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

Our github repo: https://github.com/acmucsd-projects/Team-TBD
Attendees: Catherine, Phillip, Hargen

Summary of Meeting

Action Items

Tasks by Wednesday

- Hargen - update requirements.txt
    - Push the notebook into github
    - Put headers and explain the code
    - Strip out all unnecessary things
    - Update instruction on how to set up the dataset in the repo.
- Others- -
    - Go through the notebook by hargen to understand it
    - We are using pytorch not scikitlearn
    - Learn how to do sequential models with pytorch
    - argParse library to parse arguments to automatically run the notebook
        - Want script in the repo
- Deliverables for others - suggestions to improve it (Just improve it)
- Convert the BERT from previous project team into our model https://github.com/acmucsd-projects/sp23-ai-team-1/blob/main/models/model_training.ipynb
- App: Streamlit to deploy apps (if we have more time before Wednesday's meeting)
    - Can also use gradio
- Switch model from pytorch to huggingface
- Keep model below 300m parameters
- Learn to set up Hugging Face

If not already, do the basic pipeline on a notebook and then convert the notebook into

individual .py files to automate the process.

- ☐ Download dependencies on computer and run code via VSCode (if not sure, text on discord for help!) – also set up the anaconda virtual environment.
- ☐ Put into script form to build and train model using cli args (train.py (take in hyperparameters like lr), utility.py, process.py, dataset.py, model.py)
- ☐ (Ryan) fix env

Later:

- ☐ Ready – to improve our model.
- ☐ Run training run with script using specific hyperparameters
- ☐ Save model weights, codes, base model
- ☑ ~~List of dependencies on the requirement.txt – update requirement.txt as you add libraries~~

Tip:

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

Project: Personality Test

Overview from last time:

- 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):

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

Kaggle one with various classification methods:

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