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: <u>MBTI Personality Classification</u>
Our github repo: https://github.com/acmucsd-projects/Team-TBD
Attendees: Catherine, Phillip, Ryan, Sia, Hargen

Summary of Meeting

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 – SO WE ALL KNOW.

<u>Tasks FOR EVERYONE by Saturday</u>

Model Improvement (Hugging face??) and modularize code: <u>Hargen, Catherine, Phillip, Ryan</u>

Hargen: improve model and aim for at least 50% validation accuracy and organize code better

Phillip, Ryan: Learn how to use model to predict user input/ our own text

Catherine, Sia: figure out how to tokenize data using bert-based-uncased

- [ ] Improve the model in terms of model accuracy (Keep model below 300m parameters<u>)</u>
- [ ] Incorporate hugging face to tokenize the texts (learn the basic setup of hugging face)
- [ ] Learn necessary tools just enough to improve the model
- [ ] Convert the <u>BERT or other pre-trained models</u> into our model (from previous project team [repo](#))
- [x] ~~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)
- [x] ~~Save model weights, codes, base model~~

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

- [ ] 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 MONDAY!!!

Random Ideas (add here if you have thoughts on how to improve our project):
- Before running the whole training, run one batch of data and see if the model learns anything.
- Collaborate maybe miro

Datasets:
Kaggle (MBTI) Myers-Briggs Personality Type Dataset

Notebook with various classification methods on above dataset

Hugging face model

Bert hugging face model

Resources Section (add whatever tools you have used here!!!)
- Learn the Basics — PyTorch Tutorials 2.1.0+cu121 documentation
- 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)