

ELECTRA 기반 영화평 감성분석

SDS



<T6>



201611309 최지현



201611276 이규은



201612368 이지우



201611251 공민정

김학수

담당교수님 : 김학수 교수님



Contents

High-Level Design

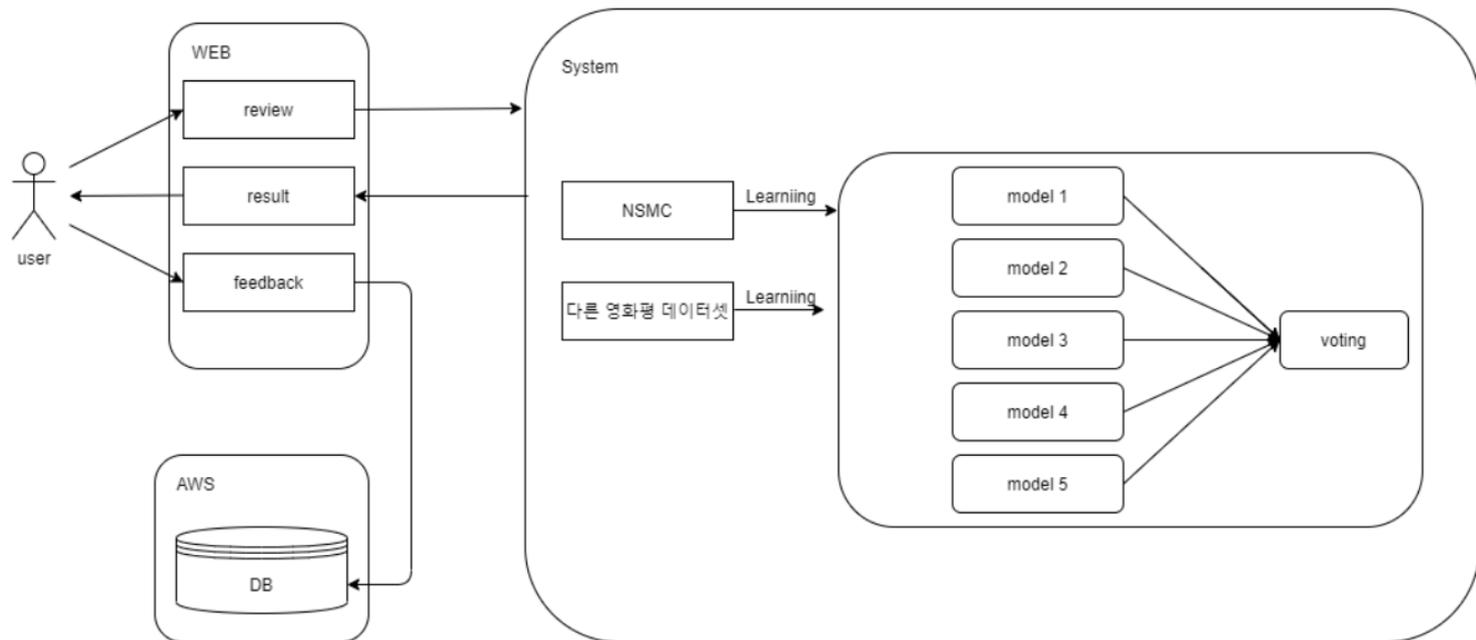
- Architecture Design
- Interface Design
- System Scenario

Low-Level Design

Traceability Matrix



High-Level Design - Architecture Design





High-Level Design - Interface Design

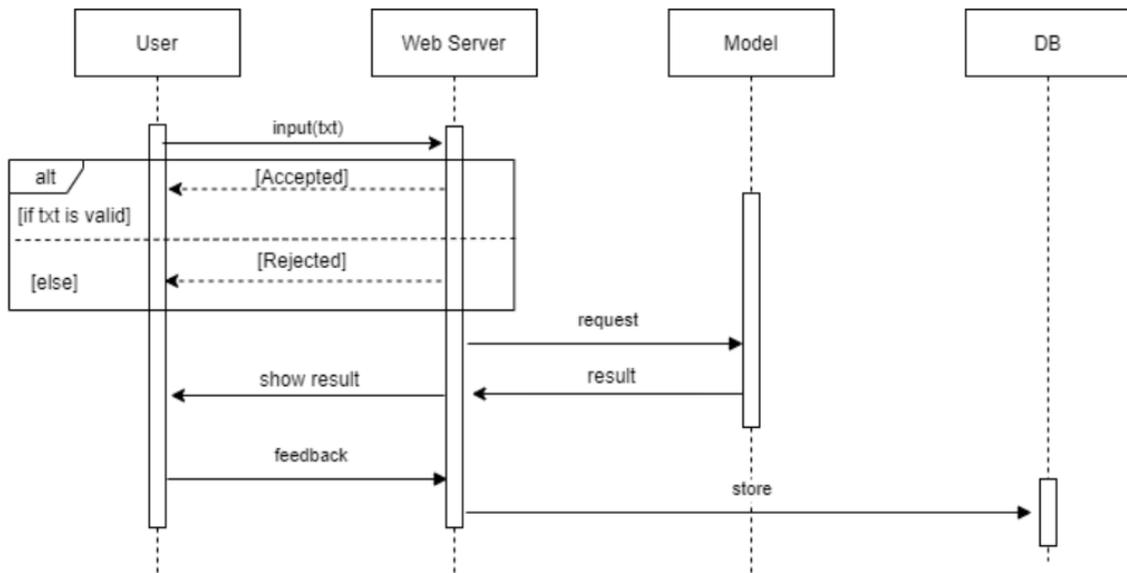
No.	Name	Definition
1	System	감정분석을 위한 전체 모델을 포함하는 시스템
2	Web	사용자의 영화평과 피드백을 입력 받고 결과를 반환하는 웹 인터페이스
3	DB (AWS RDS)	User가 새로 입력한 response 저장과 모델을 버전별로 관리하는 데이터베이스
4	AWS	웹 서버
5	Model 1~5 (Model5 = Additional)	영화평을 학습하기 위한 모델
6	voting	5개의 모델에서 얻은 결과값을 통해 결과 추출

No.	Sender → Receiver	Definition
1	NSMC/다른 영화평 데이터셋 → Model	Model이 NSMC과 다른 영화평 데이터셋의 영화평을 학습한다.
2	Model → voting	5개의 Model의 결과값으로부터 최선의 결과를 얻는다
3	review → System	사용자 영화평이 시스템으로 전달된다.
4	System → result	시스템으로부터 결과를 얻는다.
5	feedback → DB	피드백이 DB에 저장된다.



High-Level Design – System Scenario 1

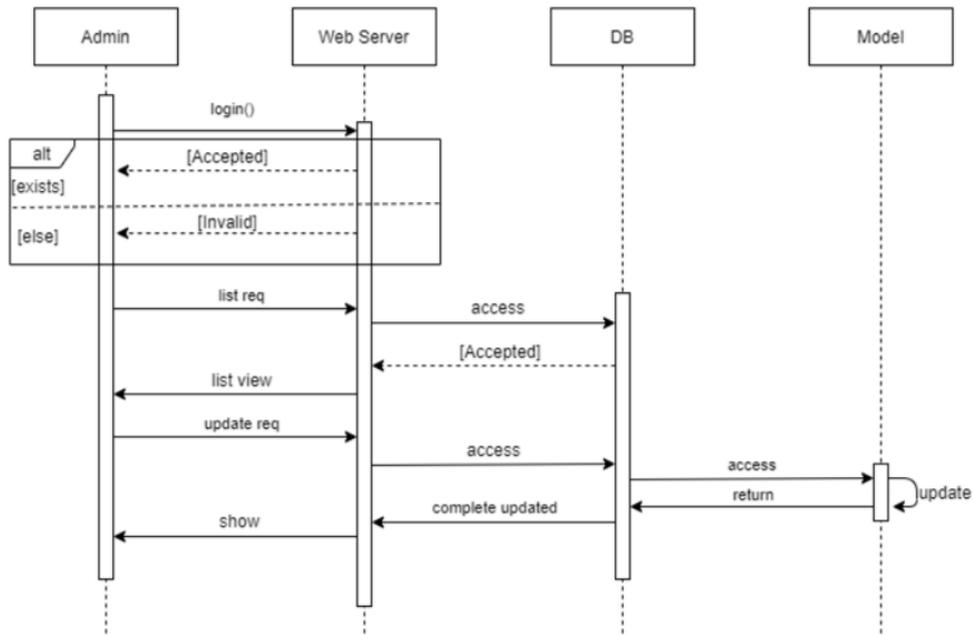
사용자: 영화 평 입력 후 피드백





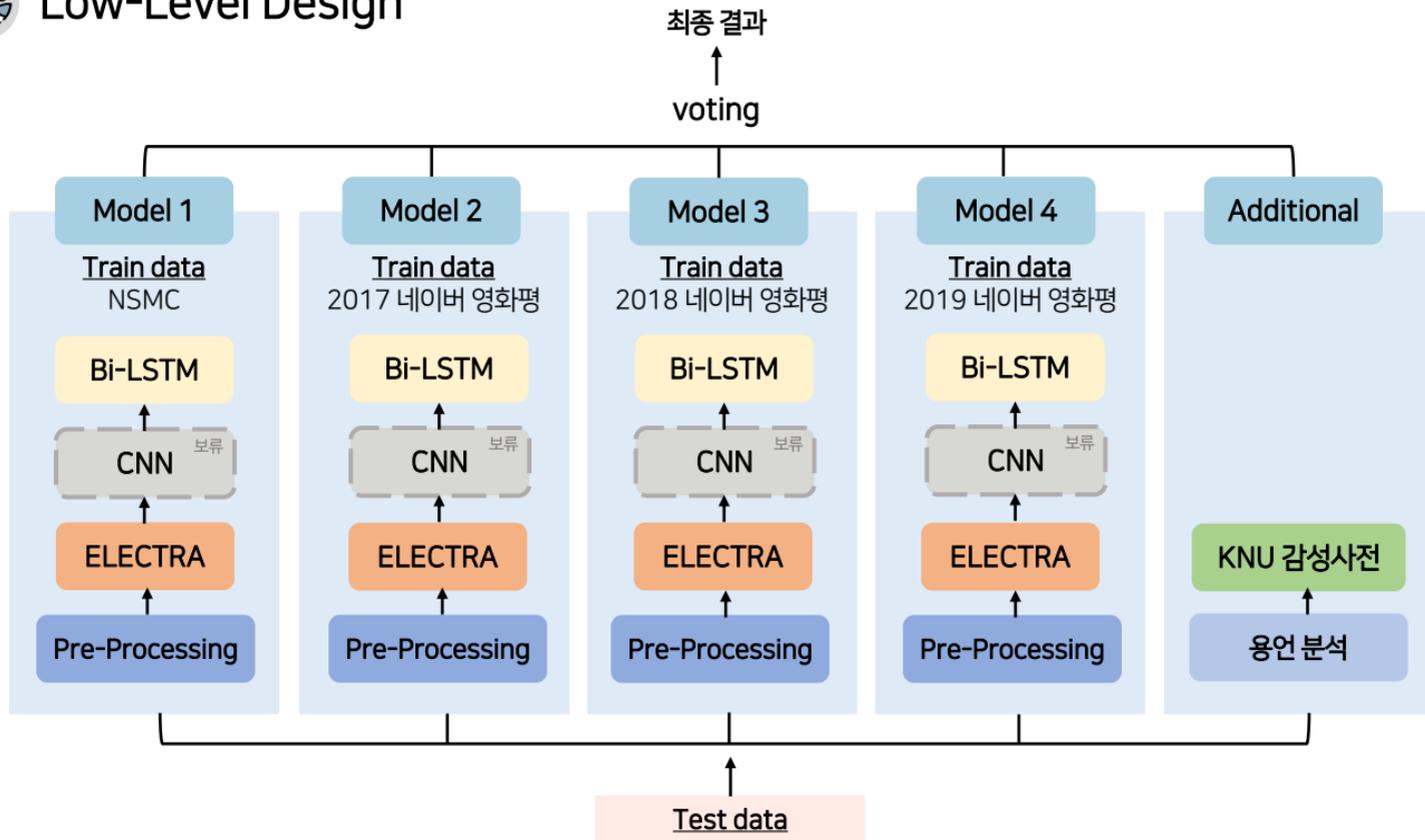
High-Level Design – System Scenario 2

관리자: Model 갱신





Low-Level Design





Traceability Matrix

Requirement	No.
불필요한 데이터를 제거한다	FR 1.1
한글, 영어, 띄어쓰기, 일부 특수 문자를 제외하고 모두 제거한다.	FR 1.2
데이터가 같은 길이가 되도록 한다.	FR 1.3
사용자가 영화평을 입력한다.	FR 3.1
사용자가 입력한 영화평에 대해 감정 분석 결과를 반환한다.	FR 3.2
사용자가 피드백을 한다.	FR 3.3
기존 ELECTRA 단일 모델의 정확도보다 높은 정확도를 기대한다.	NFR 1.1
사용자가 대기할 수 있는 응답 시간을 기대한다.	NFR 2.1
DB를 통한 model quality 향상	NFR 3.1

color match
Requirement ↓

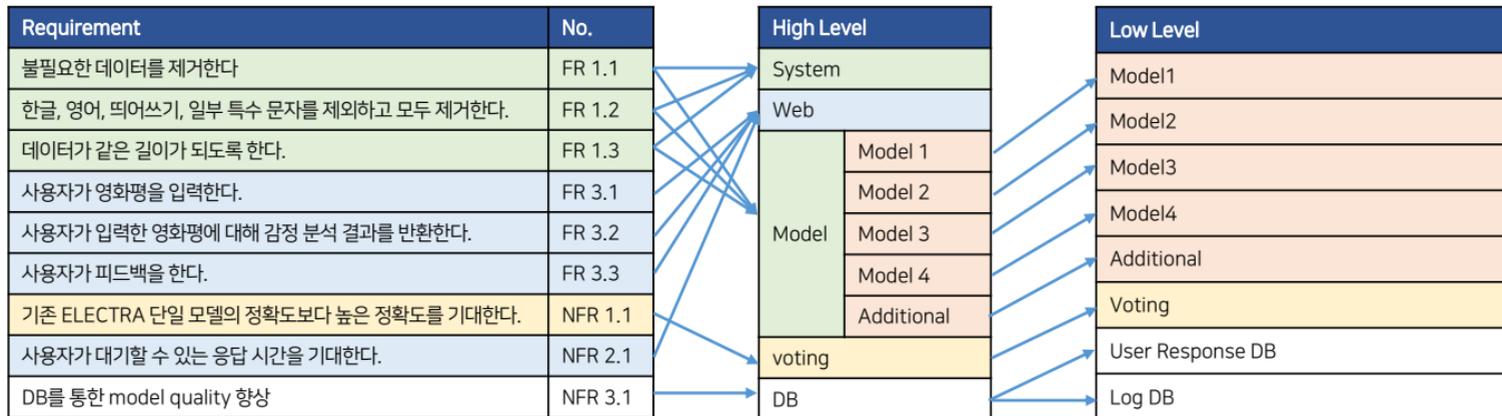
No.	Test Case
1	remove email address
2	trim jQuery Hash String
3	convert words except Korean, English, number to black
4	remove emoji not in unicode
5	padding if data is shorter than designated length
6	input right data through web
7	input black data through web
8	input too long data through web
9	input meaningless data through web
10	return result through web
11	web feedback button
12	input extremely positive data
13	input extremely negative data
14	input neutral data
15	importance of author sentiment than of others
16	importance of present's sentiment
17	check if negated negative is positive
18	check if negated of negative at the end is positive
19	high accuracy
20	short duration
21	additional learning
22	model version control

color match
Success Criteria ↓

Success Criteria
한 영화평의 85% 이상이 어절별로 나누어진다.
긍/부정 값의 정확도가 85% 이상이다.
긍/부정 값이 올바르게 도출된다.
감정 분석 결과의 정확도가 85% 이상이다.
모델의 정확도가 85% 이상이다.
5초 이내에 결과가 도출된다.
추가 학습을 통해 Model의 정확도를 향상시킨다.
정확도가 향상된 Model을 DB에 저장한다.



Traceability Matrix



감사합니다!