

# 딥러닝 기반 베트남 호텔 맞춤 추천

## 모바일 시스템 개발

오종현<sup>○</sup>, 서영수, 강현규\*<sup>1)</sup>

건국대학교 컴퓨터공학과

whdgus7592@naver.com, iulove37@naver.com, hkkang@kku.ac.kr

### A Mobile System Development which has Function of Vietnam Hotel Recommendation based on Deep Learning

Jong-Hyun Oh<sup>○</sup>, Young-Soo Seo, Hyun-Kyu Kang\*  
Department of Computer Engineering, Konkuk University

#### 요 약

본 논문은 아고다 사이트의 호텔 정보를 크롤링하여 사용자의 선호 호텔을 구글에서 제공하는 Tensorflow로 인공지능 딥러닝 학습하여 사용자가 선호하는 호텔을 맞춤 추천하는 애플리케이션의 설계 및 구현에 대하여 서술한다. 본 애플리케이션은 해외(베트남) 호텔을 취향에 맞게 추천받을 수 있도록 만들어진 애플리케이션으로 기존의 필터링 방식으로 추천하는 방식의 애플리케이션들과 달리 사용자의 취향을 딥러닝 학습을 통해 파악하고 최적의 호텔 정보를 추천하는 기능을 제공한다. 본 애플리케이션에 사용된 선호 호텔 예측 모델은 약 84%의 정확도를 보이며 추천 별점으로 표시되어 사용자가 각 호텔에 대해 얼마만큼 선호도를 갖는지 알 수 있다.

주제어: 추천, 학습형 애플리케이션, 호텔, 딥러닝

## 1. 서론

현재 우리 사회는 인공지능, 사물인터넷(IoT), 빅데이터, 모바일 등 첨단정보통신기술이 사회, 경제 전반에 융합되어 혁신적인 변화가 빠르게 일어나고 있다[1]. 그 중 인공지능에 해당하는 구글의 ‘알파고’ 이래로 기계 학습은 컴퓨터 분야에서 엄청난 주목을 받고 있다. 이러한 인기로 힘입어 전 세계적으로 기계 학습을 여러 분야에서 접목하려는 시도를 많이 볼 수 있다[2-5].

본 연구에서 초점을 둔 사용자 맞춤 추천을 숙박 시설(호텔) 시장 관점에서 봤을 때 대부분 기업이 맞춤 추천 서비스에 관심을 갖는다. 연구사례는 있으나 실질적인 기계 학습이 활용되지 않아 이용하지 못하는 실정이다. 본 연구는 기계 학습을 적극적으로 활용하여 학습한 내용에 적합한 맞춤 숙박 시설 추천만을 제공한다. 또한, 기존에 시도했던 포괄적인 형태의 맞춤 추천이 아닌 사용자 개인의 선호도를 통한 숙박 시설 맞춤 추천으로 접근한다.

## 2. 연구 목적 및 방법

### 2.1 연구 목적

본 연구의 목표는 사용자 취향에 맞춰 숙박 시설을 추천함에 있다. 숙박 시설의 crawling(크롤링) 정보와 사

용자의 선호도 정보를 활용하여 기계 학습에 기반을 둔 예측 알고리즘 모델을 학습한다. 이를 통해 사용자의 취향에 맞는 숙박 시설의 특성을 발견하고 이를 추천한다.

### 2.2 연구 방법

연구는 크롤링과 기계 학습 모델적인 측면에서 진행한다. 크롤링과 분석, 전처리, 기준선정 등의 과정을 통하여 학습 정보를 확보함과 동시에 데이터에 적합한 구조적인 기계 학습 모델을 구성하고 데이터와 모델을 접목한다. 이후 실험적 테스트 과정을 거쳐 산출된 결과물의 질적 향상을 위하여 데이터와 모델 양쪽 측면에서 결과 분석 및 개선 방안을 논의하고 이를 다시 적용한다.

## 3. 베트남 호텔 정보 수집

### 3.1 정보 크롤링 개요

해외여행 국가 중 한국인들이 많이 찾는 베트남 지역을 선정했고 해당 국가의 호텔을 조사했다. 베트남의 대표 도시인 ‘다낭’, ‘달랏’, ‘호치민’, ‘호이안’, ‘나트랑’, ‘하노이’ 6개 도시를 선정했고 크롤링했다. Python(파이썬) 언어를 이용했고 대표적으로 selenium, beautifulsoup, request 모듈을 사용했다[6].

### 3.2 호텔 URL 및 기본 정보 크롤링

아고다[7] 사이트에서 설정한 지역의 호텔 목록 HTML

1) Corresponding Author

코드를 파악해서 <a> 태그의 링크 속성을 txt 파일 형식으로 저장한다. 6개의 도시를 크롤링했고 대략 6,000개의 호텔 세부 링크를 확보했다. 호텔 URL 파일과 selenium 모듈을 이용해서 각각의 URL로 접근한다. 호텔 이름, 주소, 주변 관광지 정보의 태그와 클래스 정보를 참고해서 beautifulsoup 모듈로 크롤링한다[6].

### 3.2 평점 정보 크롤링

호텔 평점 정보의 경우 여행객 유형 정보와 후기 세부 항목 정보(평점 정보)로 나뉜다.

#### 3.2.1 후기 세부항목 정보(평점 정보)

평점 정보는 ‘숙소 청결 상태’, ‘부대시설’, ‘위치’, ‘객실의 안락함 및 쾌적함’, ‘서비스’, ‘가격대비 만족도’로 총 6가지로 구성된다. 모든 호텔이 같은 항목으로 구성되지 않고 6개의 항목 중 ‘객실의 안락함 및 쾌적함’이라는 항목이 없는 호텔이 존재한다. 그러므로 해당 항목을 제외한 5개 항목에 대해서만 크롤링했다. 평균 평점의 경우 크롤링한 5개 정보의 평균을 직접 계산해서 사용한다.

#### 3.2.2 여행객 유형 정보

‘전체 이용후기’, ‘커플/2인 여행객’, ‘나홀로 여행객’, ‘유아/아동 동반 가족 여행객’, ‘청소년 동반 가족 여행객’, ‘그룹/단체 여행객’으로 총 7가지의 여행객 유형이 존재한다. 각 유형별로 평점 정보가 다르므로 유형별 평점 정보 크롤링했다. 추가로 유형별 이용객 수도 같이 크롤링했다.

### 3.3 가격 정보 크롤링

가격 정보의 경우 호텔의 객실별, 월별로 다르다. 그래서 크롤링한 정보를 바로 사용하지 않고 평균값으로 사용한다. selenium 모듈을 이용해 하나의 호텔 URL에 대해서 임의로 7개의 월/일의 가격을 크롤링해서 7개 정보의 평균을 사용한다. 객실별 가격의 경우, 객실 세부 정보에 정원을 나타내는 아이콘이 존재한다. 해당 아이콘의 html 태그 이름을 이용해서 리스트로 분류해 가격 정보를 크롤링하였다. 총 6가지의 객실 유형을 분류했다. 분류된 파이썬 리스트에 각 7개 날짜에 대해서 크롤링한 가격이 추가되고, 리스트별로 평균 가격을 도출한 후 여행객 유형에 맞게 mapping(매핑)한다. ‘전체 이용후기’의 경우 모든 여행객 유형 가격 정보의 평균 가격으로 입력된다. 호텔별로 모든 유형에 대한 객실이 존재

하지 않는 경우 해당 유형은 호텔의 ‘전체 이용후기’의 평균 가격으로 입력한다.

### 3.4 정보 현황

크롤링 과정에서 평점 정보가 존재하지 않거나 아고다 [7]에서 서비스하지 않는 호텔은 제외하고 가격 정보가 0원인 호텔은 객실 검색이 되지 않는 호텔이므로 제외한다. 최종 크롤링 데이터는 호텔별 모든 여행객 유형에 대한 평점 정보를 포함해 27,563개이며 총 호텔 개수는 5,017개이다.

## 4. 사용자 정보 수집 및 적용

### 4.1 사용자 정보 수집

사용자의 취향 정보의 경우, 분석 및 전처리 과정을 거쳐 만들어진 정보 중 일부를 보여주고 사용자의 선호에 따라 수집되어야 하므로 사용자가 시스템을 사용하면서 입력한다. 사용자는 몇 호텔에 대해 선호도를 별점으로 입력하는데, 입력된 정보를 기반으로 사용자 취향 정보를 정의하고 수집한다.

### 4.2 학습 정보 적용

사용자 맞춤 추천에서는 사용자의 취향에 따라서 수집된 호텔 및 별점 정보를 바탕으로 모델을 학습한다. 호텔 정보 속성 중에서 호텔의 숙소 청결 상태 점수, 부대시설 점수, 위치 점수, 서비스 점수, 가격대비 만족도 점수, 평균 평점을 활용한다. 결과 값은 예측 추천 별점으로 나타난다. 추천에는 사용자가 아직 경험하지 않은 호텔들을 대상으로 적용하여 선호할 것으로 예측된 호텔들을 사용자에게 추천한다.

#### 정보 예시

국가 코드: VN(베트남)

도시 코드: DaL(달랏)

호텔 이름: 안 안 달랏(An An Dalat)

호텔 주소 : 7 KQH Nguyen Cong Tru, 깜리 폭포, 달랏, 베트남

호텔 식별 코드 : DaL10

여행객 유형 : 전체 이용후기

평점 : 4 / 6 / 6 / 6 / 6 / 5.6 (항목별 평점 및 평균 평점)

여행객 유형별 이용객 수 : 1

가격 : 12,877(원)

호텔 링크 : [https://www.agoda.com/ko-kr/an-an-dalat\\_2/h...](https://www.agoda.com/ko-kr/an-an-dalat_2/h...)

인기 명소 : 달랏 시장 - 1.48 km

근처 명소 : Domaine de Marie Church - 510m

## 5. 실험 및 분석

### 5.1 모델 개요

N개의 호텔을 사용자에게 보여주고 사용자는 해당 호텔에 대해 평가한다. 모델에 X 데이터로 호텔 정보가 입력되고, Y 데이터로 사용자 평가 정보가 입력된다. 모델은 학습을 통해 베트남 6개 도시에 존재하는 모든 호텔에 대해서 사용자 취향에 대해 예측하고 호텔을 맞춤 추천한다.

### 5.2 모델 구성

학습 정보 X는 7개로 구성한다. 숙소 청결 상태 점수, 부대시설 점수, 위치 점수, 서비스 점수, 가격대비 만족도 점수, 평균 평점, 가격으로 구성한다. 학습 정보 Y는 1개로 구성한다. 해당 호텔에 대한 사용자의 좋아요(1)/싫어요(0) 평가를 통해 입력된다. 학습 모델은 세밀한 학습을 위해 입력층, 은닉층, 출력층의 뉴럴 네트워크로 구성한다. learning rate는 0.001로 설정하고 vanishing gradient 현상을 초기에 방지하기 위해 ReLU함수를 적용했다. cost 값의 최저점을 찾는 gradient descent optimizer를 보완한 adam optimizer를 사용한다[9]. hypothesis는 0.5를 기준으로 작으면 0, 크면 1로 설정한다.

표 1. 예측 결과 일부

Predict	Percentage
True	89%
False	4%
True	100%
True	62%
False	9%
False	32%
True	51%
False	15%

테스트 데이터 셋은 150개로, 평균 평점이 0점부터 10점까지 1점대당 약 15개씩 분포된 150개 호텔에 대해 테스트 주관적으로 좋아요/싫어요 평가를 진행했다. 전체 150개를 7(105) : 3(45) 비율로 학습, 예측을 시도하였다. 정확도는 84%이다.

### 5.3 이용객 수에 따른 신뢰도 가중치 적용

이용객 수가 1명인 호텔과 10명인 호텔이 있을 때 신뢰도는 10명인 호텔이 더 높다. 평점이 아무리 높아도 이용객 수가 적으면 이보다 이용객 수가 높은 호텔부터

추천해주기 위한 메커니즘이다. 보유한 전체 데이터 셋에서 이용객 수 속성의 중간값은 10~20명이다. 모든 호텔 평균 평점의 평균은 8점이다. 이에 가중치의 중간값을 0.8로 설정하고 최저 0.6 최고 1.0으로 설정한다. 효율적인 계산을 위해 이용객 수의 기준을 16명으로 설정한다. 이용객 수에 따라서 가중치의 값이 변경되는데 0명의 경우 가중치 0.6을 기본값으로 시작해 한 명씩 증가할 때마다 기본값에 0.025를 더한다. 이용객 수가 16명 이상이면 신뢰도가 충분하다고 가정하여 가중치를 1.0으로 적용한다. 이 가중치는 입력정보인 5가지 평점에 곱해진 후 예측한다.

### 5.4 사용자 별점 평가 시스템 적용

기계 학습 모델 입장에서는 호텔 평점의 좋고 나쁨을 알 수가 없다. 이 점을 보완하고자 사용자가 별점을 매기는 방식으로 변경을 하였다. 0점~5점 사이 0.5점 단위로 평가할 수 있고 이 값이 Y 데이터로 입력이 된다. 사용자의 정확한 취향 파악을 위해 학습에 입력될 X 데이터를 12개로 설정했다. 12개의 데이터는 0~3점, 3점~6점, 6점대, 7점대, 8점대, 9점대, 10점대로 구성되고 랜덤 함수를 이용해서 범위에 있는 호텔을 무작위로 가져와서 사용자에게 보여준다. 모델은 사용자가 낮은 점수대를 시작으로 높은 점수대까지 직접 별점을 매김으로써 취향을 파악한다. 가격의 경우 사용자 취향 범위를 파악하기 위해서 더 많은 학습 정보 X가 필요하므로 학습 정보에서 가격 정보를 제외한다. 가격 정보는 추후 사용자가 선호하는 가격대를 지정할 때 필터링에 사용한다.

중요하게 생각하는 항목을 선택할 수 있도록 모델에 적용하였다. 학습 정보를 사용자로부터 받아들일 때 5가지 세부항목 중에 중요하게 생각하는 항목 최대 3개를 선택할 수 있다. 선택된 항목에 대해서는 원래값에 제공만큼 적용하여 다른 항목들보다 높은 평점으로 학습을 시켜 효과적으로 예측 추천하도록 적용하였다.

### 5.5 모델 분석

호텔 추천 연구사례로는 Expedia 사용자 정보를 기반으로 하는 연구가 있다. 약 3천 7백만의 사용자 정보로 동행 아이들 수, 예약한 방의 수, 예약 위치, 여행 패키지 여부 등 사용자 개인의 정보를 특징으로 한다[9]. 호텔 정보를 특징으로 하는 현 연구와는 차이점이 있다.

현 연구의 모델을 분석하자면 초기에 [표 2]처럼 이용객 수가 10명 이하의 낮은 인원인 사실상 신뢰도가 떨어지는 정보들이 상위권에 분포하는 경우가 많았다.

표 2. 신뢰도 가중치 적용 전 예측 정보

호텔	추천율	이용객 수
A	100%	25명
B	100%	1명
C	100%	45명
D	100%	25명
E	100%	14명

표 3. 신뢰도 가중치 적용 후 예측 정보

호텔	추천율	이용객 수
A	100%	25명
D	100%	25명
F	100%	15명
G	100%	30명
H	98%	180명

이용객 수 신뢰도 가중치 적용 이후 이용객 수가 적은 데이터가 하위권으로 내려갔고, 이용객 수에 따라서 신뢰도에 대한 16명을 충족하지 못하면 추천율에 페널티가 부여된다.

표 4. 좋아요/싫어요 평가 시스템 예측 결과 일부

추천율	이용객 수	항목별 점수(평점)				
98%	180명	9.6	9.4	9.5	9.6	9.7
96%	1명	10	10	10	10	10
95%	5명	8.4	7.6	9.2	9.2	8.4
94%	7명	9.6	7.7	7.7	8.6	8.6
94%	6명	9.3	9.0	9.0	9.7	9.7

초기 모델의 경우 좋아요(1) / 싫어요(0) 정보를 기반으로 학습한다. 따라서 점수의 좋고 나쁨을 파악하는 것이 아니라 선호하는 특정 구간을 기준으로 두고 평가를 한다. [표 4]와 같이 평점이 10점 만점인 데이터가 다른 데이터보다 낮은 순위에 있다.

표 5. 별점 평가 시스템 예측 결과 일부

추천 별점	이용객 수	항목별 점수(평점)				
5개	25명	9.7	9.8	9.9	9.9	9.9
4.5개	180명	9.6	9.4	9.5	9.6	9.7
4.5개	15명	10	9.7	9.7	9.9	9.9
4개	30명	9.5	8.7	9.0	9.5	9.7
4개	13명	9.7	9.2	9.4	10	10

사용자가 별점으로 평가하는 방식을 도입하고 나서는 완전히 해소되었다. [표 5]에서는 평점 정보를 기준으로 숙소 청결(1번 행)과 만족도(5번 행)를 중요 항목으로

설정 후 예측을 하였다. 일부 예측 결과는 이용객 수로 인해 페널티가 부여되었으며 기존에 보였던 애매한 순위가 보이지 않는다. 또 중요 항목 가중치 적용으로 인해 더 정당한 결과를 보여주고 있다.

## 6. 시스템 적용

### 6.1 눈치코칭\_베트남 호텔

“눈치코칭\_베트남 호텔”은 베트남 호텔 정보 및 예약, 관광지 등 해외여행에서 필요한 호텔 정보, 관광지 정보를 제공하는 모바일 기반의 추천 시스템이다. 호텔 정보에 따른 사용자의 선호도를 외부 데이터베이스에 저장하고 이를 토대로 모델의 학습을 통한 추천 알고리즘을 걸쳐 사용자에게 추천한다. 해당 서비스는 장치에서 3G, WIFI 환경을 통하여 외부 데이터베이스에 접근하고 정보를 받아옴으로써 수행된다.

### 6.2 베트남 호텔 인공지능 추천

“눈치코칭\_베트남 호텔”은 인공지능 알고리즘을 사용하여 호텔을 추천하는 시스템을 가진다. 인공지능을 기반으로 한 추천을 적용하여 사용자 맞춤 추천 서비스를 제공한다. 구체적으로 호텔의 항목을 보고 사용자가 “별점”을 매겨 사용자 선호 정보 수집 기능 및 ‘인공지능 추천’ 서비스를 제공하였다.

#### 6.2.1 취향파악

기기의 고유 ID를 확인하여 해당하는 정보가 존재하지 않으면 신규 사용자라 판단한다. 취향파악 화면에서는 신규 사용자의 선호도를 파악한다. 사용자는 해당 화면에서 호텔 항목 중 선호하는 항목을 최소 1가지에서 최대 3가지를 선택한다. 선호하는 가격대를 사용자가 직접 작성하여 가격을 필터링하는 데 사용한다.



그림 1. 취향파악

#### 6.2.2 취향찾기

취향파악을 완료한 신규 사용자는 취향찾기를 진행한



다. 사용자는 총 12번의 호텔 취향찾기를 한다. 임의의 호텔로 취향찾기를 하며 임의의 호텔의 여행객 유형은 ‘전체 이용후기’를 기준으로 한다. 호텔의 정보와 사용자의 별점은 기기의 고유 ID를 바탕으로 한 사용자 외부 데이터베이스 테이블에 저장하며 인공지능 모델의 학습을 할 취향 정보가 된다. 12번의 취향찾기를 모두 완료하면 취향 정보를 기반으로 모델의 학습이 이루어진다.

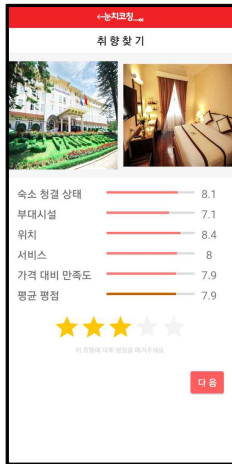


그림 2. 취향찾기

### 6.2.3 메인화면

메인화면에서는 먼저 기기의 고유 ID를 확인하고, 이를 통해 해당 사용자의 외부 데이터베이스 테이블에 접속하여 학습된 정보를 기반으로 하는 취향 정보에서 1순위로 추천하는 호텔을 가장 상단에 표시한다. 하단 레이아웃에는 모델 학습을 통한 예측 별점 순서로 전체 호텔을 추천하는 AI 호텔추천 레이아웃과 사용자가 도시, 여행객 유형 그리고 가격을 필터링해 추천받을 수 있는 AI 호텔검색 레이아웃이 존재한다.

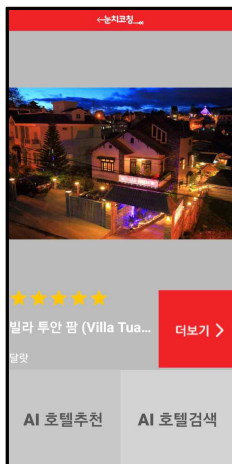


그림 3. 메인화면

### 6.2.4 AI 호텔추천

메인화면과 마찬가지로 기기의 고유 ID를 통해 사용자의 취향 정보를 확인한다. 정보가 존재한다면 외부 데이

터베이스에서 추천 호텔 목록을 가져온다. 추천 호텔 목록은 인공지능 모델의 예측 별점이 높은 순으로 기기 화면에 출력한다. (예측 별점의 경우 0~5점으로 구분되며 별 1개당 별점 1점을 나타낸다) 해당 호텔의 사진, 도시, 예측 별점 그리고 이름을 함께 표시하여 사용자에게 제공한다.

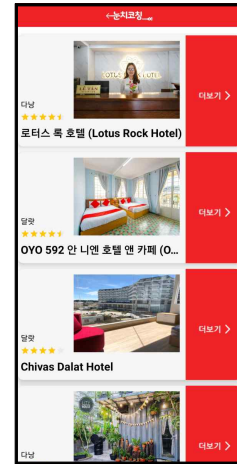


그림 4. 호텔추천 목록

### 6.2.5 AI 호텔검색 및 목록

메인화면에서 AI 호텔검색 레이아웃을 선택하면 [그림 5]를 보여준다. 사용자가 도시, 여행객 유형 그리고 가격을 설정하고 검색 버튼을 클릭하면 인공지능 모델이 학습한 취향 정보에서 해당 내용을 필터링한 추천 호텔 목록을 가져와서 [그림 6]처럼 출력한다. 이하 내용은 AI 호텔 추천과 같다.

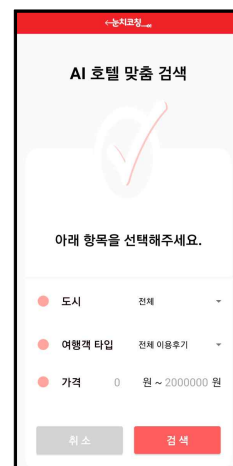


그림 5. 검색화면

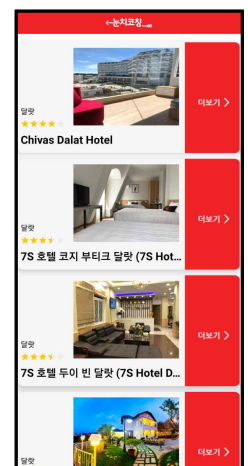


그림 6. 맞춤검색 목록

### 6.2.6 상세화면

상세화면은 목록으로 제공되는 호텔에 대해서 상세한 정보를 제공하기 위한 기능이다. 호텔명, 도시명, 주소, 항목별 점수, 가격, 인기 관광지, 주변 관광지, 예약을 위한 버튼을 제공한다. 또한 ‘평가하기’ 버튼은 사용자의 취향 정보를 수집하는 역할을 한다.



그림 7. 상세화면

### 6.2.7 상세화면 - 평가하기

평가하기는 방문 유/무를 토대로 각각 다른 화면을 제공한다. 방문하지 않은 호텔의 경우 사용자는 해당 호텔에 대해 별점밖에 매길 수 없다. 방문한 호텔의 경우 사용자는 별점과 추가로 호텔을 이용하고 해당 호텔에 대한 항목별 점수를 매길 수 있다. 이 정보는 장치 고유 ID를 토대로 사용자의 외부 데이터베이스 테이블에 추가되어 이후 취향 정보로 사용한다.

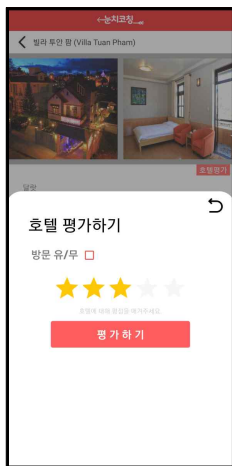


그림 8. 방문 무



그림 9. 방문 유

## 7. 결론

사용자 취향 파악 후 선호하는 베트남 호텔을 예측하기 위해 모델을 조금씩 개량해 나갔다. 초기 모델의 경우 학습이 잘 이루어지지만, 특정 기준선을 기준으로 추천이 이루어져서 높은 점수와는 상관없이 추천의 신뢰도가 떨어지는 것을 보여주었다. 이를 보완하기 위해 사용자가 별점을 부여하는 방식으로 변경하였다. 이를 통해 특정 기준선이 아닌 사용자의 선호 분포를 파악함으로써 평점에 따라 호텔 추천이 달라지는 것을 보여준다. 더 명확한 호텔 추천을 위해 이용객 수에 따른 신뢰도 페널티를 부여하고, 사용자가 선택한 중요 항목에 더 높은

가중치를 부여하였다.

많은 시중 호텔 예약 사이트는 필터링으로 추천이 이루어지거나 전체 사용자 데이터를 베이스로 추천되는 경우가 많다. 현재 구현한 시스템은 딥러닝 기반으로 다양한 항목의 호텔 평점 데이터를 이용해서 사용자의 취향을 파악한 후 추천을 시도한다. 사용자 개인별로 추천 서비스를 제공할 수 있고, 딥러닝 기반이기 때문에 세밀한 추천이 가능하다.

현재 구현한 시스템은 84% 정확도를 보인다. 전체적인 정확도를 90% 수준으로 올리고 모델의 학습 시간을 단축하는 것이 목표이다.

## 참고문헌

- [1] 한국정보통신기술협회, “제4차 산업혁명”, [http://www.tta.or.kr/data/weeklyNoticeView.jsp?k\\_num=5228](http://www.tta.or.kr/data/weeklyNoticeView.jsp?k_num=5228), TTA 신규용어 소개, 2017.
- [2] 쿠지라 히코우즈쿠에, "머신러닝, 딥러닝 실전 개발 입문", 위키북스, 2017.
- [3] 김승현, 정용주, "처음 배우는 머신러닝 (사이킷런으로 기초부터 모델링 실전 예제 문제 해결까지)", 한빛미디어, 2017.
- [4] 루카 마사론, 알베르토 보셰티, "실전활용! 텐서플로 딥러닝 프로젝트", 위키북스, 2018.
- [5] 바라스 람순다르, 레자 자데, "한 권으로 끝내는 딥러닝 텐서플로", 한빛미디어, 2018.
- [6] 아고다(Agoda), <https://www.agoda.com/>
- [7] 박정태, “파이썬으로 배우는 웹 크롤러”, 정보문화사, 2018.
- [8] 김성훈, 모두를 위한 머신러닝/딥러닝 강의, <http://hunkim.github.io/ml/>
- [9] Michael Arruza, “The Automated Travel Agent: Hotel Recommendations Using Machine Learning”, CS 229 Machine Learning Final Projects, Spring 2016, June 2016.