Weekly Schedule Schedule: Monday, 7:00 PM - 8:00 PM In-Person, Wednesday 3:00 - 5:00 PM In-Person, Saturday 1:00 PM Online.

Project: MBTI Personality Classification

Our github repo: https://github.com/acmucsd-projects/Team-TBD

Attendees: Catherine, Phillip, Ryan, Aryaman, Sia, Hargen

Summary of Meeting

Action Items

EVERYONE: Add links/tools you use to the resources section at the end of this google doc. Also, PUSH TO GITHUB WHENEVER YOU HAVE DONE SOMETHING.

Tasks by Saturday

Model Improvement (Hugging face??) and modularize code: Hargen, Catherine, Phillip

- [ ] Improve the model in terms of model accuracy (Keep model below 300m parameters)
- [ ] Incorporate hugging face to tokenize the texts (learn the basic setup of hugging face)
- [ ] Learn necessary tools as we go to improve the model
- [ ] Convert the BERT from previous project team into our model
- [ ] Modularize notebook into individual .py files (train.py (take in hyperparameters like lr), utility.py, process.py, dataset.py, model.py)
- [ ] automate the training process by writing a script to train the model with parsed hyperparameters (with argParse - argParse library to parse arguments to automatically run the notebook)
- [ ] Save model weights, codes, base model

Learn app deployment (streamlit/gradio): Aryaman, Sia, Ryan

- [ ] Learn the basics of the model (to better deploy the app?)
- [ ] Learn the basics of app development tools
- [ ] Get chatbot running on streamlit
    - [ ] Could be existing model

ADD TASKS ABOVE IF YOU THINK YOU CAN COMPLETE THAT BY SATURDAY!!!

DO WE AGREE??? If yes, type your initials!!!

RW, CZ, HZ, PW

- Deliverables for others - suggestions to improve it (Just improve it)
- https://github.com/acmucsd-projects/sp23-ai-team-1/blob/main/models/model_training.ipynb
- Switch model from pytorch to huggingface → it's not switching, we still use pytorch, but we feed

in our model into hugging face

Random Ideas:

- Before running the whole training, run one batch of data and see if the model learns anything.
- Collaborate maybe miro

Past Resources

- Learn the Basics — PyTorch Tutorials 2.1.0+cu121 documentation (In Progress)
- Sentiment Analysis Tutorial
- (1) PyTorch Prerequisites - Syllabus for Neural Network Programming Course - YouTube
- Any useful resources for the team can be organized into this resources/ folder
- **Learn PyTorch** from YouTube tutorial (as much as you can)
- Learn basic classification pipelines from Kaggle (from other people's notebook)

Kaggle Dataset we use (More to be added maybe?):

https://www.kaggle.com/datasets/datasnaek/mbti-type

Kaggle one with various classification methods:

https://www.kaggle.com/code/abhijitsingh001/mbti-test-your-personality

Hugging face model

https://huggingface.co/xlm-roberta-large

Bert hugging face model:

https://huggingface.co/bert-base-uncased

add files to commit to github:

Git add .

commit the files with a message

Git commit -m "<message>"

push the files to the main branch

Git push

get the files that people have changed

Git pull

Resources