

# System Test Plan

## for ELECTRA 기반 영화평 감성분석 System

Project Team

**Team 6 (RACCOON)**

Date

**2020-09-15**

---

### Team Information

201611251 공민정

201611276 이규은

201611309 최지현

201612368 이지우

# Table of Contents

- 1 Introduction
  - 1.1 Purpose
  - 1.2 Objective
- 2 System Case
  - 2.1 Traceability
  - 2.2 System Cases

## 1 Introduction

### 1.1 Purpose

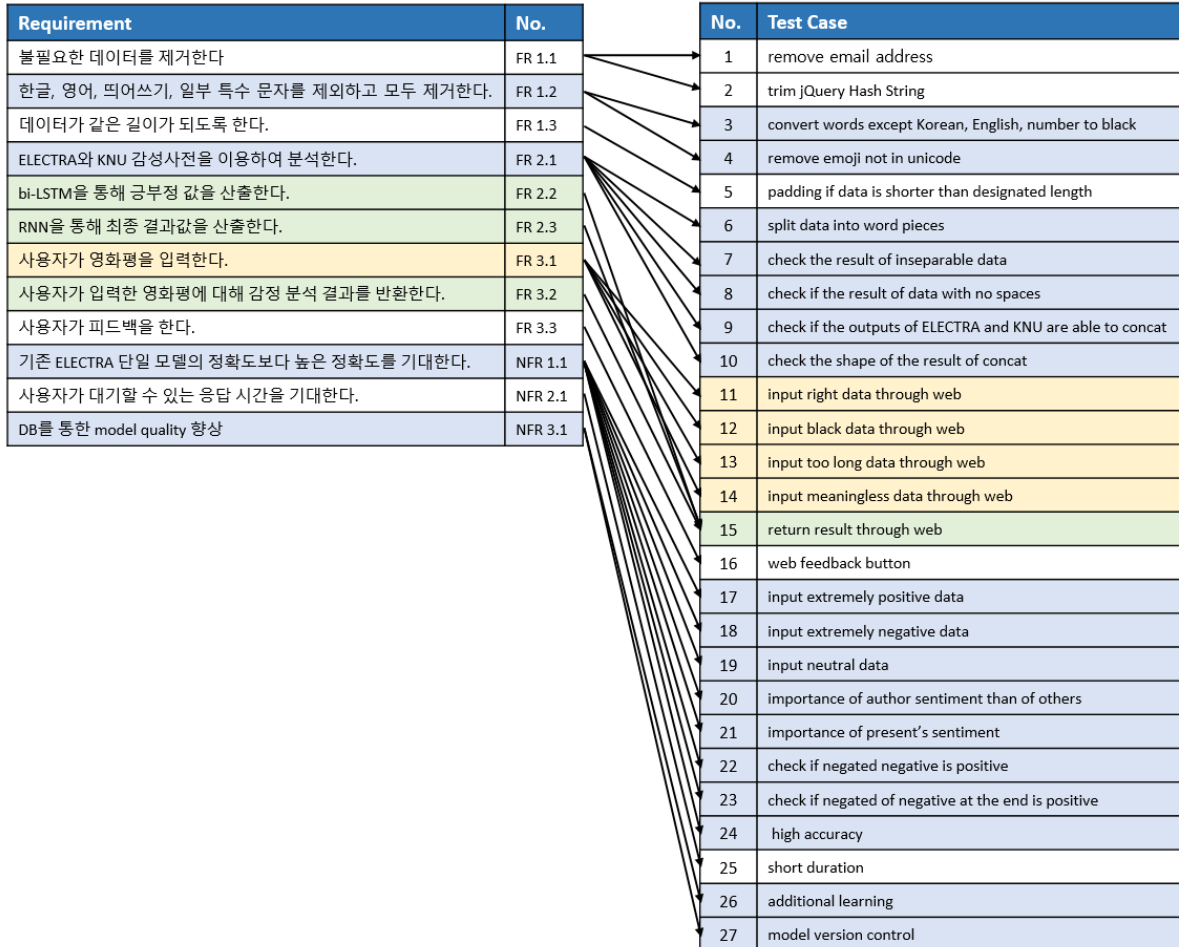
본 문서는 'ELECTRA를 이용한 한국어 영화평 감정분석 시스템' 프로젝트에 대한 시스템 테스트 계획서이다.

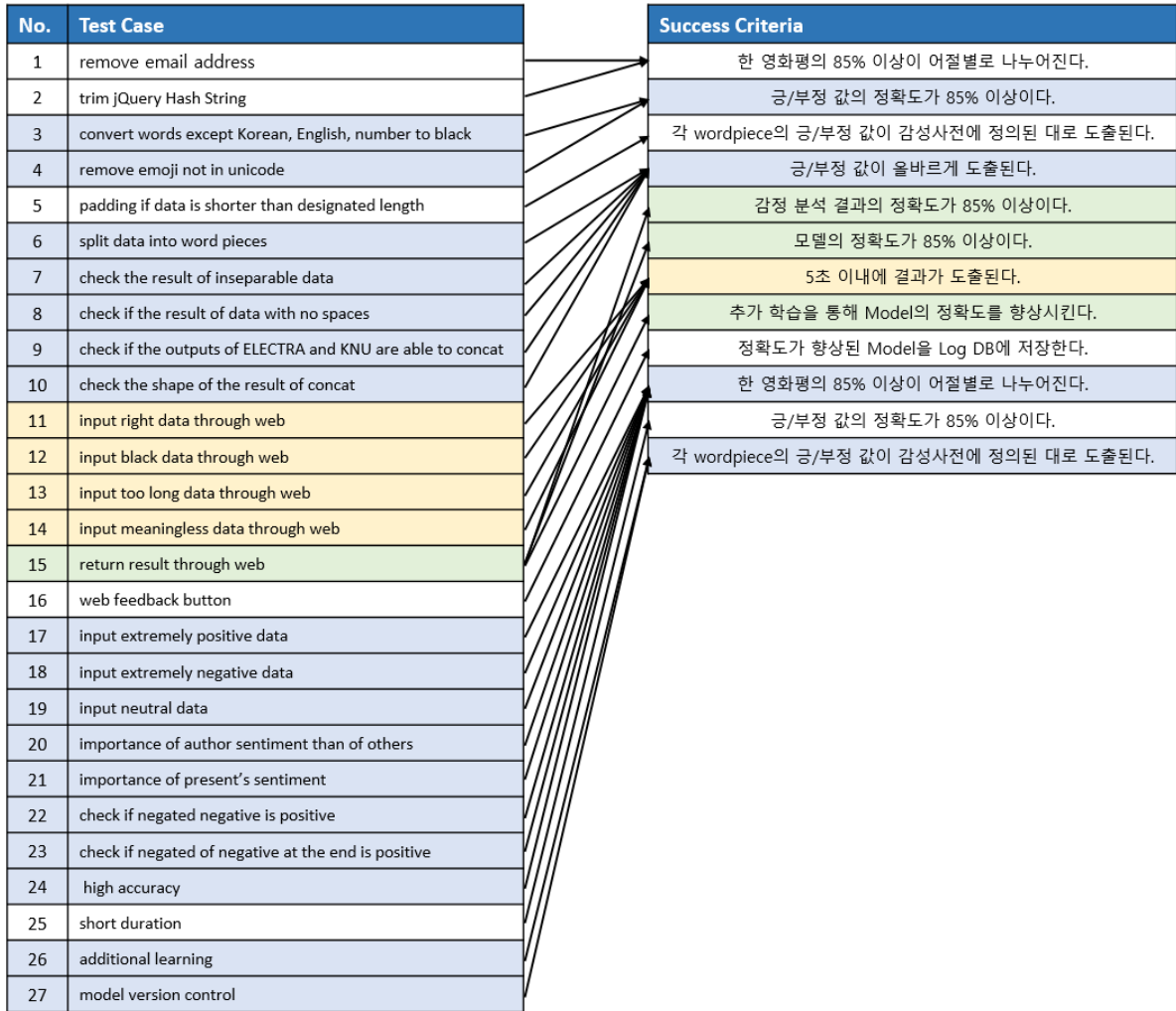
### 1.2 Objective

ICLR 2020 에서 Google 이 학습 효율을 향상시키기 위해 Replaced Token Detection (RTD)이라는 새로운 pre-training 방식을 사용하는 언어 Model 인 ELECTRA 를 발표하였다. ELECTRA 는 기존 BERT 를 비롯한 언어 Model 과 비교해 동일 조건에서의 성능 및 학습 효율이 높다. 따라서 이를 기반으로 하는 시스템을 제안한다. 본 시스템 'Sentiment Analysis System Using Movie Review Corpus(영화평 말뭉치를 사용한 감성 분석 시스템)'은 한국어 문장에 대한 긍정-부정을 판단한다. 기존 NSMC(Naver Sentiment Movie Corpus)데이터와 자체 제작한 말뭉치 데이터를 전처리 한 뒤, ELECTRA(Efficiently Learning an Encoder that Classifies Token Replacements Accurately) Model 와 KNU 감성사전 결과의 가중치 합을 수행하여 벡터를 만든다. 이 벡터를 Bi-LSTM 과 RNN 을 거쳐 학습한다. 본 시스템은 관련 연구가 많이 진행되지 않은 ELECTRA 언어 모델을 사용한다. 또한 각 단어의 긍정-부정을 5 가지의 척도(매우 부정/부정/중립/긍정 /매우 긍정)로 나타내는 한국어 감성사전을 사용해 한국어 친화적인 모델을 제안한다. 단순 문맥 표현을 사용하는 모델에서 단어의 긍정-부정 스코어를 더해 모델링을 수행함으로 기존 연구보다 감성 분석에 특화된 시스템을 기대한다

## 2 System Test case

### 2.1 System test case Traceability





## 2.2 System Case

## 2.2.1 Pre-processing Data

<b>Case No.</b>	1
<b>Name</b>	Remove email address
<b>Description</b>	이메일(ex@gmail.com) 제거한다.
<b>Input</b>	영화평
<b>Output</b>	이메일이 제거된 영화평
<b>Success Criteria</b>	이메일이 포함된 영화평이 존재하지 않는다.
<b>Reference No.</b>	기능 요구사항 1.1.1

<b>Case No.</b>	2
<b>Name</b>	Trim jQuery Hash String
<b>Description</b>	크롤링 과정에서 발생한 실제 리뷰 이외의 jQuery Hash 문자열을 삭제한다.
<b>Input</b>	영화평
<b>Output</b>	jQuery Hash 등이 제거된 영화평
<b>Success Criteria</b>	영화평에 jQuery Hash등이 포함되어 있지 않는다.
<b>Reference No.</b>	기능 요구사항 1.1.2

<b>Case No.</b>	3
<b>Name</b>	Covert words except Korea, English, number to blank
<b>Description</b>	한글,영어,숫자 이외의 문자를 공백으로 치환

<b>Input</b>	영화평
<b>Output</b>	한글,영어,숫자 이외의 문자가 포함되지 않은 영화평
<b>Success Criteria</b>	한글,영어,숫자 이외의 문자가 존재하지 않는다.
<b>Reference No.</b>	기능 요구사항 1.2.1

<b>Case No.</b>	4
<b>Name</b>	Remove emoji not in unicode
<b>Description</b>	유니코드 이외의 이모티콘은 삭제한다.
<b>Input</b>	유니코드 이외의 이모티콘이 포함된 영화평 (ex. 재밌었어요 😊 )
<b>Output</b>	유니코드 이외의 이모티콘이 삭제된 영화평 (ex. 재밌었어요)
<b>Success Criteria</b>	유나코드 이외의 문자가 포함된 영화평이 존재하지 않는다.
<b>Reference No.</b>	기능 요구사항 1.2.2

<b>Case No.</b>	5
<b>Name</b>	Padding if data is shorter than designated length
<b>Description</b>	tokenize된 영화평이 설정한 길이를 가지도록 padding한다.
<b>Input</b>	tokenize된 영화평
<b>Output</b>	config에 설정한 길이가 되도록 padding이 더해진 tokenize된 영화평
<b>Success Criteria</b>	지정한 길이로 영화평을 자른다.
<b>Reference No.</b>	기능 요구사항 1.3.1

## 2.2.2 Process ELECTRA

<b>Case No.</b>	6
<b>Name</b>	Split data into word pieces
<b>Description</b>	입력한 영화평이 제대로 잘리는지 확인한다.
<b>Input</b>	영화평 (ex. 너무 재밌어요)
<b>Output</b>	word piece (ex. ['너무', '재밌어요'] )
<b>Success Criteria</b>	입력한 영화평이 알맞은 word piece로 잘려져 있는지 확인한다.
<b>Reference No.</b>	기능 요구사항 2.1.1

<b>Case No.</b>	7
<b>Name</b>	Check the result of inseparable data
<b>Description</b>	단어 하나 등 나눌 수 없는 데이터가 알맞은 word piece로 분리되는지 확인한다.
<b>Input</b>	더 이상 나눌 수 없는 단어가 포함된 데이터 (ex. 감동)
<b>Output</b>	word piece (ex. ['감동'] )
<b>Success Criteria</b>	더 이상 나눌 수 없는 데이터가 나뉘어 지지 않은 상태의 word piece로 되는지 확인한다.
<b>Reference No.</b>	기능 요구사항 2.1.1

<b>Case No.</b>	8
<b>Name</b>	Check if the result of data with no spaces



<b>Description</b>	띄어쓰기가 안된 영화평에 대해 적절한 결과가 나오는지 확인한다.
<b>Input</b>	띄어쓰기가 안된 영화평 (ex. 영화너무재밌어요)
<b>Output</b>	긍정 또는 부정 (ex. 긍정)
<b>Success Criteria</b>	띄어쓰기가 안된 영화평에 대해 수용 가능한 결과가 나오는지 확인한다.
<b>Reference No.</b>	기능 요구사항 2.1.1

<b>Case No.</b>	9
<b>Name</b>	Check if the outputs of ELECTRA and KNU are able to concat
<b>Description</b>	ELECTRA와 KNU의 output이 concat할 수 있는 형태인지 확인한다.
<b>Input</b>	동일한 영화평
<b>Output</b>	ELECTRA와 KNU 감성사전 각각의 output의 shape
<b>Success Criteria</b>	ELECTRA와 KNU의 output이 concat이 가능한 형태인지 확인한다
<b>Reference No.</b>	기능 요구사항 2.1.2

<b>Case No.</b>	10
<b>Name</b>	Check the shape of the result of concat
<b>Description</b>	ELECTRA와 KNU의 concat 결과의 형태를 확인한다.
<b>Input</b>	ELECTRA와 KNU의 concat 결과
<b>Output</b>	알맞은 ELECTA output과 KNU 감성사전 output의 mapping vector

<b>Success Criteria</b>	ELECTRA와 KNU의 concat 결과의 형태가 맞는지 확인한다.
<b>Reference No.</b>	기능 요구사항 2.1.2

## 2.2.3 System Implementation

<b>Case No.</b>	11
<b>Name</b>	Input right data through web
<b>Description</b>	웹을 통해 정상적인 영화평을 입력한다.
<b>Input</b>	영화평
<b>Output</b>	긍/부정 결과
<b>Success Criteria</b>	지정된 길이를 초과하지 않은 영화평을 입력했을 때 정상적으로 작동하는지 확인한다.
<b>Reference No.</b>	기능 요구사항 3.1.1

<b>Case No.</b>	12
<b>Name</b>	Input blank data through web
<b>Description</b>	웹에서 영화평을 입력하지 않은 채로 결과를 확인해본다
<b>Input</b>	아무것도 입력하지 않음
<b>Output</b>	-
<b>Success Criteria</b>	아무것도 입력하지 않고 결과를 알려고 할 때 영화평을 입력하라는 경고 표시가 나온다.
<b>Reference No.</b>	기능 요구사항 3.1.1

<b>Case No.</b>	13
<b>Name</b>	Input too long data through web
<b>Description</b>	지정된 길이를 초과하는 영화평을 입력해본다.

<b>Input</b>	지정된 길이를 초과한 영화평
<b>Output</b>	-
<b>Success Criteria</b>	지정된 길이를 초과하는 영화평을 입력할 때, 영화평이 너무 길다는 경고 표시가 나온다.
<b>Reference No.</b>	기능 요구사항 3.1.1

<b>Case No.</b>	14
<b>Name</b>	Input meaningless data through web
<b>Description</b>	의미없는 영화평을 입력해본다.
<b>Input</b>	제대로 된 문자가 없는 영화평 (ex. .... , 12345678)
<b>Output</b>	-
<b>Success Criteria</b>	의미없는 영화평을 입력할 때, 제대로 된 영화평을 입력하라는 경고 표시가 나온다.
<b>Reference No.</b>	기능 요구사항 3.1.1

<b>Case No.</b>	15
<b>Name</b>	Return result through web
<b>Description</b>	웹을 통해 결과가 반환된다.
<b>Input</b>	영화평
<b>Output</b>	긍/부정 결과
<b>Success Criteria</b>	웹을 통해 결과를 확인할 수 있다.
<b>Reference No.</b>	기능 요구사항 3.2.1

<b>Case No.</b>	16
<b>Name</b>	Web Feedback Button
<b>Description</b>	웹의 피드백 버튼이 정상적으로 작동하는지 확인한다.
<b>Input</b>	영화평
<b>Output</b>	금/부정 결과
<b>Success Criteria</b>	웹의 피드백 버튼이 정상 작동하는지 확인한다.
<b>Reference No.</b>	기능 요구사항 3.3.1

## 2.2.4 Process Input Data

<b>Case No.</b>	17
<b>Name</b>	Input extremely positive data
<b>Description</b>	극단적으로 긍정적인 영화평을 입력한다.
<b>Input</b>	극단적으로 긍정적인 영화평 (ex. 너무 심각하게 재밌어요.)
<b>Output</b>	긍정
<b>Success Criteria</b>	극단적으로 긍정적인 영화평을 입력 시, 결과가 긍정인지 확인한다.
<b>Reference No.</b>	비기능 요구사항 1.1.1

<b>Case No.</b>	18
<b>Name</b>	Input extremely negative data
<b>Description</b>	극단적으로 부정적인 영화평을 입력한다.
<b>Input</b>	극단적으로 부정적인 영화평 (ex. 너무 재미없는 최악의 영화였습니다..)
<b>Output</b>	부정

<b>Success Criteria</b>	극단적으로 부정적인 영화평을 입력 시, 결과가 부정인지 확인한다.
<b>Reference No.</b>	비기능 요구사항 1.1.1

<b>Case No.</b>	19
<b>Name</b>	Input neutral data
<b>Description</b>	중립적인 영화평을 입력한다.
<b>Input</b>	중립적인 영화평 (ex. 볼만해요..)
<b>Output</b>	긍/부정의 결과 (ex. 긍정)
<b>Success Criteria</b>	중립적인 영화평을 입력했을 때 그에 따른 알맞은 결과가 반환된다.
<b>Reference No.</b>	비기능 요구사항 1.1.1

<b>Case No.</b>	20
<b>Name</b>	Importance of author sentiment than of others
<b>Description</b>	다른 사람들의 의견보다 본인의 의견에 대한 감정 분석 결과를 반환한다.
<b>Input</b>	다른 사람과 글쓴이의 의견이 같이 있는 영화평 (ex. 친구는 재미없다고 하는데 저는 재밌게 봤네요)
<b>Output</b>	긍/부정 결과 (ex. 긍정)
<b>Success Criteria</b>	비기능 요구사항 1.1.1
<b>Reference No.</b>	기능 요구사항 2.1

<b>Case No.</b>	21
<b>Name</b>	Check if negated negative is positive
<b>Description</b>	이중 부정은 긍정을 의미하므로 이를 확인한다.
<b>Input</b>	이중 부정이 포함된 영화평 (ex. 재미없지 않네요)
<b>Output</b>	긍정
<b>Success Criteria</b>	이중 부정이 포함된 영화평의 감정분석 결과가 긍정인지 확인한다.
<b>Reference No.</b>	비기능 요구사항 1.1.1

<b>Case No.</b>	22
<b>Name</b>	Check if negation of negative at the end is positive
<b>Description</b>	문장의 마지막 부분이 부정의 반대일 경우 긍정인지 확인한다.
<b>Input</b>	문장의 마지막이 부정이여서 문장 전체가 이중 부정인 영화평 (ex. 영화가 재미없을 줄 알았는데, 아니네요)
<b>Output</b>	긍정
<b>Success Criteria</b>	문장의 마지막이 부정이여서 문장 전체가 이중 부정인 영화평의 감정분석 결과가 긍정인지 확인한다.
<b>Reference No.</b>	비기능 요구사항 1.1.1

<b>Case No.</b>	23
<b>Name</b>	Importance of present's sentiment than of past's
<b>Description</b>	과거보다 현재의 감정에 초점이 맞춰졌는지 확인한다..

<b>Input</b>	과거와 현재 감정이 같이 나타난 영화평 (ex.. 옛날에는 재미없었는데, 다시보니까 재밌네요)
<b>Output</b>	긍정 또는 부정
<b>Success Criteria</b>	감정 분석 결과가 과거가 아닌 현재의 감정에 대한 것인지 확인한다.
<b>Reference No.</b>	비기능 요구사항 1.1.1

## 2.2.5 High Accuracy

<b>Case No.</b>	24
<b>Name</b>	High Accuracy
<b>Description</b>	ELECTRA 단일 모델보다 높은 정확도를 달성한다.
<b>Input</b>	영화평
<b>Output</b>	긍/부정 결과
<b>Success Criteria</b>	모델 정확도가 85% 이상인지 확인한다.
<b>Reference No.</b>	비기능 요구사항 1.1

## 2.2.6 Short Duration

<b>Case No.</b>	25
<b>Name</b>	Short Duration
<b>Description</b>	사용자가 영화평을 입력한 후, 짧은 시간 내에 결과가 도출된다.
<b>Input</b>	영화평
<b>Output</b>	긍/부정 결과
<b>Success Criteria</b>	모든 종류의 영화평에 대해 5초 이내에 결과가 도출되는지 확인한다.
<b>Reference No.</b>	비기능 요구사항 2.1.1

## 2.2.7 Additional Learning

<b>Case No.</b>	28
<b>Name</b>	Additional Learning
<b>Description</b>	User Response DB에 저장된 영화평으로 추가 학습을 한다.
<b>Input</b>	추가적인 영화평 데이터 셋
<b>Output</b>	추가 학습이 완료된 모델
<b>Success Criteria</b>	추가 학습을 통해 Model의 정확도를 향상시킨다.
<b>Reference No.</b>	비기능 요구사항 3.1.1

## 2.2.8 Model Version Control

<b>Case No.</b>	29
<b>Name</b>	Model Version Control
<b>Description</b>	Log DB에 각 Model을 버전별로 저장한다.
<b>Input</b>	현 모델
<b>Output</b>	버전별로 저장된 모델
<b>Success Criteria</b>	Log DB에 각 Model을 버전별로 저장한다.
<b>Reference No.</b>	비기능 요구사항 3.1.2