

Assignment 5

Sonnet Generator

The goal of this given assignment is to develop the poem generator. The generated sonnet must comply the style of **Abay Kunanbayev**. You have to collect the dataset yourself. The dataset should consist of 20-30 sonnets of Abay Kunanbayev in order to have well performance.

Sample sonnet (generated):

Махаббат, қызық кім көрер,
Оның да дәмін татпаса?
Біржола басын,
Оты қанақтық шарпай жүріп бақпаса.
Мал бақпақтық көрсе сөзіңдер.
Есіткенде шығып қалды.
Де айысы шап,
Кісілікті
Тала алмас әсте жуынбай.
Тәңірінің күні жасқан,
Ғылымға көң зерек-т.
Аямаған ғарптен
Қолдан жерікке болады қылығың.
Жақсыз балап, сүйісек,
Бір күйгізіп, сүйгізіп, сүйгізіп,
Аз күн қоймай ма басты көрсе сөзіңдер.
Күлкің сөзі ұрып, қыл, байлап,
Еркеленіп шығап,
Менікі көрген-білгенін
Білдірер алысы – үскірік, аяз бенер – келіп қалды.
Дем айығыз біліп сондайды, жанын дос ерер,
Жолдастық, сап-сип,
Енді кім бола ма, сойға сөйлерде өткен құрғырды қас қылпыңқы, аламайлы,
сығпасын,
От жалын боп шалысын.
Жылай-жырлай өлгендь келген көмект.
Ксге сдһтың адалы, аяғы тік,
Жаурыны етсіз, жалпақ шанар, түпсінің,
Атын сүйгендей мекі.
Күнде айлағы қылпың,
Әз басың сөз.
Сондасың ейді тауып, сүйісең,
Бір күйгізіп, сөзін біл!
Осы жалған жерді терге де,
Танданың боп айлас,
Кетті матап, тұрды ма?

Evaluation

- 1) Dataset collection - 20 points
- 2) Model implementation - 20 points
- 3) Model training - 20 points
- 4) 4 generated sample sonnets - 40 points

Tips

- 1) Is it the same as name generation?
- 2) How is the problem of larger depths solved?
- 3) How are predictions of next line and ending of the text handled?

Submission policy

You have to submit 3 files and 1 folder. If **model.pt** weights small, then put 3 files and 1 folder, **source.ipynb**, **source.html**, **model.pt** and **data** folder respectively, in a single folder named with your **student id** (for example 180102030). The folder should be compressed in ZIP format. If **model.pt** weights big, then send it in other ways.

180102030/

1. **source.ipynb**

Please write your name and surname on the top in addition

2. **source.html** *(you can miss this file, but I suggest you to upload it as well. If source code would have problems, then this file guarantees your work done)*

This is HTML version of your source code. You can export the HTML version of your work in Jupyter Notebook.

3. **model.pt**

Do not forget to save the state of you WEIGHTS!!!! You can read more about it from: https://pytorch.org/tutorials/beginner/saving_loading_models.html.

If the moodle do not allow you upload big files, then upload it into your google drive and add the link into your work, or share it in other ways.

4. **data**

This is the dataset folder which contains the sonnets of Abay Kunanbayev.