

ACM AI Team 1: MBTI Personality Project

Team Meeting Time: Every Monday 5:00 - 6:00 PM

Weekly Mentor Meeting: Every Friday 5:00-6:00 PM

Attendees: Kevin, Vanessa, Yashil, Samuel, Vincent

Important Links:

- MBTI personality type using tweets : [MBTI Personality Type Twitter Dataset | Kaggle](#)
- Our github repo: [acmucsd-projects/sp23-ai-team-1 \(github.com\)](#)
- Notebook: <https://www.kaggle.com/samuelllee1/acm-sp23-team-1-project-mbti>
- Preprocessing: <https://colab.research.google.com/drive/17kLdesMgX6Z6PdiSGeoGPvoFMk-nmP00?usp=sharing>

Action Items

- Drop index columns
- Tokenize text using HuggingFace (preprocessing)
- Then use model in HuggingFace
- Everyone can do, but use different models to tokenize and train (ex: bert-based-uncased, XLM-roberta-case, ...)

Summary of Meeting

- Check progress for past week
- Find way to tokenize the texts: HuggingFace: bert-base-uncased
- Then train with HuggingFace
- LSTM RNN layer to optimize
- XLM-roberta-case instead of bert-case-uncased

Timeline

- Step 1: Data Cleaning (28th April, Friday of Week 4)
- Step 2: Preprocessing data (1st May, Monday of Week 5)
- Step 3: Model training and validation (8th May, Monday of Week 6)
- Step 4: Decide what kind application to do and start (15th May, Monday of Week 7)
- Step 5: Continuation of the step 4- (Weeks 8,9)
- Have an app ready with Streamlit (Week 9)
- Step 6: Polish it up + presentation + showcase (Week 10)