

요구사항 명세서

이건뭐사람아니냐고

ELECTRA기반 영화 감상 분석기

TEAM 2

201310737 김주언

201611829 최보라

201712845 류한길

201411283 유병찬

목차

1. 소개 (Introduction)
 - 1.1. 목적 (Purpose)
 - 1.2. 범위 (Scope)
 - 1.3. 용어 및 약어 정의 (Definitions, Acronyms, and Abbreviations)
 - 1.4. 참고 문헌 (References)
 - 1.5. 개요 (Overview)
2. 종합 기술 (Overall description)
 - 2.1. 프로젝트 결과물 전망 (Product perspective)
 - 2.2. 제품 기능 (Product functions)
 - 2.3. 사용자 특성 (User characteristics)
 - 2.4. 제약 조건 (Constraints)
 - 2.5. 가정과 의존성 (Assumption and dependencies)
3. 상세 요구사항 (Specific requirement)
 - 3.1. 사용자 인터페이스 (User Interfaces)
 - 3.2. 기능적 요구사항 (Functional Requirement)
 - 3.2.1. 정보 흐름(Data flows)
 - 3.2.2. 프로세스 기술(Process Descriptions)
4. 추가 정보 (Supporting Information)
 - 4.1. 소프트웨어 요구사항 명세서
 - 4.2. 문서 이력

1. 서론 (Introduction)

1.1 목적 (Purpose)

본 문서는 딥러닝 기반 영화평 감성 분석 소프트웨어 요구사항 명세서 (SRS)이다.

소프트웨어는 건국대학교 컴퓨터공학부 졸업프로젝트 TEAM2 에서 설계 및 구현하기 위한 것으로 이를 위한 요구 사항을 정리, 분석하고, 기재된 내용을 바탕으로 시스템을 설계 및 구현한다.

본 문서는 건국대학교 컴퓨터공학부 졸업프로젝트 TEAM2 과 졸업 프로젝트 지도 교수님을 주요 독자로 한다.

1.2 범위 (Scope)

본 프로젝트의 개발 범위는 다음과 같다.

(기능 1) 사용자로부터 리뷰 텍스트 입력

(기능 2) 입력받은 텍스트 감성 분석(긍정/부정)

(기능 3) 감성 분석 결과를 시각적으로 출력

(기능 4) 사용자에게 감성 분석 결과에 대한 정오 정보를 입력 요청

(기능 5) 입력받은 리뷰 텍스트와 사용자가 입력한 해당 텍스트의 감성(긍정/부정)을 데이터베이스에 저장

1.3 용어 및 약어 정의 (Definitions, Acronyms, and Abbreviations)

리뷰 텍스트 : 사용자가 입력한 영화평 리뷰 데이터

감성 분석 : 모델에 리뷰 텍스트 데이터를 입력하여 긍정 또는 부정의 결과를 얻는 프로세스

ML server : 사전에 학습된 딥러닝 모델을 통해 주어진 영화평의 감성 분석을 수행하는 서버

API server : 클라이언트의 요청을 받아 데이터를 처리하여 제공하는 서버

정오 판단 : 입력한 리뷰 텍스트의 감성 분석 결과에 대해 사용자가 옳고 그름을 판단하는 행위

ELECTRA: 일반적이고 커다란 말뭉치로 미리 학습한 후, 특정한 분야에 대해 미세 조정을 통해 높은 정확도와 빠른 학습시간을 취하는 전이학습을 기반으로 하는 언어 모델

데이터 베이스 : 영화평 리뷰 데이터와 사용자에게 입력받은 정오판단 데이터를 API 서버로부터 입력받아 저장하고 있는 데이터 베이스

word piece: 단어들을 자주 발생하는 subword units으로 나누어 제한된 개수의 subword unit으로 수많은 단어들을 표현하는 방법.

1.4 참고 문헌 (References)

정인범, 서동만, 허난숙, 한우람, 방한민, "멀티미디어 콘텐츠 인식 지능형 정보 제공 서비스 소프트웨어 요구사항 명세서", 강원대학교 소프트웨어 시스템 연구실 멀티미디어 서비스 시스템 연구팀, 2007

1.5 개요 (Overview)

본 문서의 구성은 다음과 같다. 2장에서는 딥러닝 기반 영화평 감성 분석 소프트웨어의 전반적인 개요를 알아본다. 그를 위해 결과물의 향후 전망과 결과물의 세부 기능, 사용자 의 특성, 제약 조건, 가정과 의존성 등의 각 절로 나누어 살펴본다. 3장에서는 딥러닝 기반 영화평 감성 분석 소프트웨어의 상세 요구사항을 알아본다. 외적 인터페이스 요구사항과 기능적 요구사항, 성능 요구사항, 설계 제약조건, 소프트웨어 시스템 특성, 기타 요구사항을 각 절에서 살펴본다. 기능적 요구사항에서는 데이터 흐름도와 프로세스 기술, 데이터 구조 명세, 데이터 사전을 포함한다.

2. 종합 기술 (Overall Description)

2.1 프로젝트 결과물 전망 (Product Perspective)

본 제품은 언어 모델인 ELECTRA를 기반으로 만들어진 영화평 감성 분석 소프트웨어이다. 영어를 기반으로 만들어진 감성 분석 소프트웨어와는 다르게 한국어를 학습시킨 모델을 사용하였고, 추가적으로 fine tuning을 하여 한국어 영화평에 대한 정확도를 높인 소프트웨어이다.

2.2 제품 기능 (Product Functions)

본 프로젝트의 결과물은 다음과 같은 기능을 수행한다.

- (기능 1) 사용자로부터 리뷰 텍스트 입력 : 사용자가 최근에 시청한 영화에 대한 리뷰를 입력하게 한다. 해당 리뷰 텍스트를 서버에 전송한다.
- (기능 2) 입력 받은 리뷰 텍스트 감성 분석(공정/부정) : 사용자로부터 입력받은 리뷰 텍스트를 학습된 모델을 통해 분석하여 공정/부정의 결과를 얻는다.
- (기능 3) 감성 분석 결과를 시각적으로 출력 : 사용자가 입력한 리뷰 텍스트의 감성을 분석한 결과를 그래프와 수치등의 시각적 도구를 이용해서 표현한다.
- (기능 4) 사용자에게 감성 분석 결과가 맞는지 틀린지 정보를 입력 요청 : 사용자로부터 입력 받은 리뷰 텍스트의 공정/부정 결과를 보여준 후, 결과가 맞는지 틀린지의 여부를 사용자에게 추가로 입력받는다.
- (기능 5) 입력 받은 리뷰 텍스트와 사용자가 입력한 해당 텍스트의 감성(공정/부정)을 데이터베이스에 저장 : 소프트웨어에서 나온 감성 분석의 결과가 맞는지 사용자가 직접 입력해주고, 입력받은 데이터를 데이터베이스에 저장한다.

2.3 사용자 특성 (User Characteristics)

시스템 관리자는 충분히 모델에 대한 정보를 습득하고 시스템 전반에 대한 이해를 갖춘 사람으로 한정한다.

클라이언트 사용자는 일반 사용자로 한글을 원활하게 읽을 수 있고 그 의미를 파악할 수 있는 능력을 갖추고 있으며, 영화에 대한 리뷰를 문법에 맞는 한글로 작성할 수 있다고 가정한다.

2.4 제약조건 (Constraints)

본 장의 가절에서 언급한 내용을 바탕으로 시스템을 설계 및 구현한다 그 이외의 사항들은 . 개발자가 선호하는 방향을 선택하여 설계 및 구현하되 단 다음의 항목을 준수한다.

- 리뷰 텍스트의 언어는 한국어로 한정한다.
- 기반이 되는 모델은 건국대학교 김학수 교수님 연구실에서 제작한 ELECTRA 모델을 사용한다.
- 프로그래밍 언어는 Python으로 한다.
- 군산대학교에서 제작한 KNU 한국어 감성 사전을 사용하여 모델의 정확성을 높인다.

2.5 가정과 의존성 (Assumptions and Dependencies)

본 문서의 모든 시스템은 건국대학교 김학수 교수님 연구실에서 제작한 ELECTRA 모델을 기반으로 구현할 것을 가정하였다. 또한 서버는 리눅스와 오픈 소스를 기반으로 설계 및 구현할 것을 가정하였다. 따라서 서버에 관한 내용은 리눅스를 기준으로 작성되었으며, 타 운영체제에서 동일하게 적용되지 않을 수 있다. 서버의 경우 타 운영체제는 전혀 고려하지 않으며, 일부 클라이언트의 경우 타 운영체제를 고려할 수 있다.

3. 상세 요구사항 (Specific Requirements)

3.1 사용자 인터페이스(User Interfaces)

이름	화면을 통한 입력 창 출력
목적 / 내용	사용자의 영화 리뷰 텍스트를 입력받을 입력창을 화면에 출력한다.
입력 주체 / 출력 목적지	서버 / 사용자
범위 / 정확도 / 허용오차	해당 없음
단위	화면
시간 / 속도	사용자의 입력에 따른 화면 전환
타 입출력과 관계	사용자의 입력 대기
화면 형식 및 구성	텍스트 입력 창
윈도우 형식 및 구성	해당 없음
데이터 형식	이미지
명령 형식	해당 없음
종료 메시지	입력 버튼에 의한 종료

이름	사용자 리뷰 텍스트 입력
목적 / 내용	사용자의 영화 리뷰 텍스트를 입력받아 서버에 전송한다
입력 주체 / 출력 목적지	사용자 / 서버
범위 / 정확도 / 허용오차	해당 없음
단위	사용자의 키보드나 마우스에 의한 입력
시간 / 속도	비정기적인 사용자의 입력 / 즉각적인 사용자 명령 수행
타 입출력과의 관계	입력 내용을 서버로 전달
화면 형식 및 구성	해당 없음
윈도우 형식 및 구성	해당 없음
데이터 형식	utf-8 인코딩의 한글 문자열
명령 형식	입력
종료 메시지	입력 버튼에 의한 종료

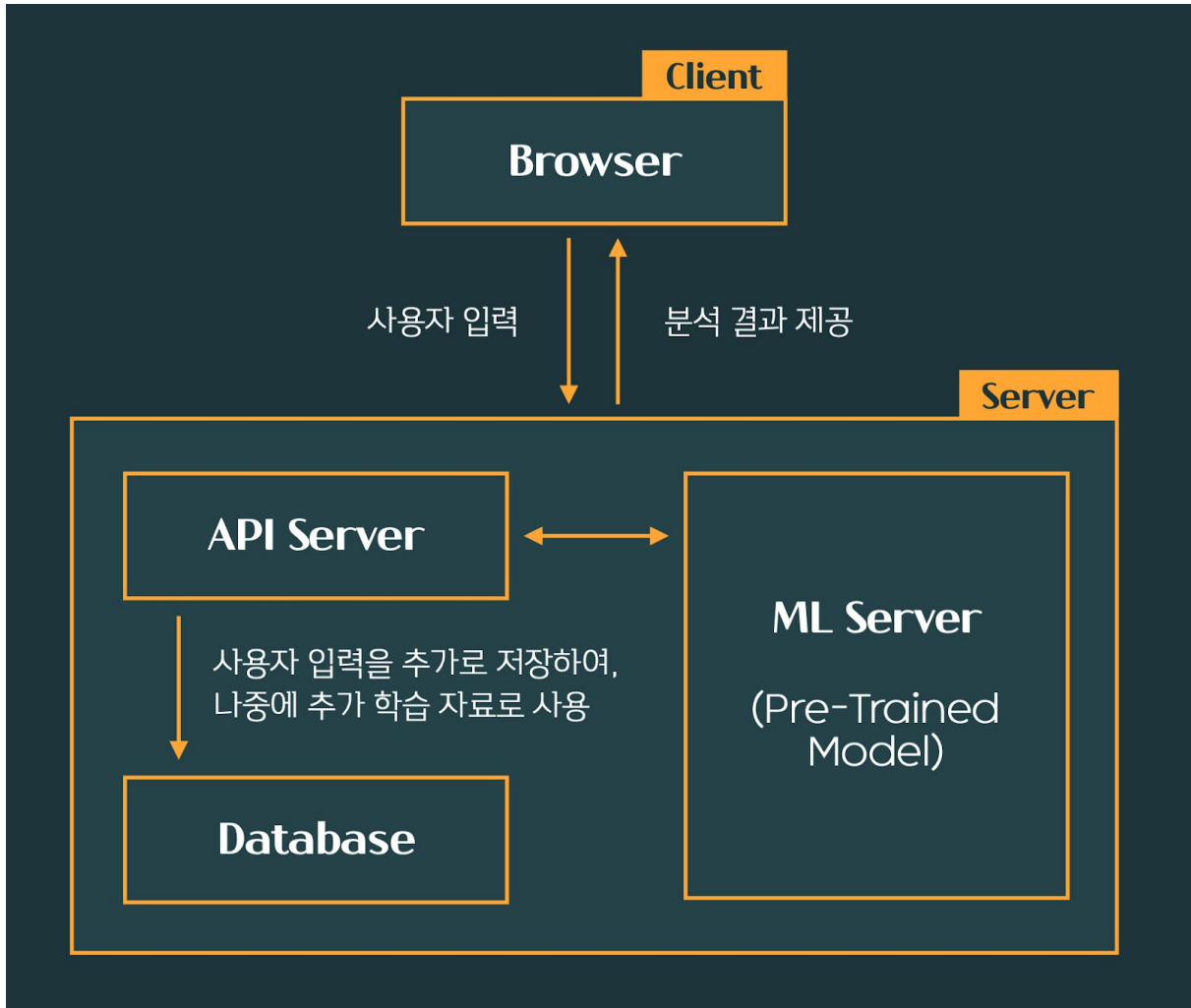
이름	사용자 리뷰 텍스트 감성 분석 결과 출력
목적 / 내용	사용자의 영화 리뷰 텍스트에 대한 감성 분석 결과를 출력한다.
입력 주체 / 출력 목적지	서버 / 클라이언트
범위 / 정확도 / 허용오차	해당 없음
단위	화면
시간 / 속도	사용자의 입력에 따른 화면 전환
타 입출력과의 관계	사용자의 입력에 대한 처리 결과를 출력
화면 형식 및 구성	사용자가 입력한 리뷰 텍스트와 감성 분석 결과와 근거를 제공하는 창
윈도우 형식 및 구성	해당 없음
데이터 형식	이미지
명령 형식	해당 없음
종료 메시지	입력 버튼에 의한 종료

이름	사용자 리뷰 텍스트 감성 분석 결과 정오 판단 입력
목적 / 내용	사용자의 영화 리뷰 텍스트에 대한 감성 분석 결과에 대한 사용자의 정오 판단 데이터를 입력받는다
입력 주체 / 출력 목적지	사용자 / 서버
범위 / 정확도 / 허용오차	해당 없음
단위	버튼 클릭
시간 / 속도	사용자의 1회 입력 / 즉각적인 서버 전송
타 입출력과의 관계	사용자 리뷰 텍스트 감성 분석 결과 출력 후 사용자의 입력 대기
화면 형식 및 구성	입력 버튼

윈도우 형식 및 구성	입력 버튼
데이터 형식	해당 없음
명령 형식	해당 없음
종료 메시지	입력 버튼에 의한 종료

3.2 기능적 요구사항 (Functional Requirements)

3.2.1 정보 흐름 (Information Flows)



3.2.2 프로세스 기술 (Process Descriptions)

(기능 1) 사용자로부터 리뷰 텍스트 입력

브라우저에서 사용자의 리뷰 텍스트를 입력 받을 수 있는 화면을 띄운다

사용자가 리뷰 텍스트를 입력하고 '입력' 버튼을 누르면 리뷰 텍스트에 대한 유효성 검사를 진행한다.

유효성 검사에 통과하지 못한 경우, 재입력을 요청한다.

(기능 2) 입력받은 텍스트 감성 분석(긍정/부정)

API 서버를 통해 입력된 리뷰 텍스트를 ML 서버에 전송한다.

전송받은 리뷰 텍스트를 ML 서버에서 훈련이 완료된 모델을 사용하여 감성 분석 결과를 얻는다.

(기능 3) 감성 분석 결과를 시각적으로 출력

감성 분석 데이터를 API 서버를 통해 클라이언트에 전달한다.

브라우저에 감성 분석의 근거가 되는 word piece 와, word piece 별 긍정/부정 판단 결과와 가중치를 시각적으로 제공한다.

(기능 4) 사용자에게 감성 분석 결과에 대한 정오 정보를 입력 요청

브라우저에서 감성 분석 결과에 대한 입력을 요청한다.

버튼을 통해 입력을 받고, 입력을 받은 후에는 버튼이 비활성화되며, 감성 분석 결과 화면은 유지한다.

(기능 5) 입력받은 리뷰 텍스트와 사용자가 입력한 해당 텍스트의 감성(긍정/부정)을

데이터베이스에 저장

기능 4에서 입력받은 정오 정보를 리뷰 텍스트와 함께 데이터 베이스에 저장한다.

4. 추가 정보 (Supporting Information)

4.1 소프트웨어 요구사항 명세서

본 소프트웨어 요구사항 명세서는 IEEE의 권고안(IEEE Recommended Practice for Software Requirements Specifications) 을 따라 작성되었음.

4.2 문서 이력

날짜	버전	설명	작성자
2020년 7월 30일	1.0	초기 아이디어를 기반으로 작성	김주언, 최보라, 류한길, 유병찬