
졸업 프로젝트 최종 발표

- 일기 감정 기반 영화 추천 시스템 -

김예지 201814246

박인우 201512265

손소영 201711649

(조장) 정창해 201710304

CONTENTS

01 프로젝트 소개

02 팀원 소개 및 역할

03 모델 설명

04 앱 및 서버 설명

05 확장 가능성

01 프로젝트 소개

01 프로젝트 소개

02. 팀원 소개 및 역할

03. 모델 설명

04. 앱 및 서버 설명

05. 확장 가능성

개발 배경

- 내면의 감정을 억압하는 게 일상적인 사회
→ 자신의 감정을 이해하고 표출할 줄 아는 것이 정신 건강에 이로움
➡ 일기의 감정을 분석하여 그 날의 감정을 이해할 수 있도록 하자!
- 감정을 표출하는 방법에는 일기 작성, 문화 콘텐츠 시청 등이 있음
- OTT 플랫폼이 활성화된 오늘날 영화 시청에 시공간적 제약이 없어지면서
보다 많은 사람들이 여가 시간에 영화를 시청
➡ 일기의 감정을 분석하여 그 날의 감정에 맞는 영화를 추천해주자!

01 프로젝트 소개

01 프로젝트 소개

02. 팀원 소개 및 역할

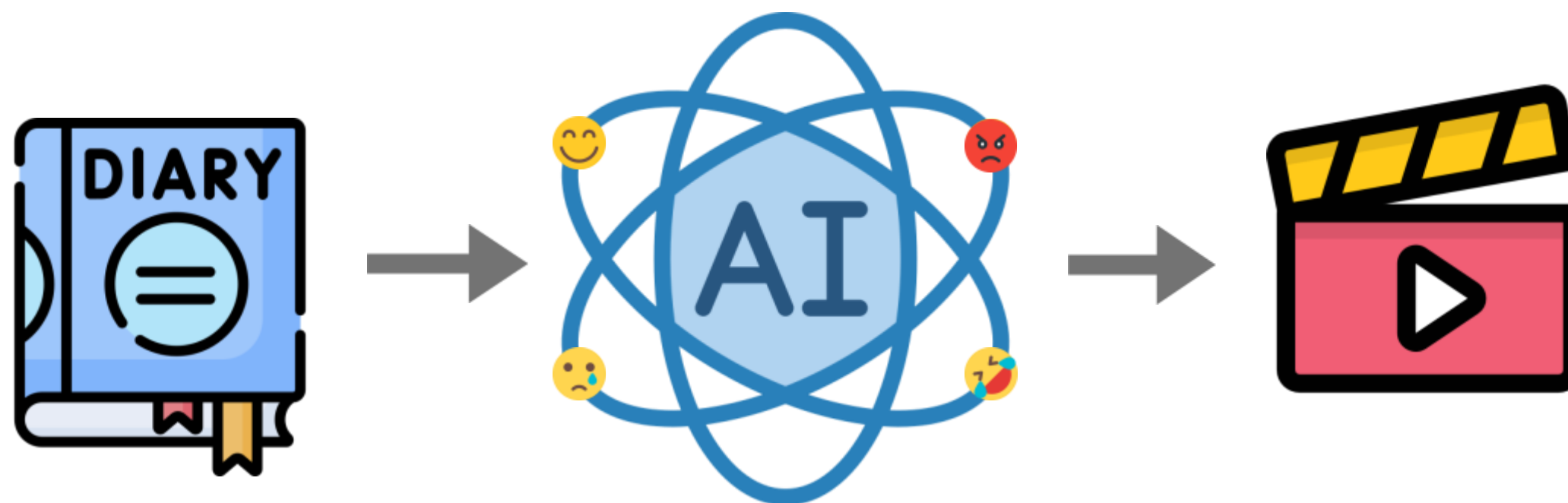
03. 모델 설명

04. 앱 및 서버 설명

05. 확장 가능성

주제 소개

사용자가 작성한 일기의 감정을 분석하여 영화를 추천해주는 iOS 애플리케이션



01 프로젝트 소개

01 프로젝트 소개

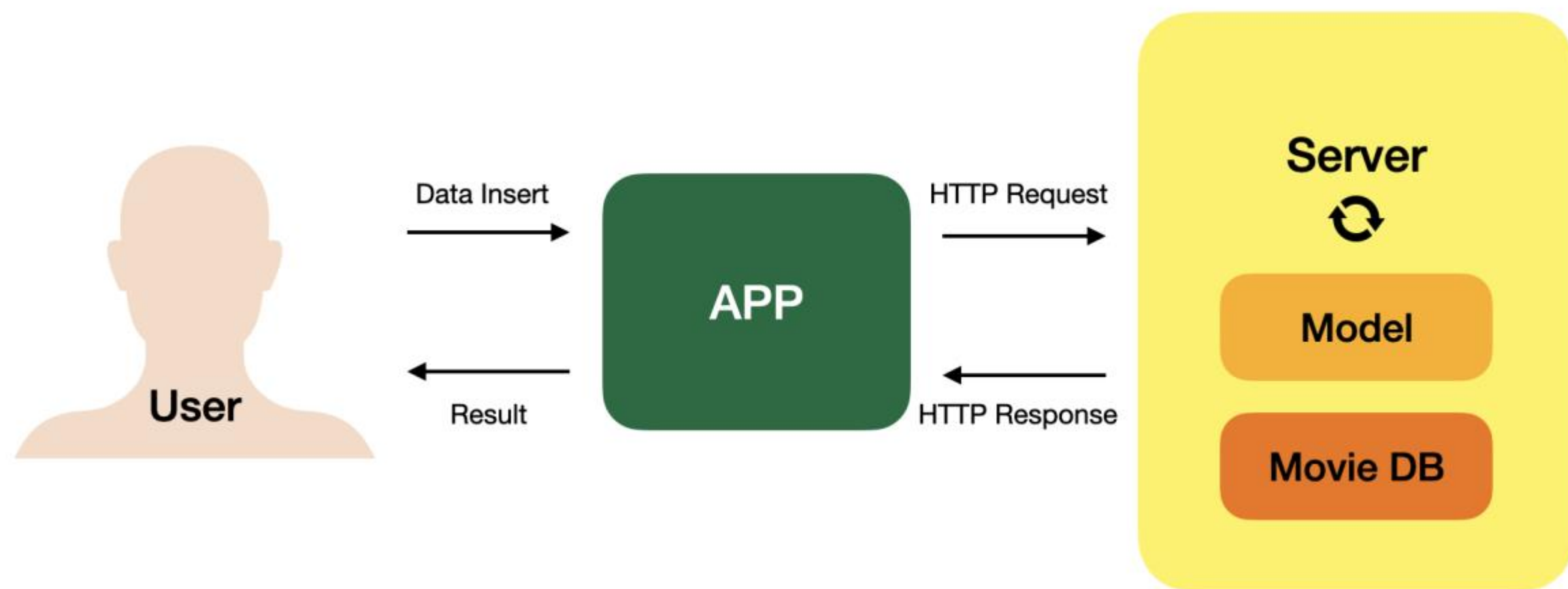
02. 팀원 소개 및 역할

03. 모델 설명

04. 앱 및 서버 설명

05. 확장 가능성

시스템 아키텍처



- 앱에서 사용자가 일기를 입력하면 서버에서 사전에 학습된 모델을 이용해 일기의 감정 벡터 반환
- 이를 미리 모델에 돌려서 db에 저장한 영화 벡터와 유사도를 계산하여 감정 확률과 영화 추천 결과를 반환

02 팀원 소개 및 역할

01. 프로젝트 소개

02 팀원 소개 및 역할

03. 모델 설명

04. 앱 및 서버 설명

05. 확장 가능성

팀원 소개 및 역할



김예지

감정 분석 모델 설계 및
데이터 라벨링



손소영

감정 분석 모델 설계 및 구현



박인우

iOS 앱 설계 및 개발
서버 개발



정창해

감정 분석 모델 설계 및 구현

03 모델 설명

01. 프로젝트 소개

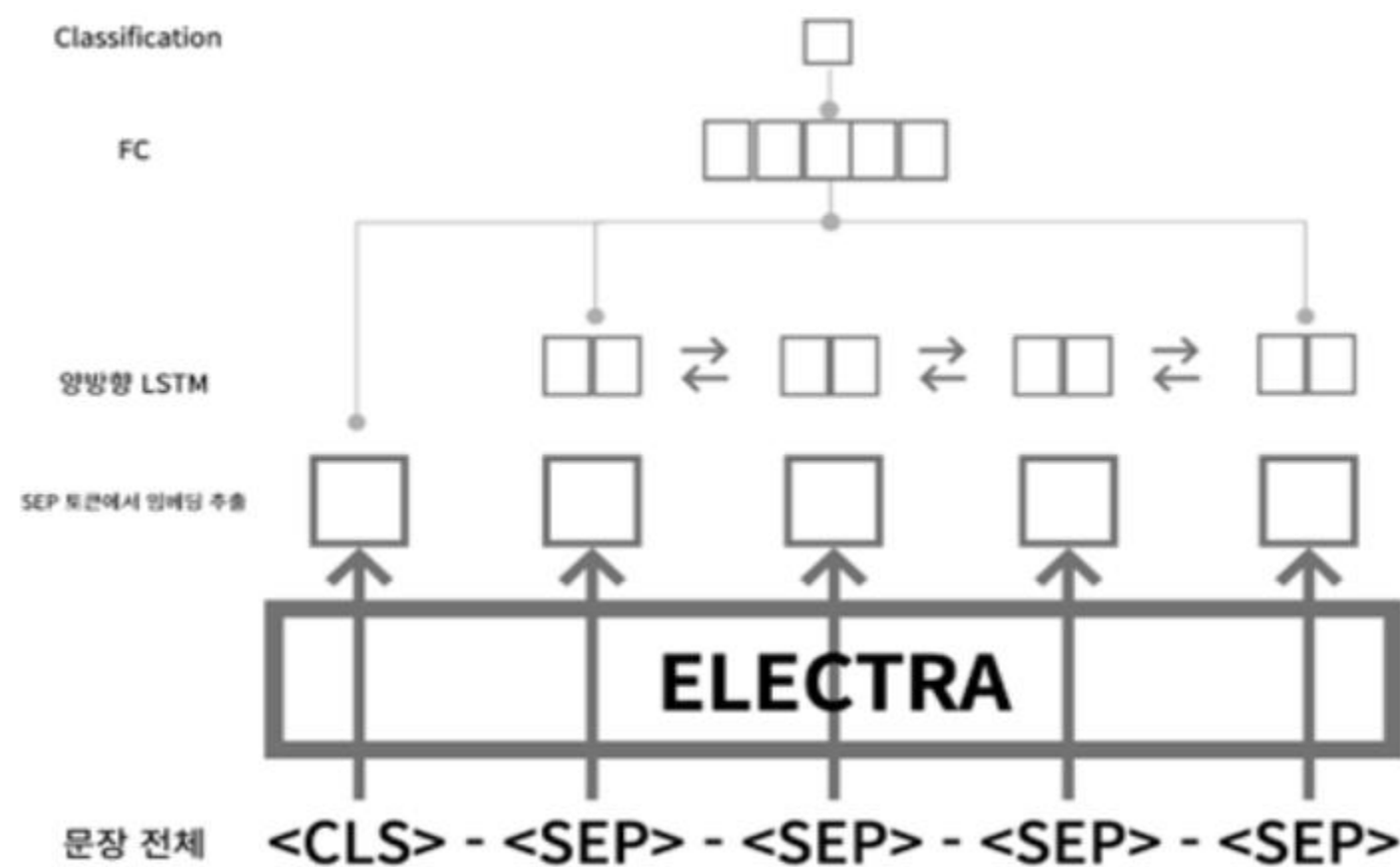
02. 팀원 소개 및 역할

03 모델 설명

04. 앱 및 서버 설명

05. 확장 가능성

감정 분석 모델 - ELECTRA + LSTM



- ELECTRA: 사전훈련된 언어모델(PLM)로, BERT의 파생모델 중 하나
- 일기의 문장의 흐름을 반영하기 위해 bi-LSTM 레이어를 추가
- KSS(Korean Sentence Split)를 사용하여 문장 사이에 <SEP> 토큰 추가하고, 첫번째 문장과 마지막 문장의 <SEP> 토큰에서 추출한 임베딩들을 합쳐 bi-LSTM에 넣음
- <CLS> 토큰과 첫 문장과 마지막 문장의 <SEP> 토큰에서 계산된 loss를 합쳐서 학습 진행

03 모델 설명

01. 프로젝트 소개

02. 팀원 소개 및 역할

03 모델 설명

04. 앱 및 서버 설명

05. 확장 가능성

감정 분석 모델 - Fine Tuning

- 앱에 작성하는 일기의 특성 상 pre-train에 사용된 언어는 실생활에서 많이 사용되고, 온라인에서 많이 사용되는 단어들이어야 함
 - > 한국어 댓글로 pre-train한 kcELECTRA 모델을 사용
- 감정 분류를 하기 위해 문장을 44개의 감정으로 multi-labeling 한 KOTE 데이터셋을 사용
- 각 문장의 감정을 모델에 넣어 silver label을 생성한 후 직접 라벨링을 하여 gold label을 만듦

KOTE (Korean Online That-gul Emotions) Dataset

paper

<https://arxiv.org/pdf/2205.05300.pdf>

dataset

- 다양한 플랫폼에서 수집한 50,000개의 댓글에 44개 정서로 레이블링한 데이터셋.
 - 한 댓글 당 5명이 레이블링 --> 25만 케이스
 - 이것 저것 해 보기 좋은 25만 케이스에 관한 raw data: [raw.json](#)
 - 정서 레이블 종류 [불평/불만, '환영/호의', '감동/감탄', '지긋지긋', '고마움', '슬픔', '화남/분노', '존경', '기대감', '우울함/무시함', '안타까움/실망', '비장함', '의심/불신', '뿌듯함', '편안/쾌적', '신기함/관심', '아껴주는', '부끄러움', '공포/무서움', '절망', '한심함', '역겨움/징그러움', '짜증', '어이없음', '없음', '패배/자기혐오', '귀찮음', '힘듦/지침', '즐거움/신남', '깨달음', '죄책감', '증오/혐오', '호뭇함(귀여움/예쁨)', '당황/난처', '경악', '부담/안_내김', '서러움', '재미없음', '불쌍함/연민', '놀람', '행복', '불안/걱정', '기쁨', '안심/신뢰']

03 모델 설명

01. 프로젝트 소개

02. 팀원 소개 및 역할

03 모델 설명

04. 앱 및 서버 설명

05. 확장 가능성

감정 분석 모델 - Cosine Annealing LR Scheduler

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min}) \left(1 + \cos \left(\frac{T_{\text{cur}}}{T_{\max}} \pi \right) \right)$$

- Local Minimum에 빠지는 문제를 해결하기 위해 사용한 스케줄러
- Learning Rate가 Cosine 함수를 따라서 eta_min까지 떨어졌다가 다시 초기 Learning Rate까지 올라옴

03 모델 설명

01. 프로젝트 소개

02. 팀원 소개 및 역할

03 모델 설명

04. 앱 및 서버 설명

05. 확장 가능성

감정 분석 모델 - 실험 결과

모델	micro F1	차이
kcELECTRA	0.57	
kcELECTRA + bi-LSTM	0.62	+0.05
kcELECTRA + bi-LSTM + CosineAnnealingLR Scheduler	0.65	+0.03

- kcELECTRA + bi-LSTM 아키텍처를 적용했을 때 kcELECTRA 단독 모델에 비해 성능이 5% 향상
- 여기에 스케줄러를 Cosine Annealing LR 스케줄러로 변경한 결과 성능이 추가로 3% 향상되어 총 8% 향상

03 모델 설명

01. 프로젝트 소개

02. 팀원 소개 및 역할

03 모델 설명

04. 앱 및 서버 설명

05. 확장 가능성

영화 추천 - 코사인 유사도

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

- 두 벡터가 얼마나 유사한지 나타내는 척도
- 사용자가 작성한 일기와 영화 줄거리의 감정의 유사도를 측정
- 코사인 유사도를 바탕으로 일기의 감정과 제일 유사한 5개의 영화 추천

04 앱 및 서버 설명

01. 프로젝트 소개

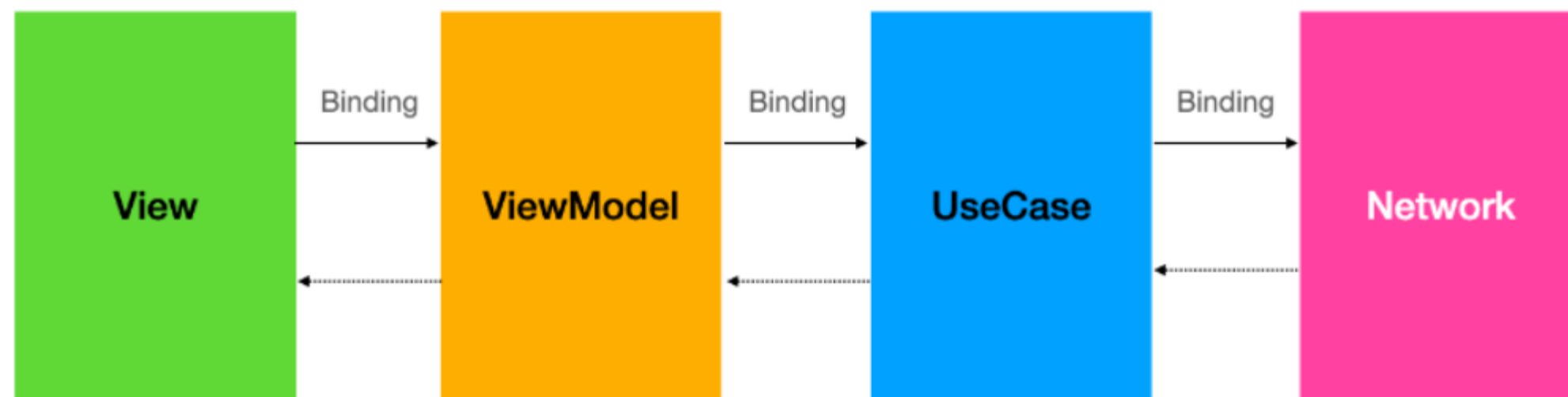
02. 팀원 소개 및 역할

03. 모델 설명

04 앱 및 서버 설명

05. 확장 가능성

앱 및 서버 설명



앱

- MVVM 아키텍처 기반 시스템 구성
- 중간에 UseCase 계층 → 동일한 작업에 대한 중복 코드를 줄임
- Combine을 통한 데이터 Binding → 비동기적으로 데이터를 받아올 수 있는 형태

서버

- Flask를 통해 구현
- 앱에서 데이터를 요청했을 경우 그에 맞는 데이터를 확인하고 보내도록 처리
- 요청은 REST API 형태의 HTTP Request로 구성

05 확장 가능성

01. 프로젝트 소개

02. 팀원 소개 및 역할

03. 모델 설명

04. 앱 설명

05 확장 가능성

확장 가능성

서비스 측면

- 일기로부터 도출된 감정을 기반으로 정신과 상담 연계
- 영화 뿐만 아니라 드라마, 도서 등의 콘텐츠 추천으로 확장

성능 측면

- 데이터셋부터 다시 제작하여 정확도 향상 도모
- 유저의 영화 시청 내역, 리뷰 등 유저 히스토리를 반영하여 추천시스템 고도화
- 모델에 대한 새로운 접근 및 실험

감사합니다
