

EARIN project Midterm report

Krzysztof Rudnicki

Jakub Kliszko

May 29, 2023

1 Progress

We have implemented reading data from csv files, preprocessing them with optional showing of some of the information about the data and used model/learner for implementing neighbour searches

Right now the model predicts random anime chosen from the table

2 Results

2.1 Presentation

2.1.1 Plots

2.1.2 Tables

Seed We added seed in predict function for choosing random anime, using the same seed always returns same recommendations and choosing random anime is the only random part of our code

User can specify their own seed by using -s or --seed flag by entering in command line:

```
python -s 42
```

3 Challenges

3.1 Failed attempts

Biggest challenge was realizing how overcomplicated and unnecessary difficult to implement is the first code we based on: Kaggle code with tensorflow. This solution runs for almost 10 minutes on kaggle and implementing it to run on our local devices was a real chore that took us a good day and a half to implement.

This implementation is based around very powerful Tensor Processing Unit from google and while it is possible to change it to run on local graphics card it requires downloading both cuda and cudnn to a downgraded version supported by tensorflow (11.8) and downgrading graphics card drivers. Running it with CPU results in the model training for over 3 hours.

3.2 Corrections

Surprisingly even though we based our preliminary report around different example code we managed to not make any corrections to preliminary report. All of functionality that we want to implement is available in sklearn and scipy.

3.3 Results and findings

4 Finishing project