

양방향 순환 신경망 언어 모델을 이용한 Fuzzy-AHP 기반 영화 추천 시스템

오재택¹, 이상용^{2*}

¹공주대학교 컴퓨터공학과 박사, ²공주대학교 컴퓨터공학부 교수

A Fuzzy-AHP-based Movie Recommendation System with the Bidirectional Recurrent Neural Network Language Model

Jae-Taek Oh¹, Sang-Yong Lee^{2*}

¹Doctor, Department of Computer Science & Engineering, Kongju National University

²Professor, Division of Computer Science & Engineering, Kongju National University

요약 다양한 정보가 대량으로 유통되는 IT 환경에서 사용자의 요구를 빠르게 파악하여 의사결정을 도와줄 수 있는 추천 시스템이 각광을 받고 있다. 그러나 현재 추천 시스템은 사용자의 취향이나 관심사가 바뀌었을 때 선호도가 즉시 시스템에 반영이 되지 않을 수가 있으며, 광고 유도로 인하여 사용자의 선호도와 무관한 아이템이 추천될 수가 있다는 문제점이 있다. 본 연구에서는 이러한 문제점을 해결하기 위해 양방향 순환 신경망 언어 모델을 이용한 Fuzzy-AHP 기반 영화 추천 시스템을 제안하였다. 본 시스템은 사용자의 취향이나 관심사를 명확하고 객관적으로 반영하기 위해 Fuzzy-AHP를 적용하였다. 그리고 사용자가 선호하는 영화를 예측하기 위해 양방향 순환 신경망 언어 모델을 이용하여 실시간으로 수집되는 영화 관련 데이터를 분석하였다. 본 시스템의 성능을 평가하기 위해 그리드 서치를 이용하여 전체 단어 집합의 크기에 대한 학습 모델의 적합성을 확인하였다. 그 결과 본 시스템의 학습 모델은 전체 단어 집합의 크기에 따른 평균 교차 검증 지수가 97.9%로 적합하다는 것을 확인할 수 있었다. 그리고 본 모델은 네이버의 영화 평점 대비 평균 제공근 오차가 0.66, LSTM 언어 모델은 평균 제공근 오차가 0.805으로, 본 시스템의 영화 평점 예측성이 더 우수함을 알 수 있었다.

주제어 : 양방향 순환 신경망, 언어 모델, 퍼지 계층적 분석 방법, 스크래핑, 추천 시스템

Abstract In today's IT environment where various pieces of information are distributed in large volumes, recommendation systems are in the spotlight capable of figuring out users' needs fast and helping them with their decisions. The current recommendation systems, however, have a couple of problems including that user preference may not be reflected on the systems right away according to their changing tastes or interests and that items with no relations to users' preference may be recommended, being induced by advertising. In an effort to solve these problems, this study set out to propose a Fuzzy-AHP-based movie recommendation system by applying the BRNN(Bidirectional Recurrent Neural Network) language model. Applied to this system was Fuzzy-AHP to reflect users' tastes or interests in clear and objective ways. In addition, the BRNN language model was adopted to analyze movie-related data collected in real time and predict movies preferred by users. The system was assessed for its performance with grid searches to examine the fitness of the learning model for the entire size of word sets. The results show that the learning model of the system recorded a mean cross-validation index of 97.9% according to the entire size of word sets, thus proving its fitness. The model recorded a RMSE of 0.66 and 0.805 against the movie ratings on Naver and LSTM model language model, respectively, demonstrating the system's superior performance in predicting movie ratings.

Key Words : Bidirectional Recurrent Neural Network, Language Model, Fuzzy-AHP, Scraping, Recommendation System

*This work was supported by the research grant of the Kongju National University in 2020.

*Corresponding Author : Sang-Yong Lee(sylee@kongju.ac.kr)

Received September 4, 2020

Revised September 22, 2020

Accepted December 20, 2020

Published December 28, 2020

1. 서론

5G 이동통신의 시작으로 정보통신 인프라 구축이 더욱 견고해지고 있으며, AI 기반 플랫폼을 활용한 시스템이 빠르게 등장하고 있다[1-3]. 추천 시스템은 그중의 하나의 시스템으로써 사용자의 니즈를 빠르게 파악하여 사람 대신 의사결정을 할 수 있도록 적용되고 있다. 과거에는 사용자가 가장 적합하고 필요한 제품이나 콘텐츠 등을 직접 고민하면서 구매하였지만, 최근에는 추천 시스템이 추천해주는 상품을 직접 구매하고 있다. 이러한 경향은 여러 분야에 적용되고 있다. 대표적으로 축구 경기 영상 인공지능 분석 서비스로 해당 축구 프로팀 관계자들이 다음 경기를 준비할 때 많은 시간을 소비한다는 점에 착안하여 인공지능 기반 기술을 적용해 팀과 선수들에게 추천 전략뿐만 아니라 분석 결과 등을 제공하고 있다[2].

이렇게 추천 시스템은 해당 플랫폼의 참여율을 높이기 위해 사용자에게 가장 적합하고, 알맞은 콘텐츠를 추천하고 있다[3, 4]. 그러나 이러한 추천 시스템은 사용자의 취향이나 관심사가 바뀌었을 때 이에 따른 선호도가 반영이 되지 않을 수가 있으며, 사용자의 선호도와 무관하게 콘텐츠(Item)가 추천될 수가 있다는 문제가 제기되고 있다. 또한 선호하지 않는 콘텐츠를 지속적으로 유도함으로써 인위적인 간섭 시스템을 제공한다는 우려가 있으며, 전반적인 추천 알고리즘의 적용 방식에 대한 불확실성을 증가시킬 수 있다는 문제가 제기되고 있다[3, 5].

이에 본 연구에서는 이러한 문제 제기를 해소하기 위해 양방향 순환 신경망 언어 모델을 이용한 Fuzzy-AHP 기반 영화 추천 시스템을 제안하고자 한다.

Fuzzy-AHP를 통해 사용자의 취향이나 관심사를 명확하고 객관적으로 반영할 수 있으며, 양방향 순환 신경망 언어 모델을 이용한 학습 방법으로 사용자가 선호하는 영화를 예측하여 추천할 수 있을 것이다.

2. 관련연구

2.1 추천 시스템

최근 각국 정부에서 AI 시장에 대한 투자를 적극적으로 단행하면서 정보통신 기술 사업자들이 맞춤형 추천 서비스 및 개인화 시스템을 의료·금융·가전·모빌리티·로봇·에너지 등 산업 전반 분야에 적용시키고 있다[3, 6]. 이를 기반으로 추천 서비스를 적용한 가전 제품이

나 사용자의 여러 위험 리스크를 최소화할 수 있는 소프트웨어를 출시하고 있다.

즉, 추천 시스템은 사용자의 취향을 이해하여 맞춤 상품이나 콘텐츠를 제공하며, 사용자가 관심을 보일만 한 새로운 관심사를 소개하는 시스템을 말한다[7]. 현재 추천 시스템은 기계학습을 적용하고, 이를 심화시키는 방향으로 나아가고 있는데 그 이유는 빠른 연산이 가능한 하드웨어의 개발이 급격하게 발전하고 있기 때문이다[5].

S. Sivaprasad와 3인은 아카데미 수상작에서 획득한 대본 및 시놉시스, 리뷰 등을 이용하여 LSTM(Long-short Term Memory Network) 즉, 순환 신경망의 문제점인 시퀀스 데이터가 변형되거나 사라지는 문제를 해결할 수 있는 장·단기 기억망으로 영화 속 감정을 예측하거나 장르를 추천하는 시스템을 제안하였다[8].

G. Liu와 1인은 사용자가 선호하는 콘텐츠(Item)를 파악할 수 있는 추천 기법인 협업 필터링(Collaborative Filtering)과 Doc2Vec(Document to Vector) 즉, 문장 자체를 실수 형태로 임베딩하는 기법을 활용하여 사용자 선호도 기반 영화 시놉시스 추천 시스템을 제안하였다[9].

앞으로 추천 시스템 및 개인화 시스템은 일상생활에 빠르게 스며들 것이며, 이를 통해 추천 시스템의 가치는 더욱 높아질 것으로 예상된다.

본 연구에서는 영화 추천 시스템의 구현을 통해 사용자가 선호하거나 관심이 있는 영화를 추천하고자 한다.

2.2 Fuzzy-AHP

Fuzzy-AHP는 사용자가 선호하거나 판단, 평가하는 불확실한 콘텐츠(Item)의 정보의 대안을 평가하고, 판단하여 보다 정확한 의사결정을 진행하는 기법을 말한다. Fuzzy-AHP는 사용자의 선호도에 대해 특정 값으로 표현하지 않고, 범위 값으로 표현하기 때문에 보다 객관적으로 사용자의 의사를 반영할 수 있다[10, 11].

Fuzzy-AHP는 e-Business를 위해 업무 프로세스 계획이나 전략을 수립할 때 사용되며, 집단 의사결정이나 문제 해결을 위한 전략적 합의를 위해 사용된다[12].

본 연구에서는 Fuzzy-AHP를 이용하여 사용자가 선호하는 요소의 영화를 반영하고자 하며, 이를 통해 보다 정확한 사용자의 의사를 확인할 수 있다.

2.3 양방향 순환 신경망

양방향 순환 신경망(Bidirectional Recurrent

Neural Networks)은 Fig. 1과 같이 임베딩된 형태의 레이블을 두 개의 순환층이 양방향으로 학습하는 신경망을 말한다[13, 14].

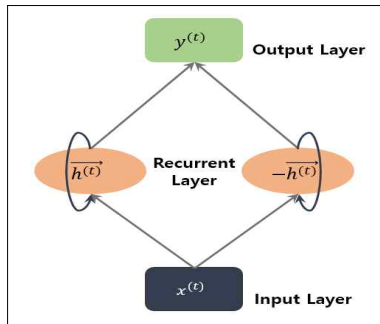


Fig. 1. Bidirectional Recurrent Neural Networks Model

두 개의 순환층은 서로 연결되어 있지 않으며, 전향 순환층(\vec{h})과 후향 순환층($-\overleftarrow{h}$)을 나타낸다. 양방향 순환 신경망은 순환 신경망의 학습 방법을 따르지만 이전 타임 스텝의 셀의 출력을 입력값으로 전달할 때 전향 순환층은 시간 순서로 임베딩된 시퀀스 데이터를 전달받으며, 후향 순환층은 반대의 시간 순서로 임베딩된 시퀀스 데이터를 전달받는다.

양방향 순환 신경망은 기존 순환 신경망이나 LSTM이 효과적으로 처리할 수 없었던 시퀀스 데이터를 처리할 수 있으며, 다른 형태로 변하게 되는 시퀀스 데이터를 발견할 수 있어 학습 효과를 더욱 높일 수 있다[15].

T. N. Duc와 3인은 스트리밍 비디오 포맷과 비디오 버퍼링, 버퍼링 메모리 관련 데이터를 양방향 순환 신경망을 사용하여 품질이 좋은 스트리밍 비디오를 추천하는 시스템을 제안하였다[16].

X. Han와 4인은 사용자가 선택하였던 의상 데이터로 양방향 순환 신경망 학습을 통해 기존 구성 스타일에 추가로 추천할 만한 패션 아이템을 추천하는 시스템을 제안하였다[17].

본 연구에서는 양방향 순환 신경망을 적용하여 스크래핑(Scraping)으로 수집한 영화 관련 데이터를 학습하고자 한다.

2.4 언어 모델

언어 모델(Language Model)은 이전 시퀀스 데이터가 주어졌을 때 다음 시퀀스의 데이터를 예측하는 모델을 말한다[15, 18].

여러 순환 신경망 언어 모델을 이용하여 다음 시퀀스의 데이터를 예측할 수 있으며, 효과적인 신경망을 선택하여 시스템을 구성하면 된다. 모델의 생성은 언어 모델로 예측된 시퀀스 데이터를 샘플링 과정으로 생성할 수 있다. 보통 무작위 시퀀스 데이터를 생성하여 구성하는 확률적인 샘플링 과정을 선택한다[15, 18].

본 연구에서는 양방향 순환 신경망으로 학습된 영화 데이터를 언어 모델로 구성하여 사용자가 선호하는 영화를 추천할 수 있는 추천 시스템을 구축하고자 한다.

3. 시스템 설계

본 연구에서 제안하는 추천 시스템은 Fig. 2와 같이 사용자 정보 수집 모듈과 학습 모듈, 매핑 모듈, 추천 모듈 그리고 데이터베이스로 구성된다.

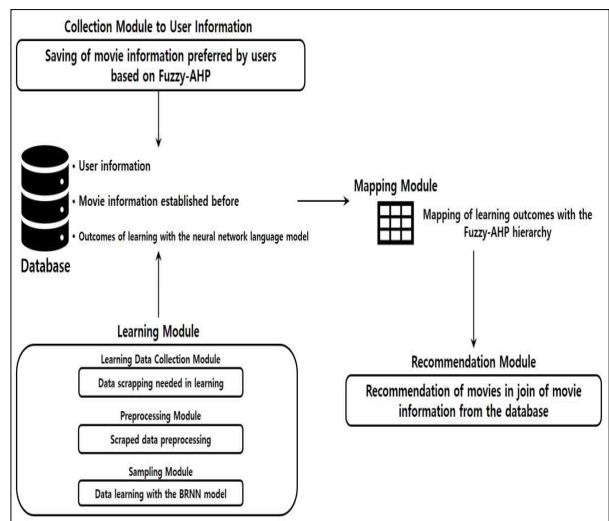


Fig. 2. System Structure

사용자 정보 수집 모듈은 사용자가 선호하는 영화 정보를 판단하기 위해 Fuzzy-AHP를 이용하여 선호 기준(Criteria)에 가장 적합한 사용자의 정보를 파악하는 모듈이다. 선호 기준은 Fig. 3과 같이 Reddy 외 2인이 분석한 영화가 성공하기 위한 요인들을 계층 구조로 구성하여 모형화하였으며, 각각의 요소들은 선호 기준에 대한 하위 요소로 나타내어 계층별 비교를 통해 사용자가 선호하는 영화 정보를 파악할 수 있다. 영화의 장르는 동일 계층에 있는 요소들간의 비교를 위해 비교 대상을 최대 7±2 가지로 제안하여 장르의 유사성 및 명칭의 다양화를 통합하여 나타내었다[19-21].

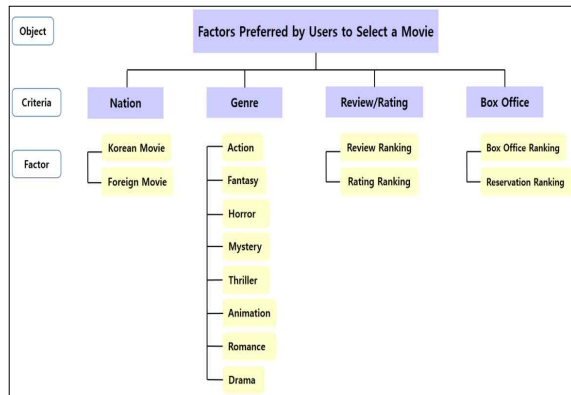


Fig. 3. Hierarchical Structure of Collection Module to User Information

학습 모듈은 영화를 추천하기 위해 학습에 필요한 데이터를 스크래핑하여 그 결과를 전처리하고, 양방향 순환 신경망 언어 모델로 학습하는 모듈이다. 학습 데이터 수집 모듈에서 학습에 필요한 영화와 관련된 뉴스 기사 및 포털 사이트의 영화 리뷰와 평점 댓글 등이 스크래핑되며, 전처리 모듈에서 스크래핑된 데이터를 하나의 파일로 병합하여 학습에 불필요한 특수 문자나 기호 등을 제거한다. 학습에 불필요한 문자열 등은 이후 샘플링 모듈에서 학습 저하를 일으켜 해당 모델이 과적합이나 과소적합이 될 수 있다. 해당 학습 데이터의 전처리가 완료된 이후에는 샘플링 모듈을 통해 양방향 순환 신경망 언어 모델로 학습 과정이 진행된다. 언어 모델 적용 방법(샘플링 기법)은 신경망 모델의 구성과 학습 레이블 형태에 따라 선택할 수 있다. 탐욕적 샘플링(Greedy Sampling) 기법은 예측률이 좋은 시퀀스 데이터를 추천하는 방법이며, 확률적 샘플링(Stochastic Sampling) 기법은 무작위 추천 기법을 말한다[18]. 본 연구에서는 확률적 샘플링 기법을 통해 학습된 데이터를 Fig. 2의 선호 요소에 따라 구성하였다.

매핑 모듈은 학습 모듈에서 학습된 결과와 사용자 정보 수집 모듈에서 처리된 사용자가 선호하는 영화 정보를 서로 매핑하는 모듈이다. 해당 테이블들의 속성값(선호 요인)이 서로 일치하면, 학습 결과를 변수에 트랜잭션하여 이후 추천 모듈을 통해 사용자에게 추천될 수 있도록 한다.

추천 모듈은 매핑 모듈에서 트랜잭션한 결과를 데이터베이스에 저장되어 있는 영화 정보와의 일치 여부를 판단하여 사용자에게 영화를 추천하는 모듈이다. 추천된 인터페이스에는 본 시스템이 예측한 평점과 포스터가 제공되며, 해당 포스터의 터치를 통해 사용자는 더욱 자세한

영화 정보를 확인할 수 있다.

본 연구의 추천 시스템의 데이터베이스에는 추천 시스템을 사용하는 사용자의 기본 정보와 사용자 정보 수집 모듈에서 처리된 사용자가 선호하는 영화 정보, 학습 모듈에서 학습된 학습 결과와 사용자에게 영화를 추천할 수 있도록 포털 사이트의 정보를 바탕으로 구성된 영화 정보가 저장되어 있다. 영화 정보는 Naver 영화의 정보를 바탕으로 구성하였으며, Fig. 2에서 각 선호 기준의 하위 요소에 따른 영화를 4개씩 저장하였다.

4. 시스템 구현

본 연구에서 제안하는 추천 시스템을 구현하기 위해 설계된 모듈들을 바탕으로 시스템을 구현하였으며, 인터넷 정보 서비스(Internet Information Service)를 이용하여 웹 서버와 이를 반환할 수 있는 페이지를 구축하였다. 본 연구에서 제안하는 추천 시스템을 구현하기 위한 시스템 개발 환경은 Table. 1과 같다.

Table 1. Development Environment

Mobile Device	Samsung Galaxy S8+
Operating System	Android 10.0, Windows 10
Language	Java for Android, Python, ASP.NET 4.8, HTML5, CSS3.0
Database	MS SQL 2012
Tool	Android Studio, Pycharm, Visual Studio 2015

Table 1과 같이 클라이언트를 구축하기 위한 도구로는 Galaxy S8+와 Android Studio, Visual Studio 2015를 사용하여 구축하였으며, 양방향 순환 신경망 언어 모델 기반으로 해당 시퀀스 데이터를 학습하기 위한 모델을 구축하기 위해 Pycharm으로 이를 구현하였다. 그리고 학습 결과와 사용자가 선호하는 영화를 저장하고, 이를 서로 매핑하여 선호하는 영화를 추천하기 위한 데이터베이스는 MS SQL을 이용하여 해당 저장소를 구축하였다.

사용자 정보 수집 모듈은 ASP.NET에서 제공하는 응용 프로그램인 웹 폼(Web Form)을 이용하였다. 웹 폼을 구성하여 Fig. 3의 선호 기준 및 선호 요인들을 일대일 비교하고, 이를 Fuzzy-AHP로 처리한 가중치들을 정렬하여 데이터베이스에 저장한다. 우선순위가 가장 높은 가중치는 사용자가 선호하는 영화 정보를 나타내며, 0이 아

닌 가장 낮은 가중치는 사용자가 선호하지 않는 영화 정보를 나타낸다. 본 연구에서 사용자가 선호하는 영화의 요인에 대한 종합 가중치를 나타낸 결과는 Table 2와 같다.

Table 2. Fuzzy-AHP weights for movie preference factors

Preference Ranking(Ratio)	Preference Factor	Total weights
1(0.9)	Movies whose ranks are high on the reservation list	0.889
2(0.8)	Thriller Movie	0.064
3(0.7)	Mystery Movie	0.047
Joint bottom ranks(0)	The remaining factors	0

Table. 2의 결과를 통해 종합 가중치가 0이 아닌 선호 요인들은 데이터베이스에 저장되며, 매핑 모듈에서 학습된 결과와 일치 비교를 통해 추천 모듈로 트랜잭션된다.

학습 모듈은 Python을 이용하여 학습에 필요한 데이터를 스크래핑하기 위한 시스템을 구축하고, 영화와 관련된 뉴스 기사 및 포털 사이트의 영화 리뷰와 평점 댓글 등을 스크래핑한다. 스크래핑된 데이터는 Python으로 구현된 프로그램을 통해서 특수 문자나 기호 등이 제거되고, 이후 프로그래밍을 통해 해당 시퀀스 데이터는 토큰화(단위화) 과정과 고정된 길이로 임베딩되어 양방향 순환 신경망 언어 모델로 학습된다. 학습된 결과는 데이터베이스에 저장되며, 매핑 모듈에서 사용자가 선호하는 영화 정보와 매핑될 수 있도록 Fig. 3의 선호 요소에 맞게 저장된다.

매핑 모듈은 신경망 학습이 완료된 즉시 학습된 결과를 트랜잭션할 수 있도록 구현하였으며, 사용자 정보 수집 모듈의 테이블에서 우선 선호 요인이 가장 높은 학습 결과만을 트랜잭션하여 이를 변수에 저장한다.

이후 추천 모듈에서 데이터베이스에 저장되어 있는 영화 정보와 매핑 모듈의 트랜잭션한 결과가 서로 일치하면, 이를 사용자 인터페이스로 추천하게 된다. 일치 여부의 판단은 매핑 모듈에서 트랜잭션한 결과와 데이터베이스에 저장되어 있는 영화 제목의 문자열이 서로 일치하는지를 확인한다. 서로 일치하면 이를 변수에 담아 사용자 인터페이스에 추천하게 되며, 일치하지 않으면 후 순위의 영화가 추천된다. 추천된 영화는 시스템이 예측한 평점과 함께 해당 영화의 포스터가 제공되는데, 해당 영화의 포스터를 터치하면 해당 영화의 정보 페이지로 로

딩되어 영화의 자세한 정보를 알 수 있다.

본 시스템에서 추천된 영화의 평점 예측 방법은 학습 데이터에서의 추천된 영화의 단어 빈도수 비율(F)과 학습 결과의 정확도(A), Table. 2의 선호 순위 비율(R)을 이용하여 영화의 평점을 예측한다. 해당 식은 다음과 같으며, 사용자 인터페이스로 영화가 추천된 결과는 Fig. 4와 같다.

$$10 - \left(\frac{A \cdot F}{R} \right) \quad (\text{단, } F \geq 1) \quad \text{식(1)}$$

$$10 - \left(\frac{R}{A \cdot F} \right) \quad (\text{단, } F < 1) \quad \text{식(2)}$$

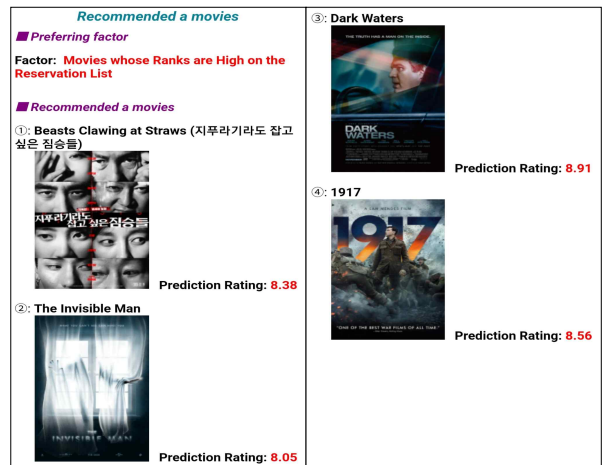


Fig. 4. Program Capture] User Interface to Recommend a Movie to Users

Fig. 4의 결과를 통해 사용자에게 영화를 선호하는 요인과 그 요인에 맞는 영화를 추천하게 된다. 본 연구에서 추천된 영화 1순위는 선호 요인과 신경망 학습 결과에 따른 문자열이 데이터베이스와 서로 일치한 ‘지푸라기라도 잡고 싶은 짐승들’ 영화가 추천되었으며, 나머지 영화는 선호 요인에 따른 포털 사이트의 우선순위에 맞는 영화가 추천되었다. Fig. 4의 결과는 3월 16일 기준으로 스크래핑된 데이터를 사용하였으며, 포털 사이트는 Naver 영화를 기준으로 하여 영화를 추천하였다.

5. 실험 및 평가

본 연구에서 제안한 영화 추천 시스템의 성능을 확인하기 위해 그리드 서치(Grid Search)로 신경망 학습에 사용된 스크래핑 데이터를 이용하여 전체 단어 집합의

크기에 대한 학습 모델의 적합성을 확인하였으며, 본 시스템이 예측한 평점과 다른 언어 모델인 LSTM을 적용하여 예측한 평점으로 실제 영화와의 평점 간 차이를 비교하였다.

그리드 서치는 신경망 학습 과정에서 필요한 매개변수의 조합을 판단하여 학습 모델에 적용하였을 때 학습 결과가 최적화될 수 있도록 검증하는 방법을 말한다[15, 18, 22]. 본 연구에서 전체 단어 집합의 크기를 10으로 학습했을 때 학습 모델의 교차 검증 지수를 나타낸 결과는 Fig. 5와 같다. 교차 검증 지수는 학습 데이터가 모델에 얼마나 적합인지 판단하는 지수로 80~90% 이상이면 해당 학습 모델이 적합하다고 판단한다[22].

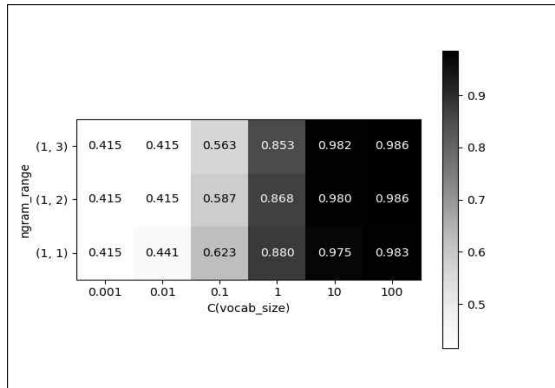


Fig. 5. Cross validation index of the learning model based on grid searches

Fig. 5에서 x 축의 매개변수 $C(\text{vocab_size})$ 는 전체 단어 집합의 크기를 나타내는 것으로 해당 학습 모델의 과적합이나 과소적합을 방지하기 위해 이를 조절하는 역할을 하는 제한 변수이다. y 축의 매개변수 ngram_range 는 시퀀스 데이터의 토큰(단위) 분할 방식을 나타내는 것으로 토큰의 길이가 1인 시퀀스 데이터를 유니그램(Unigram), 2인 시퀀스 데이터를 바이그램(Bigram) 그리고 3인 시퀀스 데이터를 트라이그램(Trigram)으로 나타낸다[15, 18, 22]. 즉, 그리드 서치는 해당 학습 모델에서 제한 변수와 분할 변수를 입력받아 학습 모델의 적합성을 측정할 수 있다. Fig. 5의 결과와 같이 본 연구의 학습 모델은 전체 단어 집합의 크기가 10일 때 평균 교차 검증 지수가 97.9%이므로 적합하다고 볼 수 있다.

Table. 3과 Table. 4는 사용자가 영화를 선호하는 요인 중 우선순위가 가장 높은 영화인 예매 순위가 높은 영화에 대해 실제 영화와의 평점 간 차이를 비교하였다. 해

당 영화의 평점은 2020년 3월 16일 기준, 네이버 영화를 참고하였다.

Table. 3과 Table. 4의 결과를 통해 LSTM 언어 모델은 평균 제공근 오차가 0.805이며, 본 연구에서 제안한 모델의 평균 제공근 오차는 0.66으로 본 연구에서 제안한 시스템의 영화 평점 예측성이 더 우수함을 알 수 있었다.

Table 3. Differences in movie ratings between the LSTM language model and Naver(as of March 16)

Movie	LSTM Model		Naver Movie Ratings
	Average	RMSE	
Beasts Clawing at Straws	8.45	1.33	6.56
The Invisible Man	8.13	0.48	7.94
Dark Waters	8.95	0.66	8.56
1917	9.38	0.75	8.94

Table 4. Differences in movie ratings between the proposed model and Naver(as of March 16)

Movie	BRNN Model		Naver Movie Ratings
	Average	RMSE	
Beasts Clawing at Straws	8.4	1.03	6.56
The Invisible Man	8.04	0.42	7.94
Dark Waters	8.91	0.52	8.56
1917	9.31	0.67	8.94

그러나 평점이 다소 낮은 영화에 대해 두 시스템 모두 평균 제공근 오차가 높게 나타나 예측성이 떨어지는 것을 알 수 있었다.

6. 결론

본 연구에서는 양방향 순환 신경망 언어 모델을 이용한 Fuzzy-AHP 기반 영화 추천 시스템을 제안하였다. Fuzzy-AHP를 통해 사용자의 취향이나 관심사를 명확하고 객관적으로 반영할 수 있었으며, 양방향 순환 신경망 언어 모델을 이용한 학습 방법으로 사용자가 선호하는 영화를 예측할 수 있었다.

본 연구에서 제안한 추천 시스템의 성능을 확인하기 위해 그리드 서치를 이용하여 전체 단어 집합의 크기에 대한 학습 모델의 적합성을 확인하였으며, 본 시스템이 예측한 평점과 다른 언어 모델인 LSTM을 적용하여 예측

한 평점으로 실제 영화와의 평점 간 차이를 비교하였다. 그 결과 본 연구에서 제안한 학습 모델은 전체 단어 집합의 크기에 따른 평균 교차 검증 지수가 97.9%로 적합하다는 것을 확인할 수 있었으며, 실제 영화와의 평점 간 차이가 LSTM 언어 모델에 비해 상대적으로 우수함을 확인할 수 있었다.

REFERENCES

[1] Y. K. Kim. (2020. 02). Samsung Electronics Laughs at U.S 5G Expansion. *Economist*, 1522, 34-36.

[2] M. Y. Lee. (2020.02). 'Dream Force 2019' Conference Site Report. *Dong-A Business Review*, 291, 36-48.

[3] Y. K. Kim. (2019. 10). How far can we trust artificial intelligence?. *Economist*, 1503, 24-28.

[4] L. Adam. (2019. 10). Facebook wants to become a source for 'Trustworthy News'. *News Week*, 1392, 36-37.

[5] J. B. Bae. (2020. 01). Smart Robot Eyes with Deep Fake AI just like the Real Person. *Economist*, 1519, 50-52.

[6] Y. J. Kim. (2020. 01). The Age of New Gold Rush. *Economy Chosun*, 332, 10-11.

[7] C. M. Kwon. (2019). *Python Machine Learning Perfect Guide*. Paju: Wiki books.

[8] S. Sivaprasad, T. Joshi, R. Agrawal & N. Pedanekar. (2018). Multimodal Continuous Prediction of Emotions in Movie using Long-short Term Memory Networks. *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*, 413-419.

[9] G. Liu & X. Wu. (2019). Using Collaborative Filtering Algorithms Combined with Doc2Vec for Movie Recommendation. *2019 IEEE 3rd Information Technology, Networking, Electronics and Automation Control Conference*, 1-4.

[10] C. M. Lim, S. D. Kim & D. K. Yoon. (2011). AHP/DEA Decision Model using Triangular Fuzzy Number. *Proceedings of the Korean Institute of Industrial Engineers Spring Conference*, 2011(5), 612-617.

[11] B. Pang & S. Bai. (2013). An Integrated Fuzzy Synthetic Evaluation Approach for Supplier Selection based on Analytic Network Process. *Journal of Intelligent Manufacturing*, 24(2013), 163-174.

[12] E. Kinoshita, T. Oya. (2012). *Strategic Decision Making - AHP*. Seoul: Cheong Ram

[13] M. Schuster & K. K. Paliwal. (1997). Bidirectional Recurrent Neural Networks. *IEEE Transactions on Signal Processing*, 45(11), 2673-2681.

[14] A. Graves, A. Mogamed & G. Hinton. (2013). Speech Recognition with Deep Recurrent Neural Networks. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 6645-6649.

[15] C. Francois. (2018). *Deep Learning with Python*. Seoul: Gilbut.

[16] T. N. Duc, C. M. Tran, P. X. Tan & E. Kamioka. (2019). Bidirectional LSTM for Continuously Predicting QoE in HTTP Adaptive Streaming. *Proceedings of the 2019 2nd International Conference on Information Science and Systems*, 156-160.

[17] X. Han, Z. Wu, Y. G. Jiang & L. S. Davis. (2017). Learning Fashion Compatibility with Bidirectional LSTMs. *Proceedings of the 25th ACM International Conference on Multimedia*, 1078-1086.

[18] R. Sebastian, M. Vahid. (2019). *Python Machine Learning Second Edition*. Seoul: Gilbut.

[19] S. K. Reddy, V. Swaminathan & C. M. Motley. (1998). Exploring the Determinants of Broadway Show Success. *Journal of Marketing Research*, 35(3), 296-315.

[20] G. T. Cho, Y. G. Cho, H. S. Kang. (2005). *The Analytic Hierarchy Process*. Seoul: Dong Hyeon.

[21] J. T. Oh & S. Y. Lee. (2020). A Movie Recommendation System based on Fuzzy-AHP and Word2Vec. *Journal of Digital Convergence*, 18(1), 301-307.

[22] A. C. Müller, S. Guido. (2019). *Introduction to Machine Learning with Python*. Seoul: Hanbit Media.

오 재 택 (Jae-Taek Oh)

[상학원]



- 2015년 2월 : 대전대학교 IT경영공학과 (공학사)
- 2017년 2월 : 공주대학교 컴퓨터공학과 (공학석사)
- 2020년 8월 : 공주대학교 컴퓨터공학과 (공학박사)
- 관심분야 : 인공지능, 추천 시스템

· E-Mail : ohjt15@kongju.ac.kr

이 상 용 (Sang-Yong Lee)

[상학원]



- 1984년 2월 : 중앙대학교 전자계산학과 (공학사)
- 1988년 2월 : 일본동경공업대학대학원 총합이공학연구과 (공학석사)
- 1988년 3월 ~ 1989년 2월 : 일본 NEC 중앙연구소 연구원
- 1993년 2월 : 중앙대학교 일반대학원 전자계산학과 (공학박사)

- 1996년 9월 ~ 1997년 8월 : University of Central Florida 방문교수
- 1993년 8월 ~ 현재 : 공주대학교 컴퓨터공학부 교수
- 관심분야 : 인공지능, 컨텍스트 예측, 추천 시스템
- E-Mail : sylee@kongju.ac.kr