

Bi-Weekly Report of (Team #13)

Week #8

What were the goals for the last 2 weeks?

- Create an algorithm that reflects other features in addition to self-introduction in the recommendation algorithm
- Create data interface with modeling and API design
- Help create an algorithm and create a data interface with modeling
- Finish the Figma model and share it with team members

What goals were accomplished this week?

[Data preprocessing]

```
def generate_introduction(jobs, personalities, likes, num_sentences):
    introductions = []
    for _ in range(num_sentences):
        job = random.choice(jobs)
        personality = random.choice(personalities)
        like = random.choice(likes)
        introductions.append(f"나는 {job}이다. 나의 MBTI는 {personality}이다. 나는 {like}을(를) 좋아한다.")
    return introductions
```

```
jobs = ["음악가", "작곡가", "가수", "DJ", "음향 엔지니어", "게임 개발자",
```

```
likes = ["화초 돌보기", "동물 돌보기", "새 관찰", "별보기", "천문학", "지리학", "역사 탐구", "문화 탐구",
```

The code that was written previously allowed people under the age of 10 to have a variety of jobs, but the code was modified to limit the occupation of people under the age of 10 to be a student. It seemed like the occupations were too focused on professionals, so we increased the diversity of occupations by adding more diverse occupations such as DJs and craftsmen. Hobbies also seemed to be focused on sports, so I added humanistic hobbies such as writing poetry and watching musicals.

[Recommendation system]

- recommend 10 top users based on content based recommendation
 - filtered by mother tongue language, language to learn, gender, language level
 - self-introduction embedding with TfidfVectorizer and stopword preprocessing
 - cosine similarity between self-introduction (e.g. job, MBTI, hobby)
 - it can resolve the cold start problem, which can make recommendation possible without knowing who the user likes.

```
499,Addisynn,남,30대,나는 음악가이다. 나의 MBTI는 INFJ이다. 나는 스탬프 수집을(를) 좋아한다.,영어,일어,하
500,Addlee,여,30대,나는 특수교육 교사이다. 나의 MBTI는 ENFP이다. 나는 글쓰기를(를) 좋아한다.,한국어,중국어,상
501,Addley,여,60대 이상,나는 해양생물학자이다. 나의 MBTI는 ESTP이다. 나는 천문학을(를) 좋아한다.,영어,일어,상
```

Example Input: User_ID 500, Addlee wants user who is male, has high level language, has mother tongue she wants to learn.

Output:

User_ID	이름	성별	연령대	자기 소개	모국어	외국어	외국어 수준	상
26697	26698	Rebeckah	남	50대	나는 영화 감독이다. 나의 MBTI는 ENFP이다. 나는 글쓰기를(를) 좋아한다. 중국어 한국어	상		
3276	3277	Ariona	남	20대	나는 특수교육 교사이다. 나의 MBTI는 ENFP이다. 나는 뮤지컬 보기를(를) 좋아한다. 중국어 한국어	상		
13721	13722	Jahmal	남	20대	나는 특수교육 교사이다. 나의 MBTI는 ENTJ이다. 나는 BMX을(를) 좋아한다. 중국어 한국어	상		
43193	43194	Adamarys	남	20대	나는 특수교육 교사이다. 나의 MBTI는 ENFP이다. 나는 개인 금융 관리를(를) 좋아한다. 중국어 한국어	상		
38989	38990	Baiden	남	10대 이하	나는 학생이다. 나의 MBTI는 ENFP이다. 나는 카약을(를) 좋아한다. 중국어 한국어	상		
24031	24032	Myshawn	남	60대 이상	나는 중학교 교사이다. 나의 MBTI는 INTP이다. 나는 족구를(를) 좋아한다. 중국어 한국어	상		
13011	13012	Iory	남	30대	나는 군인이다. 나의 MBTI는 ENFP이다. 나는 농구를(를) 좋아한다. 중국어 한국어	상		
20266	20267	Leam	남	50대	나는 약사이다. 나의 MBTI는 ENFP이다. 나는 카누를(를) 좋아한다. 중국어 한국어	상		
49793	49794	Yader	남	30대	나는 군인이다. 나의 MBTI는 ENFP이다. 나는 수영을(를) 좋아한다. 중국어 한국어	상		

[Backend]

- Collect and store user data with sign-up.
- Login authentication/authorization via jwt cookie issued.
- Logout process.
- Multiple endpoint settings.
- Distribute user lists through render-template.

[Frontend]



Design for the authentication and sign up, log in.

Reflect critically on any goals not accomplished.

[Recommendation System]

- When MBTI enters the hobby, it seems that more meaningful results will come out if it is recommended considering MBTI compatibility, but considering its reliability and unwanted users, only natural language similarity has been applied for now. We will decide again whether to apply this later.

[Backend]

- Postpone interface with Modeling.
- data management via reference keys.

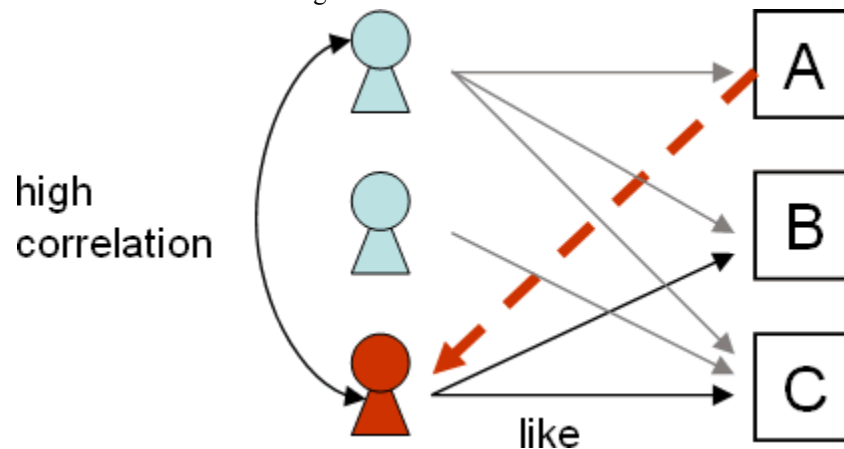
[Frontend]

- The design needs improvement.
- Improvement in user affinity should be made.

What are the goals for the next two weeks?

[Recommendation system]

- will apply an item collaborative filter with a rating function to reflect the taste of the user
 - and will create rating data (like, dislike, NA)
 - after rating the 10 recommended users above, there will be users(A,B,C) with high ratings. Re-recommend other users(A) selected by the user who gave that user(A) a high rating and exclude users with low ratings.



[Data Preprocessing]

- The data pre-processing process to be used in the service seems to have been roughly completed, so we will consider how to run the recommendation service more efficiently using this data.

[Backend]

- Chatting function with translation API.
- Friend Request/Accept function.
- Find face-talk, and audio translation features.

[Frontend]

- Develop design in more detail
- Improve code

How many hours were spent on each goal noted above?

Recommendation system

- 20 hours per two weeks

Data preprocessing

- 5 hours per two weeks

Backend

- 5 hours per two weeks

Frontend

- 5 hours per two weeks