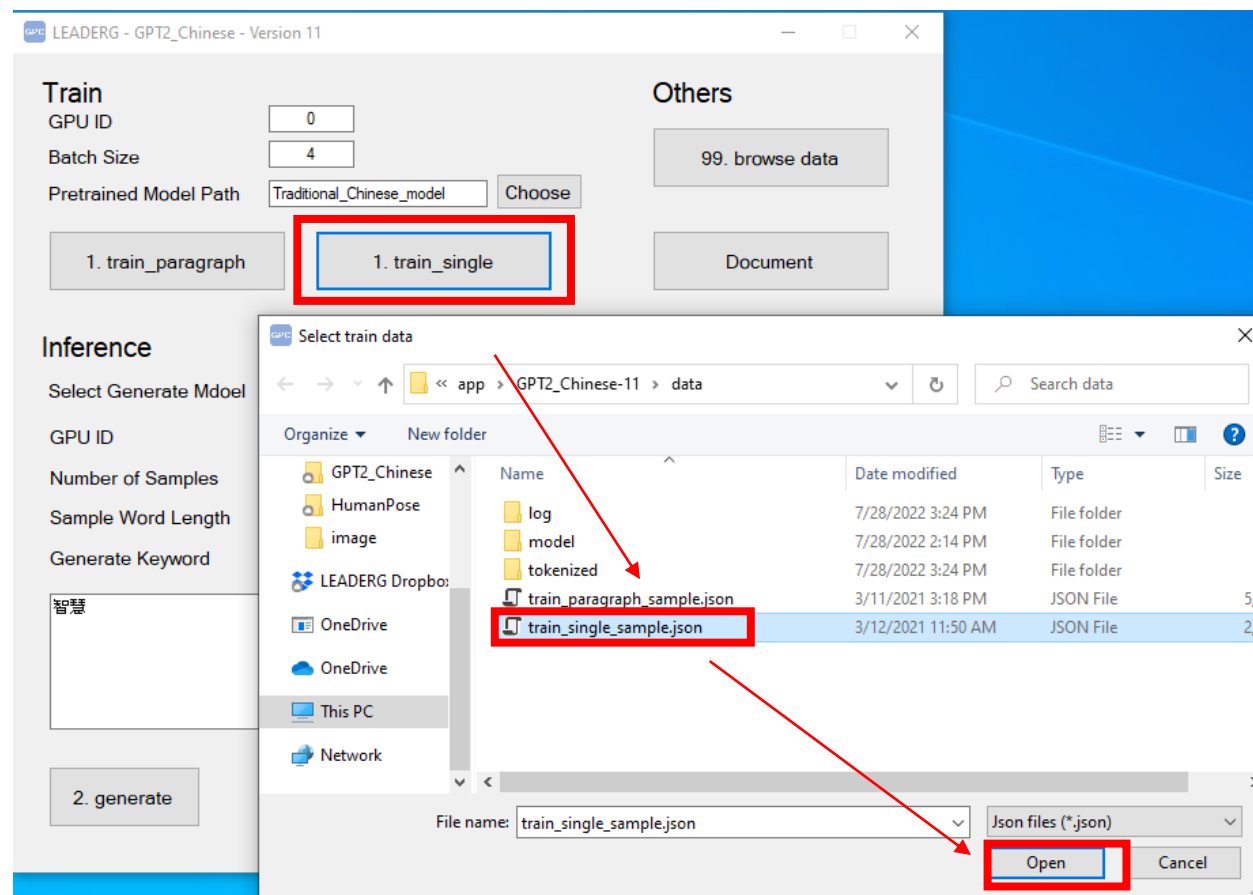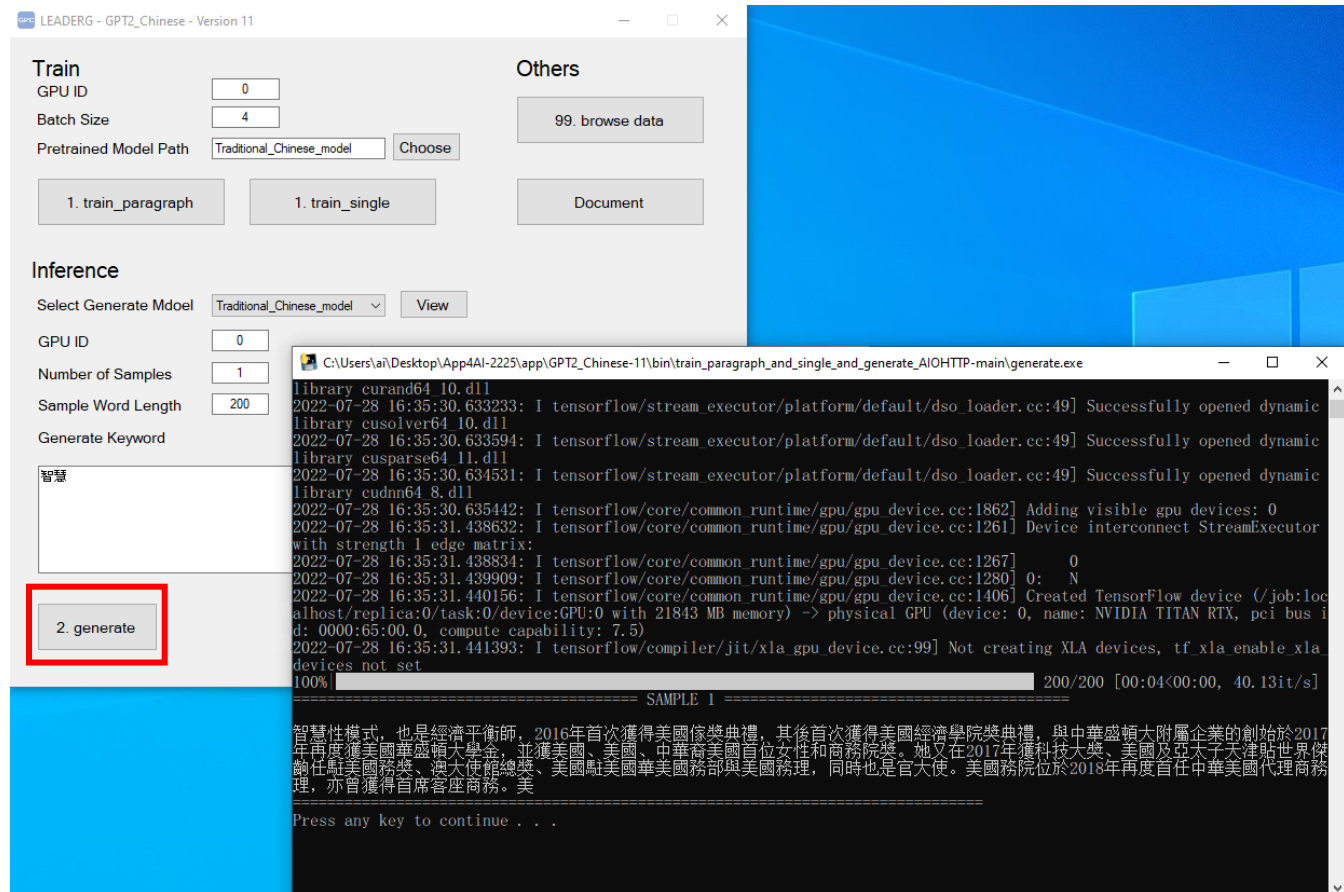
The trained model will exist in "final_model".

1. train_single

If the prepared text type is an entire article type, you can press [1. train_single] to select your own sample for training. The default Traditional_Chinese_model (Traditional Chinese model) is a pre-trained model.The trained model will exist in "final_model".

2. generate

Select the model to generate text (Select Generate Model), you can refer to the model introduction to select the model, set the number of articles you want to generate (Number of Samples), and a reasonable number of words to generate (Sample Word Length), and finally enter as the beginning of text, you can press [2. generate] to generate the article.

Reference

- Please refer to the readme.txt in the APP folder.
- LEADERG AppForAI: https://www.leaderg.com/appforai-windows
- Copyright © LEADERG INC. All rights reserved.