

# 지식 기반 추천 시스템을 이용한 식물 추천 시스템

전기범<sup>1</sup>, 이유진<sup>2</sup>, 안민하<sup>3</sup>, 김용규<sup>4</sup>, 김영종<sup>✉</sup>

<sup>1,2,3,4</sup>승실대학교 소프트웨어학부

✉승실대학교 소프트웨어학부

kibeom1145@gmail.com, yujin038@soongsil.ac.kr,

dks4857@gmail.com, e2yong98@gmail.com,

opensys@gmail.com

## Plant Recommendation System Using Knowledge-Based Recommender System

Kibeom Jeon<sup>1</sup>, Yujin Lee<sup>2</sup>, Minha Ahn<sup>3</sup>, Yonggyu Kim<sup>4</sup>, Youngjong Kim<sup>✉</sup>

<sup>1,2,3,4</sup>School of Software, Soongsil University

✉School of Software, Soongsil University

### 요 약

‘홈가드닝’ 시장의 성장이 가속화되면서 식물에 관심을 가지는 사람들이 많아지고 있으나, 초보자의 경우 식물 육성에 있어 어려움이 존재한다. 본 논문에서는 이러한 한계를 극복하기 위한 방안으로 일조량, 계절, 크기 등 세분화된 조건을 바탕으로 하는 지식 기반 추천 시스템을 적용한 식물 추천 시스템을 제안하고 있다. 해당 시스템을 통해 사용자의 환경에 맞는 식물을 추천하여 육성 및 관리를 돕고자 한다.

### 1. 서론

코로나-19로 집에서 시간을 보내는 사람들이 늘어나면서, 식물 키우기가 인기 취미로 떠오르고 있고, 관련 산업 매출은 급격하게 증가하는 중이다. [1]

최근에는, 식물 키우기에 대한 사람들의 높은 관심을 증명이라도 하는 것처럼, 집을 뜻하는 홈(Home)과 원예를 뜻하는 가드닝(Gardening)의 합성어인 ‘홈가드닝’, 친구처럼 정서적인 교감과 위안을 얻게 하는 식물을 뜻하는 ‘반려 식물’과 같은 신조어도 등장하였다.

‘반려 식물’과 ‘홈가드닝’에 대한 사람들의 관심이 증가하면서, 식물을 키워보고자 하는 사람들이 늘어나는 중이지만, 방대한 종류의 식물 중에 자신의 수준과 주위 환경에 맞는 ‘반려 식물’을 찾기가 어렵다는 점이 진입장벽이 되면서 사람들로 하여금 어려움을 겪게 만들고 있다.

이와 같은 문제점을 해결하기 위해 본 논문에서는 지식 기반 추천 시스템(Knowledge-Based Recommender System)을 이용한 식물 추천 시스템(Plant Recommendation System)을 제안하고자 한다.

### 2. 연구 동향

추천 시스템에서 사용되는 기법에는 콘텐츠 기반 필터링, 협력 필터링, 지식 기반 추천 시스템 등 다

양한 종류가 존재한다. 본 단락에서는 콘텐츠 기반 필터링, 협력 필터링, 지식 기반 추천 시스템 각각의 방식에 대해 살펴보고자 한다.

#### 2.1 협력 필터링 (Collaborative Filtering)

협력 필터링은 전자 문서를 위해 설계된 메일 시스템인 태피스트리(Tapestry)에서 문서 필터링을 위해 처음 소개된 기법이다. [2] ‘특정 아이템에 대해 선호도가 유사한 고객들은 다른 아이템에 대해서도 비슷한 선호도를 보일 것’이라는 기본 가정을 바탕으로 사용자 혹은 아이템 간의 유사도를 기반으로 선호도를 예측하는 방법이다. [3]

해당 기법은 사용자의 평가에 기반하여 동작하기 때문에 사용자 경험을 더 많이 축적할수록 정확성이 상승한다는 특징을 가진다.

#### 2.2 콘텐츠 기반 필터링 (Content-Based Filtering)

콘텐츠 기반 필터링은 사용자가 아닌 아이템을 중심으로 하는 필터링 기법이다. 아이템 사이의 유사성을 측정하여, 사용자가 과거에 평가했던 아이템과 유사성이 높은 아이템을 추천하는 방식이기 때문에, 사용자 개인의 독립적 정보를 필터링의 기준으로 삼는다는 점에서 정확도를 위해 다수의 사용자 경험을 필요로 했던 협력 필터링 기법과는 차이가 존재한다.

다.

### 2.3 지식 기반 추천 시스템 (Knowledge-Based Recommender System)

지식 기반 추천 시스템은 사용자가 아이템을 탐색할 때 선호도를 명시하여 탐색할 수 있도록 검색 기준을 세분화하는 기법이다. 사용자는 요구사항에 대한 조건을 설정하기 때문에, 해당 기법은 앞서 나왔던 콘텐츠 기반 필터링, 협력 필터링 기법과 달리 사용자의 과거 정보나 취향, 아이템 간의 유사성이나 등급에 의존하지 않는 기법이라고 할 수 있다.

해당 기법을 적용한 추천 시스템의 경우, 사용자가 설정한 조건에 맞추어 동작하므로 필터링 과정은 사용자 개인의 취향과 무관하다. 따라서 특정 사용자에 대한 정보를 수집하지 않으며, 그렇기 때문에 지식 기반 추천 시스템은 사용자 데이터 규모에 정확성이 좌우되는 협력 필터링의 단점을 보완하는 기법이라고 할 수 있다. [4]

### 3. 식물 추천 시스템

지식 기반 추천 시스템은 사용자의 수가 적거나, 사용자의 과거 이용 기록이 없더라도 사용자에게 적합한 아이템을 추천할 수 있는 기법이므로 본 논문에서는 지식 기반 추천 시스템을 적용하여 식물 추천 시스템을 만들고자 한다.

#### 3.1 식물 데이터베이스

식물 데이터베이스는 자가 구축한 데이터베이스를 사용하며, 각 식물 데이터는 이름, 종, 키우는 난이도, 물주는 주기 등, 여러 값을 포함하고 있다.

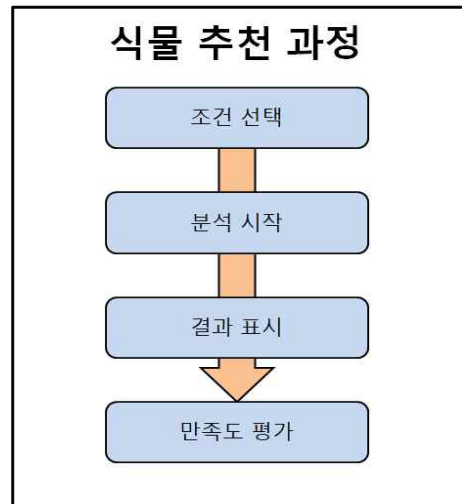
이름	종	조건1	조건2	조건3	조건4
식물1	종1	상	3	4	3
식물2	종2	중	7	5	3
식물3	종3	하	4	3	4
식물4	종4	상	5	2	4
식물5	종5	중	7	1	2

<표 1> 식물 데이터베이스 예시

데이터베이스에 새로운 데이터를 추가하거나, 데이터베이스의 기존 데이터를 변경, 삭제할 때는 다음과 같은 규칙을 지켜야 한다.

1. 데이터의 키(Key) 값은 식물의 이름이며, 식물의 이름은 중복될 수 없다.
2. 데이터를 추가할 때, 데이터 안에는 Null 값이 없어야 한다.
3. 데이터를 변경할 때, 데이터를 변경하는 합당한 이유가 존재해야 한다.(새로운 연구결과 발표, 잘못된 데이터 값 입력, 데이터 값과 사용자들의 체감 수치가 상이 할 경우 등)
4. 데이터를 삭제하는 경우는, 해당 식물이 세상에 존재하지 않게 되었거나, 원래 세상에 존재하지 않는 식물일 경우이다.

#### 3.2 식물 추천 방식



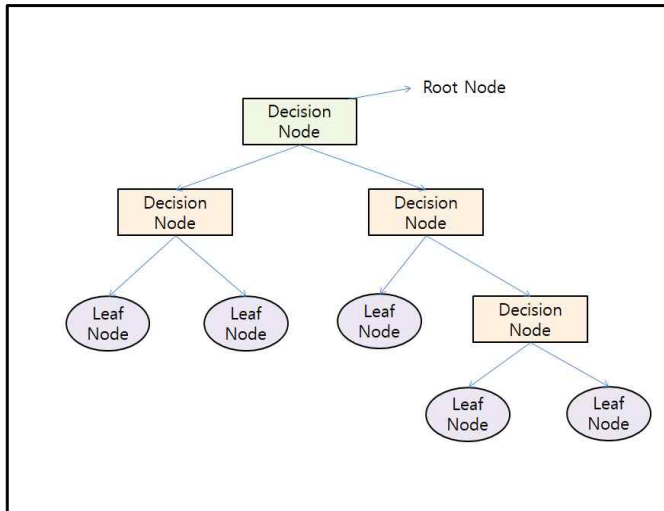
<Figure 1> 식물 추천 과정

<Figure 1>은 식물 추천 과정을 간략하게 보여주는 그림이다. 먼저, 사용자는 식물을 추천받기 위해 사용자에게 맞는 조건을 선택한다. 각 조건은 식물 데이터베이스에 있는 조건들과 동일하며, 식물 추천 시스템은 각 조건에서 선택된 값을 이용하여 식물을 필터링하고, 시스템의 개발자 측에서 정한 기준(검색 순, 평점 순 등)에 따라, 필터링이 된 식물 중 일정 개수만큼의 식물을 선택하여 추천 결과에 표시한다. 추천이 종료되고, 사용자는 추천 결과에 대한 만족도를 평가하게 된다.

#### 3.3 피드백(Feedback)을 통한 추천 개선

3.2에서 설명한 방법으로는 특정 조건에 대한 결

과로 계속 같은 식물들을 추천해 주기 때문에, 추천 결과의 다양성이 떨어지고, 새로운 식물을 추천받고 싶은 사용자들이 불만족할 확률이 높다. 따라서 3.3에서는 3.2의 추천 방법을 의사결정 나무(Decision Tree) 기법을 통해 개선하여 사용자가 추천 결과에 만족할 확률을 높이고자 한다.



<Figure 2> 의사결정나무 예시

의사결정 나무란 가장 유명한 분류 기법으로, <Figure 2>와 같이 나무의 맨 위 노드는 뿌리 노드(Root Node)라고 한다. 나무 안에 있는 노드 중 잎 노드(Leaf Node)가 아닌 노드는 속성에 대한 테스트를 나타내고, 가지(Branch)는 테스트의 결과를 나타내며, 잎 노드는 예상 결과를 나타낸다. [5]

의사결정 나무에서 의사결정 노드(Decision Node)들이 배치되는 순서는 정보 획득량(Information Gain)이 높은 순서로 정해지는데, 정보 획득량은 엔트로피(Entropy)를 통해 구할 수 있다.

먼저, 엔트로피는

$$Entropy(S) = - \sum_{i=1}^n P_i \log P_i$$

으로 정의한다. S는 사례들의 집합이고, n는 사례들이 속하는 클래스의 총 개수이며, P<sub>i</sub>는 S중에서 클래스 i에 속하는 사례들의 비율이다.

다음으로, 정보 획득량은

$$IG = Entropy(P) - [W_{avg} Entropy(C)]$$

로 정의한다. Entropy(P)는 부모노드에서 얻은

엔트로피를 의미하고  $W_{avg} Entropy(C)$ 는 자식노드들에서 얻은 엔트로피의 가중평균을 의미한다.

조건1	조건2	조건3	조건4	조건5	만족
상	3	4	3	2	yes
중	7	5	3	4	no
하	4	3	4	4	no
상	5	2	4	3	yes
중	7	1	2	3	yes

<표 2> 사용자 만족/불만족 데이터 예시

지금부터, 개선된 방법에 대해 설명하려고 한다. 먼저, <Figure 1>의 만족도 평가에서 얻은 사용자 만족/불만족 데이터를 의사결정 나무로 모델링 하고, 나무의 크기를 조건의 개수만큼만 자라게 하여 과대적합 되는 것을 방지한다. 이후, 사용자가 추천 조건을 선택하였을 때, 사용자가 선택한 조건 데이터는 의사결정 노드에 의해 분류되며 잎 노드까지 내려간다. 마지막으로, 데이터가 최종 분류된 위치의 결과가 만족이라면, 추천 결과는 3.2의 방식과 동일하게 표시되지만, 예측 결과가 불만족이라면 추천 결과는 기존의 방식이 아닌 시스템 개발자가 정한 다른 방식으로 변경되어 표시된다.

이와 같이 사용자의 만족/불만족을 예측하여 추천 결과를 다르게 보여준다면, 추천 결과의 다양성을 보장할 수 있고, 사용자가 추천 결과에 만족할 확률 또한 높일 수 있을 것이다.

#### 4. 결론

본 논문에서는 콘텐츠 기반 필터링, 협력 필터링, 지식 기반 추천 시스템 등 다양한 추천 시스템의 기법에 대해 알아보면서 각각의 특징에 대해 살펴보았다. 또한 지식 기반 추천 시스템을 적용한 식물 추천 시스템의 절차에 대해 설명하고, 의사결정 나무 기법을 사용한 개선 방식을 제안하였다. 해당 시스템을 통해 사용자가 식물 추천을 필요로 할 때마다 설정된 조건을 기반으로 자신의 환경에 맞는 식물을 추천 받을 수 있다는 효과를 기대할 수 있을 것이다.

그러나, 본 논문에서 제안한 식물 추천 시스템은 사용자의 수가 협업 필터링 기법을 적용하기에 충분하지 않고, 사용자의 과거 경험 데이터가 없는 경우

를 고려하여, 지식 기반 추천 시스템만을 적용하였기 때문에, 축적된 사용자 데이터가 많아진다면 2가지 이상의 추천 기법을 융합한 추천 시스템에 비해 추천의 성능이 낮을 수밖에 없다는 한계가 존재한다.

따라서 본 논문에서 제안한 시스템은 사용자 데이터를 수집하기 시작하는 초기부터 사용자 데이터가 충분히 쌓여 협업 필터링 기법과 콘텐츠 기반 필터링 기법을 적용할 수 있기 전까지의 기간에 적용되면 적절할 것이라 생각된다.

### 참고문헌

[1]“빅 데이터로 알아보는 반려식물”, 농사로, 2021년9월29일 수정, 2022년4월13일 접속, <https://nongsaro.go.kr/portal/ps/psv/psvr/psvre/curationDtl.ps?menuId=PS03352&srchCurationNo=1696>

[2] Goldberg, D., Nichols, D., Oki, B. M., and Terry, D, Using collaborative filtering to weave an information tapestry, Communications of the ACM, 35(12), 61-70, 1992

[2] 손지은, 김성범, 김현중, 조성준, 추천 시스템 기법 연구 동향 분석, 대한산업공학회지, 41(2), 185-208, 2015

[4] Allen Kent, Encyclopedia of Library and Information Science, New York, CRC Press, 2000

[5] 박득희, 김혜경, 최일영, 김재경, 추천시스템관련 학술논문 분석 및 분류, 지능정보연구, 17(1), 144, 2011