

영화 리뷰 감성 분석을 통한 키워드 추출 및 시각화

박종찬⁰, 김성진*, 윤영현**, 백재순(교신저자)*

⁰명지전문대학 ICT융합공학과,

*명지전문대학 ICT융합공학과,

**명지전문대학 정보통신공학과

e-mail: {2022821007⁰, ict214548*, yhyoon**, hisoon99*}@mj.ac.kr

Keyword Extraction and Visualization of Movie Reviews through Sentiment Analysis

Jong-Chan Park⁰, Sung Jin Kim*, Young Hyun Yoon**, Jai Soon Baek*

⁰Dept. of ICT Convergence Engineering, Myongii College,

*Dept. of ICT Convergence Engineering, Myongii College,

**Dept. of Information Technology communication, Myongii College

● 요약 ●

본 연구에서는 감성 분석 기반의 키워드 도출형 영화 리뷰 웹사이트를 개발하였다. 사용자들은 영화에 대한 리뷰를 작성할 때, 자동으로 키워드를 추출하는 기능을 활용하여 다양하면서도 빠르게 정보를 얻을 수 있다. 사용자가 작성한 리뷰를 시스템에 입력하면, 내부적으로 ChatGPT를 활용하여 텍스트를 분석하고 키워드를 추출한다. 이를 통해 사용자는 별다른 노력 없이도 키워드를 통해 영화의 장르, 감독, 배우, 플롯 요소 등 다양한 정보를 빠르게 확인할 수 있다. 추출된 키워드는 저장되어 시각화에 활용되며, 사용자들은 리뷰에 대한 원하는 정보를 쉽게 얻을 수 있다. 개발된 키워드 도출형 영화 리뷰 웹사이트는 사용자들에게 빠르고 다양한 정보를 제공하며, 영화 관련 결정을 내리는 데에 도움을 줄 것으로 기대된다.

키워드: ChatGPT API, Django, 키워드 리뷰(keyword review), 감성 분석(sentiment analysis)

I. Introduction

기존의 영화 리뷰는 점수나 서술형 형식으로만 작성되어 점수만으로는 자세한 평가를 확인하기 어렵고, 서술형 리뷰를 작성하는 경우에도 전체 내용을 파악하기 어려웠다. 이에 리뷰 키워드 추출 및 시각화 시스템을 도입하여 사용자에게 빠른 정보 제공을 목표로 하는 서비스를 개발하게 되었다. 키워드를 추출하여 사용자에게 제공함으로써 영화의 주요 내용을 간략하게 파악할 수 있고, 장르, 감독, 배우 등의 정보를 신속하게 확인할 수 있다. 이에 본 연구에서는 기존의 리뷰 형식의 한계를 극복하기 위해 키워드 추출 및 시각화 시스템을 개발하고자 한다.

II. Preliminaries

1. ChatGPT API

ChatGPT API는 OpenAI에서 개발한 강력한 자연어 처리(NLP) 모델이다. 이 API는 GPT(Generative Pre-trained Transformer)

아키텍처와 고급 언어 이해 기술을 활용하여 대화 기능을 제공한다. API와 상호작용함으로써 사용자는 인간과 유사한 응답을 생성하고 언어 번역을 수행하고, 감성 분석을 수행하는 등 텍스트 기반의 다양한 자연어 처리 작업에 유용하게 활용될 수 있다. 이 연구에서는 ChatGPT API를 활용하여 리뷰 작성과 관련된 키워드 도출을 수행한다.

2. Django

Django는 Python으로 작성된 고성능 웹 애플리케이션 프레임워크이다. 이 프레임워크는 웹 개발을 위한 다양한 기능과 도구를 제공하여 개발자들에게 효율적인 웹 애플리케이션 구축을 지원한다. Django는 MVC(Model-View-Controller) 패턴을 기반으로 하여 데이터 모델링, URL 라우팅, 사용자 인증, 세션 관리 등을 간편하게 처리할 수 있다. 이 연구에서는 Django를 활용하여 웹사이트의 기능 개발과 사용자 리뷰의 관리를 수행한다.

III. The Proposed Scheme

본 연구에서는 영화 정보 제공 및 리뷰 작성이 가능한 영화 리뷰 웹사이트를 개발하였다. 사용자는 영화에 대한 리뷰를 작성할 때, 웹사이트에서 제공하는 프롬프트를 통해 ChatGPT API를 호출한다. 이때, ChatGPT API를 사용하여 질문을 던지고 받은 답변을 정규식을 통해 알맞게 추출하여 DB에 저장하는 방식을 구현하였다.

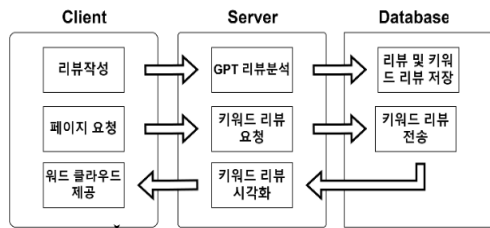


Fig. 1. 시스템 구조도

1. Page organization

본 연구에서 구현한 영화 리뷰 웹사이트는 메인 화면, 검색 결과 화면, 영화 상세 정보 화면으로 구성되어 있다. 사용자는 상단의 네비게이션 바를 통해 메인 화면으로 이동하고, 검색 기능을 통해 원하는 영화를 검색할 수 있다. 검색 결과 화면에서는 사용자의 검색에 따른 결과를 제공한다. 또한, 메인 화면과 검색 결과 화면에서는 일일 박스오피스 정보, 최신 리뷰, 최신 개봉 예정 영화 정보를 확인할 수 있다.

2. Provide movie information

영화에 대한 상세한 정보는 KMDB API를 통해 가져온다. 메인 화면 및 검색 결과 화면에서는 영화의 요약 정보와 함께 포스터 및 제목을 제공하며, 영화 상세 정보 화면에서는 해당 영화에 대한 상세 정보를 제공한다. 또한, 사용자는 영화 상세 정보 화면에서 댓글을 작성할 수 있다.

3. Writing and analyzing reviews

사용자가 영화 리뷰를 작성할 때 사용자가 작성한 리뷰와 서버측에서 준비된 프롬프트를 결합한 결과를 통해 ChatGPT API를 호출한다. ChatGPT API는 사용자의 답변을 이해하고 분석하기 위해 자연어 이해(NLU) 기술을 활용한다. 이를 통해 리뷰의 감정 분석과 키워드 추출이 이루어진다. 추출된 키워드는 정규식을 통해 알맞은 형식으로 추출되어 데이터베이스(DB)에 저장된다. 이러한 키워드는 영화의 장르, 감독, 배우, 플롯 요소 등 다양한 측면을 포함하며, 이후의 시각화를 위해 활용된다.

4. Visualization

저장된 긍정적 및 부정적 키워드 리뷰는 사용자에게 AnyChart 라이브러리를 활용하여 워드 클라우드 형태로 제공된다. 구현된 워드

클라우드는 fig2과 같다.

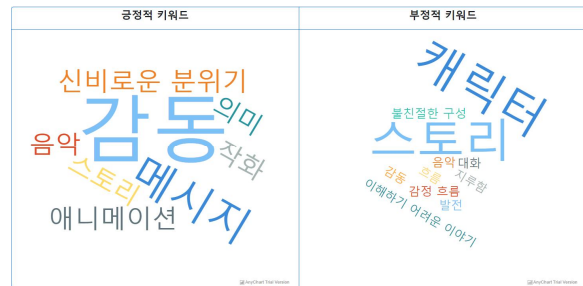


Fig. 2. 키워드 워드클라우드

이러한 영화 리뷰 웹사이트의 개발은 사용자들에게 영화 정보 제공과 함께 효율적인 리뷰 분석을 가능하게 하여 더욱 유용한 서비스를 제공할 수 있게 한다.

IV. Conclusions

본 연구에서는 키워드 도출형 영화 리뷰 웹사이트를 개발하여 사용자들이 영화 리뷰를 읽을 때 유용한 정보를 제공하였다. 향후 연구에서는 키워드 추출 알고리즘을 더욱 정확하게 개선하고, 사용자 경험을 개선하기 위한 기능을 추가할 것이다. 이를 통해 사용자들에게 더욱 유용한 영화 리뷰 서비스를 제공할 수 있을 것으로 기대된다.

REFERENCES

[1] <https://platform.openai.com/docs/api-reference>
 [2] <https://docs.djangoproject.com/ko/4.2/>