

ELECTRA 기반 영화평 감성분석

요구사항 분석서



<T6>

201611251 공민정

201611276 이규은

201611309 최지현

201612368 이지우

담당교수님 : 김학수 교수님



Requirements

Functional Requirement

FR 1 데이터 가공

- FR 1.1 입력된 영화평을 어절 별로 나눔
- FR 1.2 불필요한 문장 부호, 반복 제거
- FR 1.3 ELECTRA와 KNU 감성사전을 이용해 분석
- FR 1.4 bi-LSTM을 통해 긍정/부정 값을 산출

FR 2 감정분석 시스템 구현

- FR 2.1 임의의 영화평 입력 시 감정 분석 결과 반환

Non-Functional Requirement

NFR 1 85프로 이상의 정확도 달성

- NFR 1.1 영화평 데이터셋을 이용한 성능 측정

NFR 2 짧은 응답 시간

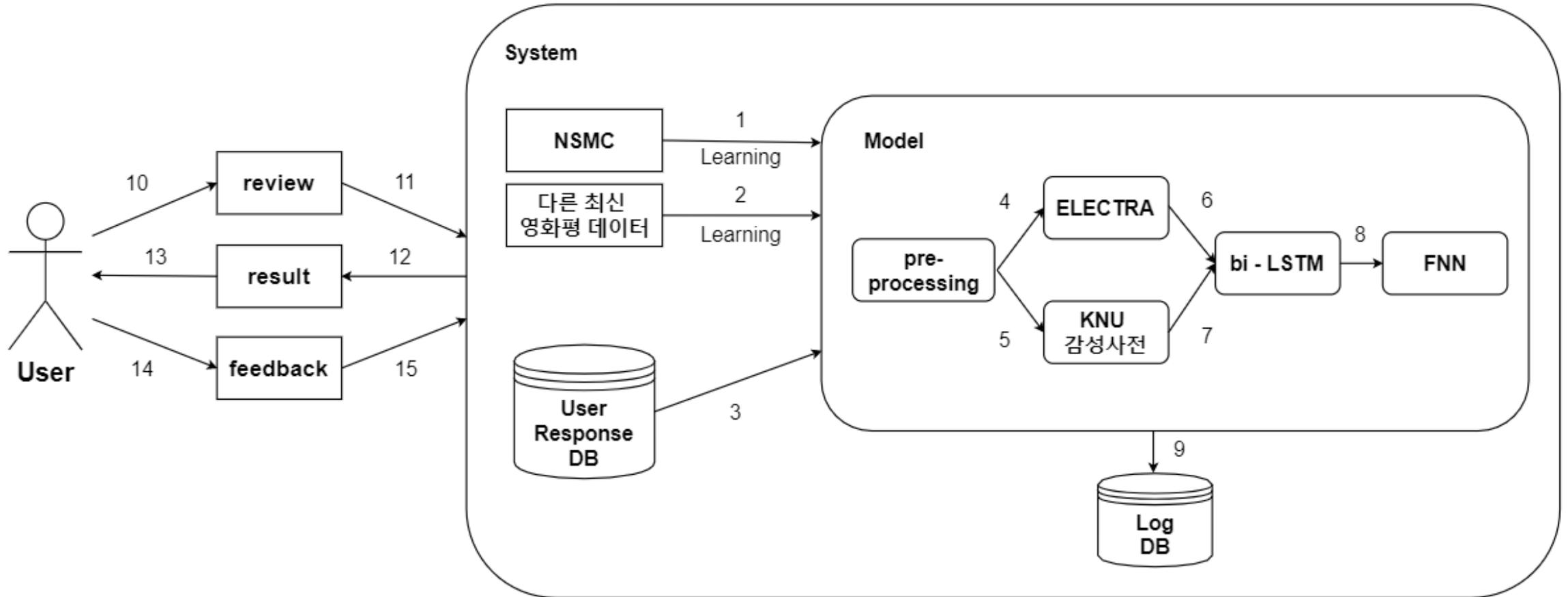
- NFR 2.1 5초 이내에 긍정/부정 결과를 도출하여 사용자에게 보여줌

NFR 3 DB를 통한 Model Quality 향상

- NFR 3.1 User Response DB를 통해 User가 입력한 새로운 영화평을 저장 후 Model 학습
- NFR 3.2 Log DB를 통해 Model을 버전별로 관리

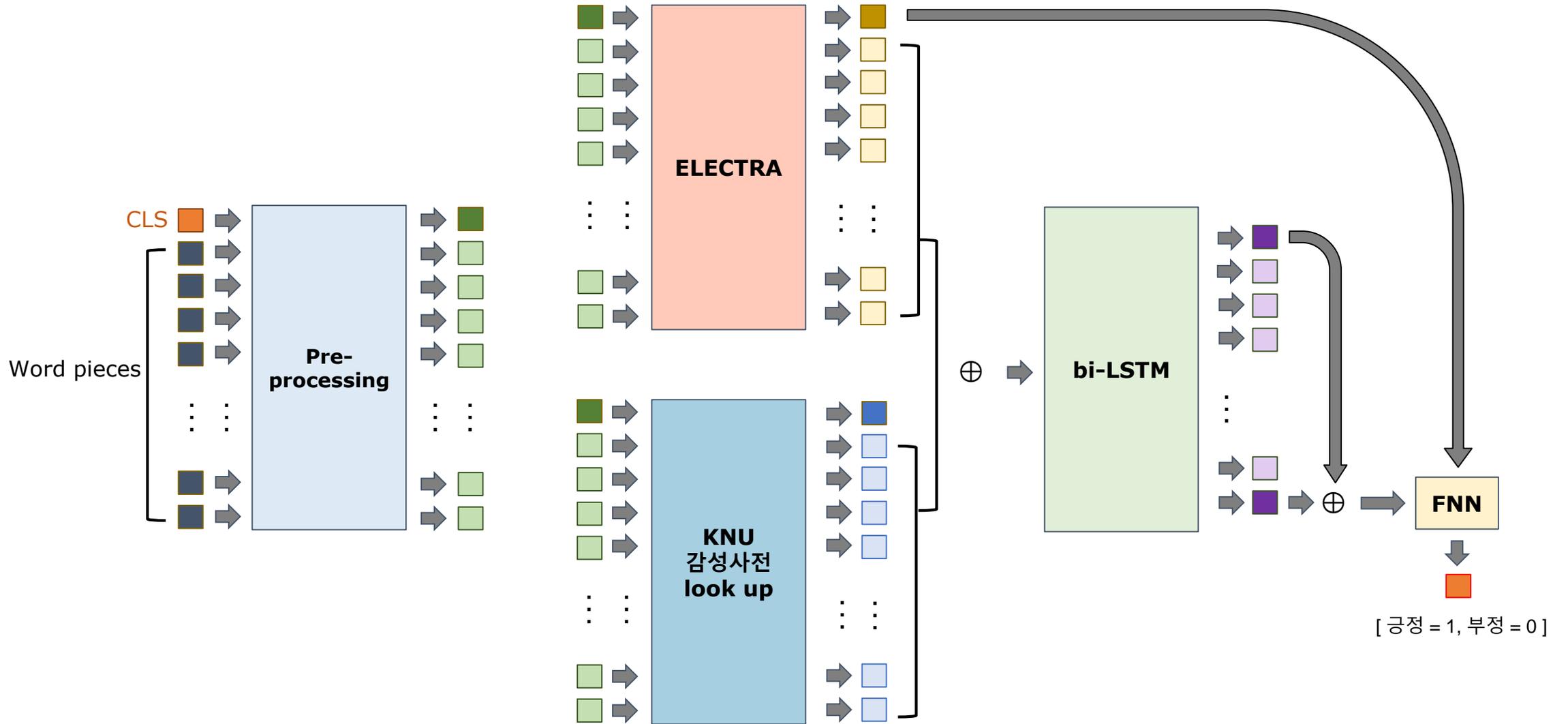


Architecture Diagram





Model Layer Description





Component Description & Requirement

No.	Name	Definition	Requirement No.
1	System	감정분석을 위한 전체 모델을 포함하는 시스템	FR 1 / FR 2
2	NSMC	Naver Sentiment Movie Corpus	FR 1 / FR 2
3	다른 최신 영화평 데이터	성능 개선을 위한 Naver 외 다른 Site의 영화평 데이터	FR 1 / FR 2
4	User Response DB	User가 새로 입력한 response를 저장해두는 database	NFR 3.1
5	Log DB	변화하는 model을 버전별로 관리하는 database	NFR 3.2
6	Model	영화평을 학습하기 위한 model	FR 1
7	pre-processing	데이터 전처리	FR 1.1
8	ELECTRA	token(→real/fake)을 통해 학습하는 모델	FR 1.2
9	KNU 감성사전	각 단어의 입력을 통해 긍/부정 정도값 출력	FR 1.3
10	bi-LSTM	영화평 전체의 문맥 분석을 통해 긍/부정 결과를 도출하는 인공신경망	FR 1.4
11	FNN	감정분석 결과값을 위한 인공신경망	FR 2.1
12	User	시스템에 접근하는 사용자	NFR 1.1 / NFR 2.1
13	review	User가 입력하는 영화평	NFR 3.1
14	result	User에게 반환하는 결과값	NFR 1.1 / NFR 2.1
15	feedback	User가 result의 옳고 그름을 판단하여 입력	NFR 3.1



Interface Definition

No.	Sender → Receiver	Definition
1	NSMC → Model	Model이 NSMC의 영화평을 학습한다.
2	다른 최신 영화평 데이터 → Model	Model이 NSMC 이외의 다른 영화평을 학습한다.
3	User Response DB → Model	Model이 User Response DB에 저장된 영화평을 학습한다.
4	pre-processing → ELECTRA	ELECTRA가 전처리한 word pieces로 학습한다.
5	pre-processing → KNU 감성사전	KNU 감성사전으로 word pieces의 긍/부정을 판단한다.
6	ELECTRA → bi-LSTM	학습한 word pieces들의 뭉치를 통해 문맥을 파악한다.
7	KNU 감성사전 → bi-LSTM	판단한 word pieces들의 뭉치를 통해 문맥을 파악한다.
8	bi-LSTM → FNN	문맥을 분석해 결과를 도출한다.
9	Model → Log DB	Model을 버전 별로 나누어 저장한다.

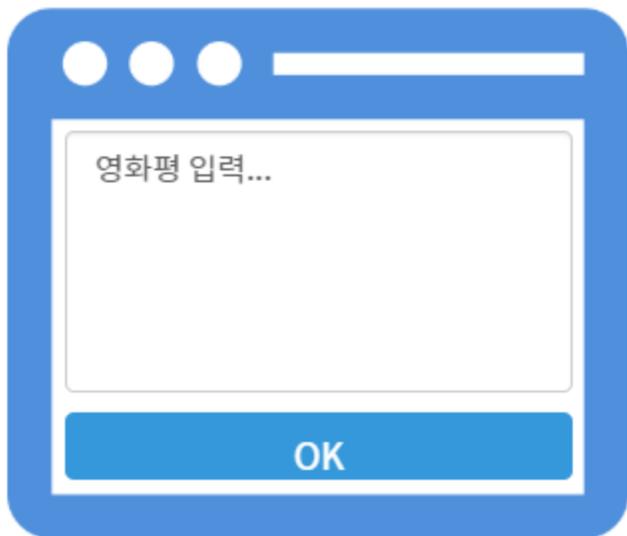


Interface Definition

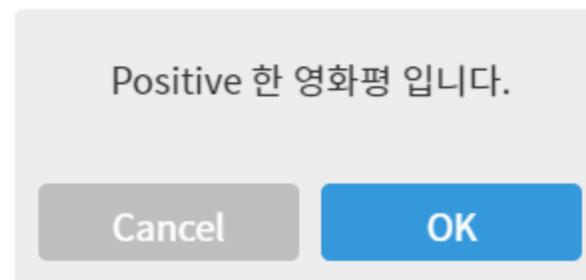
No.	Sender → Receiver	Definition
10	User → review	User가 영화평을 작성한다.
11	review → System	User가 작성한 영화평이 System으로 전달된다.
12	System → result	System이 분석한 영화평의 감정 결과값을 반환한다.
13	result → User	결과값이 User에게 전달된다.
14	User → feedback	User의 감정분석 result에 대한 feedback이 입력된다.
15	feedback → System	User의 feedback이 System으로 전달된다.



Prototype Illustration



<Review>
User가 영화평 입력



<Feedback >
올바른 결과의 경우, OK 버튼
올바르지 않은 결과의 경우, Cancel 버튼



System Test Case

No.	Description	Success Criteria	Requirement No.
1	입력된 영화평이 어절별로 나뉜다.	한 영화평의 85% 이상이 어절별로 나누어진다.	FR 1.1
2	ELECTRA를 통해 분석 결과가 도출된다.	긍/부정 값의 정확도가 85% 이상이다.	FR 1.2
3	KNU 감성사전을 이용해 분석 결과가 도출된다.	각 wordpiece의 긍/부정 값이 감성사전에 정의된 대로 도출된다.	FR 1.3
4	bi-LSTM을 통해 결과가 도출된다.	긍/부정 값이 올바르게 도출된다.	FR 1.4
5	임의의 영화평 입력 시, 감정 분석 결과가 반환된다.	감정 분석 결과의 정확도가 85% 이상이다.	FR 2.1
6	ELECTRA 단일 모델보다 높은 정확도를 달성한다.	모델의 정확도가 85% 이상이다.	NFR 1.1
7	사용자가 영화평을 입력한 후, 짧은 시간내에 결과가 도출된다.	5초 이내에 결과가 도출된다.	NFR 2.1
8	User Response DB에 저장된 영화평으로 추가 학습을 한다.	추가 학습을 통해 Model의 정확도를 향상시킨다.	NFR 3.1
9	Log DB에 각 Model을 버전별로 저장한다.	정확도가 향상된 Model을 Log DB에 저장한다.	NFR 3.2



감사합니다