

추천 시스템 개요

I. 도입

II. 추천 시스템 기법

- 가. 콘텐츠 기반 필터링 방식
- 나. 협업 필터링 방식
- 다. 딥러닝 기반 추천 시스템

III. 유튜브 추천 시스템

IV. 추천 시스템과 인간 행동

- 가. 콘텐츠 소비자의 측면
- 나. 콘텐츠 생산자의 측면

V. 결론



김명필
KAIST 기술경영학부

I. 도입

지인 중에 좋은 영화, 책, 음악, 웹툰을 추천해 주는 이들이 있을 것이다. 영화광 친구가 오랜만에 연락해 영화를 추천해 주면 고맙다. 내가 좋아할 만한 영화를 선별해 추천해 주면 더욱 고맙다. 이렇듯 좋은 추천은 두 가지 요소를 함께 고려해야 한다. 콘텐츠의 품질을 평가할 수 있어야 할 뿐만 아니라, 내 취향까지도 고려해야 한다.

예전부터 우리 경험의 많은 부분은 추천에 의해 결정되어 왔다. 온라인 상의 개인 맞춤형 서비스가 일상화되기 전까지는 편집자에 의한 큐레이션(curation)이 주된 추천 방식이었다. 뉴스 방송 편집자가 뉴스의 순서와 시간을 정하고, 신문사 '데스크'가 지면에서의 위치와 그 분량을 정한다. 이러한 큐레이션은 여전히 사회 전반에 걸쳐 활용된다. 온라인 쇼핑몰이나 앱 스토어는 편집자 추천(editor's pick) 또는 MD(merchandise) 추천 항목을 눈에 띄는 곳에 배치한다. 뉴스레터 서비스는 특정 주제에 관한 기사들만 추려서 보내 주고 있다. 하지만 큐레이션은 개인 취향을 고려하지는 못한다. 그래서 개개인이 알아서 자신의 취향에 맞는 큐레이션 결과를 찾아야 한다. 누구나 한 번쯤 영화 평론가가 5점을 준 영화를 보고 즐겼던 경험이 있을 것이다. 자신의 취향과는 맞지 않는 영화 평론가를 믿은 결과다.

원래 개인 맞춤형 추천은 값비싼 서비스이다. 백화점 VIP 고객을 위한 퍼스널 쇼퍼 서비스, 고객 자산 보유자를 위한 은행 프라이빗 बैं킹 서비스가 그 예이다. 하지만 빅데이터와 머신러닝 기술은 개인 맞춤형 추천의 비용을 크게 낮추었다. 온라인 서비스들은 내가 관심이 있을 법한 상품, 투자 기회, 영화, 도서, 동영상을 추천해 준다.

인류는 이제 하루에만 250만 테라바이트(TB)의 데이터를 생산하고 있다고 한다.¹⁾ 추천 시스템이 없다면 이렇게 방대한 데이터 중에서 우리가 필요한 콘텐츠를 찾아서 소비하는 일은 불가능하게 되었다. 넷플릭스에서 시청하는 영화의 80%는 추천의 결과이다. 나머지 20%만이 이용자 검색으로부터 나온다.²⁾ 이처럼 추천시스템은 우리 삶의 핵심적 부분이 되었다.

이러한 추천 시스템의 발전과 더불어 그 법적, 윤리적 위험에 대한 우려도 제기된다.³⁾ 자극적·선정적 콘텐츠 또는 허위 정보가 추천되거나, 맞춤형 추천을 위해 과도한 개인정보를 수집하는 문제가 지적된다. 추천 시스템이 콘텐츠 소비 시간을 늘리고, 개인이 이에 '중독'된다는 우려도 있다. 추천 시스템의 불투명성, 사회적 차별 및 편향 문제, 추천 시스템의 피드백 효과로 인해 반대 견해를 접할 기회를 잃게 되는 문제 등도 지적된다. 그 결과 추천 시스템에 대한 법적 규율의 필요성이 대두되고 있기도 하다.

이 글은 이러한 법적 논의의 전제로서, 현재 활용되고 있는 추천 시스템의 기술적 기초를 전반적으로 소개하고자 한다. 특히 최근 관심이 높아진 유튜브의 추천 시스템의 구조를 살펴보고, 추천 시스템이 콘텐츠 소비자와 생산자에게 어떠한 영향을 가져올 수 있는지 검토한다.

1) Bernard Marr, "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read", Forbes, 2018. 5. 21.
2) Carlos A Gomez-Urbe and Neil Hunt, "The Netflix recommender system: Algorithms, business value, and innovation", TMIS 6, 4(2016).
3) 추천 시스템의 윤리적 쟁점에 관한 정리된 글로는 Silvia Milano, Mariarosaria Taddeo and Luciano Floridi, "Recommender systems and their ethical challenges", AI & SOCIETY 35.4 (2020) 참조.

II. 추천 시스템 기법

가. 콘텐츠 기반 필터링 방식

여러분이 영화 추천 알고리즘을 작성한다고 생각해 보자. 가장 쉽게 떠올릴 수 있는 방법은 무엇인가? 아마도 영화의 장르, 감독, 배우에 관한 데이터베이스를 이용하는 방법을 떠올릴 것이다. 이용자가 이전에 시청한 영화 목록을 토대로 어떠한 장르, 감독, 배우를 선호하는지 정보를 파악한 다음, 그의 프로필에 부합하는 영화를 추천하고자 할 것이다. 요컨대, 사용자 프로필 특성과 유사한 영화를 추천하는 것이다.

이처럼 해당 아이템의 내용과 이용자의 프로필 간의 유사성에 기반한 추천을 콘텐츠 기반 필터링(Content-based Filtering)이라 한다.⁴⁾ 예컨대 도서 추천 시스템은 이용자가 과거 구매하였거나 관심 목록에 포함시켜 둔 도서의 목록을 입력 정보로 활용한다. 여기서 이용자가 어떠한 장르나 주제의 도서에 관심이 있는지 추출하여 그 사용자 특성에 부합하는 도서를 추천하는 것이다.⁵⁾

콘텐츠 기반 필터링은 (1) 콘텐츠 분석, (2) 사용자 프로필 학습, (3) 필터링의 세 가지 단계로 구성된다.⁶⁾

우선, 콘텐츠 분석 단계에서는 추천 대상 아이템으로부터 관련된 특성(feature)을 추출한다. 영화 등과 같이 관련 정보 데이터베이스가 잘 구축되어 있는 경우에는 이를 활용하면 되겠지만, 웹페이지, 뉴스, 상품 설명 등과 별도의 특성 추출이 필요한 경우도 있다. 콘텐츠 분석기(content analyzer)는 비정형화된 정보로부터 이를 잘 대표할 수 있는 특성을 추출하여 구조화된 형태의 정보로 변환한다. 예컨대 영화 '타이타닉'은 드라마와 로맨스 장르로, '주라기 공원'은 액션, 어드벤처, 공상과학 장르로 구분할 수 있을 것이다. 논의를 간단히 하기 위해 영화의 특성을 드라마(x_1), 로맨스(x_2), 액션(x_3), 어드벤처(x_4), 공상과학(x_5) 5개의 장르로만 구분한다고 해 보자. 그러면 '타이타닉'은 (1, 1, 0, 0, 0), '주라기 공원'은 (0, 0, 1, 1, 1)와 같은 형태로 표현할 수 있다.⁷⁾

다음은 사용자 프로필 학습 단계이다. 이용자가 선호할 콘텐츠 특성을 추출하기 위하여 머신러닝 기법이 널리 활용된다. 머신러닝을 위한 주된 학습 데이터는 이용자의 인적 사항 및 이용자가 이전에 긍정 또는 부정 피드백을 제공한 콘텐츠이다(영화, 구매한 제품, 방문한 웹페이지 등).⁸⁾ 예를 들어 어떤 이용자가 '타이타닉'은 높은 별점을 주었지만, '주라기 공원'은 낮은 별점을 주었다면, 그 이용자는 드라마(x_1), 로맨스(x_2) 장르의 선호도는 높는데 비해 액션(x_3), 어드벤처(x_4), 공상과학(x_5) 장르의 선호도는 낮은 것

으로 파악할 수 있을 것이다. 한편, 만약 그 이용자는 로맨스와 공상과학이 결합된 영화 '그녀(Her)'를 좋아한다면, 공상과학 장르는 중립적 선호를 갖는 것으로 생각할 수 있다. 그러면 그 사용자 프로필은 (0.7, 0.9, 0.1, 0.1, 0.5)와 같은 형태로 표현할 수 있다. 이러한 숫자 배열은 그 이용자는 로맨스와 드라마를 주로 선호하지만 공상과학도 싫어하지 않는다는 점을 보여준다.

마지막으로 필터링 단계에서는 앞서 콘텐츠 분석 단계와 사용자 프로필 학습 단계를 통해 추출된 정보를 비교해서, 이용자가 이제까지 본 적이 없는 아이템 중 이용자가 선호할 아이템을 매칭(matching)한다. 매칭을 위해서는 해당 아이템 특성과 사용자 프로필의 유사도 점수(similarity score)를 계산해야 한다.⁹⁾ 가령 약간의 로맨스가 포함된 액션 어드벤처 영화인 '분노의 질주' 영화가 (0.0, 0.2, 1.0, 1.0, 0.0)의 변수 값을 갖는다고 해 보자. 이 값은 앞서 사용자 프로필 학습단계에서 도출한 사용자 프로필 변수 값 (0.7, 0.9, 0.1, 0.1, 0.5)과는 큰 차이가 있다. 따라서 양자 간 유사도 점수를 계산해 보면 매우 낮게 나온다(0.21). 이에 비해 로맨스 영화인 '로미오와 줄리엣'이 (1, 1, 0, 0, 0)으로 표현된다고 해 보자. 이는 사용자 프로필과 비슷하므로 유사도 점수가 높게 계산된다(0.90).

일단 유사도 점수가 산정되면, 그 점수에 따라 아이템의 순위(ranking)를 결정한다. 이용자에게는 제한된 개수의 아이템만이 추천될 수 있으므로 여러 아이템 중 순위가 가장 높은 몇 개만 추려내는 것이다. 순위 결정에 있어서는 유사도 점수가 중요한 역할을 하지만, 그 이외에도 다양한 요소들이 고려될 수 있다. 가령 전문적 영화 평론가의 평점이 좋은 영화가 높은 순위를 갖도록 고려할 수 있다. 한편, 고전 영화보다 최신 영화에 더 높은 가중치가 부여될 수도 있다. 요컨대 다양한 요소들에 서로 다른 가중치를 적용하여 점수를 매기고, 그 점수를 정렬하여 제한된 수의 추천 결과를 도출하게 된다.

이상에서 설명한 콘텐츠 분석, 사용자 프로필 학습 및 필터링 단계는 새로운 콘텐츠가 추가되고, 이용자의 피드백(콘텐츠 소비, 제품 구매 등)이 제공됨에 따라 지속적으로 개선된다.

나. 협업 필터링 방식

협업 필터링(Collaborative Filtering) 방식은 콘텐츠 기반 필터링과 달리 다른 이용자와의 유사도를 기반으로 추천을 수행한다. 즉, 피드백, 평점, 선호 및 활동 등에 관한 사용자 행태 정보를 바탕으로 다른 이용자와의 유사성을 계산해 내는 것이다.¹⁰⁾ 예를 들어 다음과 같은 상황이 주어졌다고 하자.

• 사용자 P1은 영화 1, 영화 2, 영화 3을 좋아한다.

• 사용자 P2는 영화 1, 영화 2, 영화 4를 좋아한다.

• 사용자 P3는 영화 1을 좋아한다.

4) Debashis Das, Laxman Sahoo, Sujoy Datta, "A Survey on Recommendation System", International Journal of Computer Applications, vol 160(7), 2017, pp.6-10.

5) Shah Khusro, Zafar Ali and Irfan Ullah, "Recommender Systems: Issues, Challenges, and Research Opportunities", Information Science and Applications (ICISA) 2016, p. 1181.

6) Marwa Hussien Mohamed, Mohamed Helmy Khafagy, Mohamed Hasan Ibrahim, "Recommender systems challenges and solutions survey", International Conference on Innovative Trends in Computer Engineering (ITCE), IEEE, 2019, p. 150.

7) 콘텐츠로부터 추출된 특성은 이처럼 인간이 인식할 수 있는 변수로 표시되는 경우도 있지만, 인간이 인식할 수 없는 그저 일련의 숫자 배열의 형태로 되어 있을 수도 있다. 특히 아래에서 설명하는 바와 같이 딥러닝 기법을 활용하여 임베딩(embedding)을 추출하는 경우가 그러하다. 이러한 경우 그 추천 알고리즘의 작동 원리를 이해하는 것이 쉽지 않다. 더욱이 추출된 변수 값은 수백개에서 수천개에 이를 수 있는데, 이러한 고차원 데이터는 인간이 이해하기 더욱 어렵게 된다. Shuai Zhang et al., "Deep learning based recommender system: A survey and new perspectives", ACM Computing Surveys (CSUR) 52.1, (2019), p. 6.

8) Shuai Zhang et al., 앞의 논문

9) 이러한 숫자 배열간의 유사도는 두 벡터 간의 코사인 유사도(cosine similarity)를 계산하는 방식으로 이루어질 수 있다. 숫자 배열이 서로 유사할수록 코사인 유사도는 높은 값을 갖는다. Shuai Zhang et al., 앞의 논문, p. 151.

10) Shuai Zhang et al., 앞의 논문, p. 151.

P3는 영화 2, 영화 3 또는 영화 4를 좋아할 것인가? 사용자 P1, P2는 모두 영화 1을 좋아한다는 점에서 사용자 P3와 유사하다. 그런데 사용자 P1, 사용자 P2가 모두 영화 2를 좋아했다. 따라서 사용자 P3도 영화 2를 좋아할 가능성이 높다고 예측할 수 있다. 사용자 P3가 영화 3이나 영화 4를 좋아할 수도 있겠지만 영화 2만큼 가능성이 높지는 않을 것이다. 이렇듯 협업 필터링 방식은 다수 사용자 피드백 등에 기반하여 해당 사용자와 유사한 사용자들을 찾고, 그 유사한 사용자들의 정보를 이용해서 추천을 수행한다.

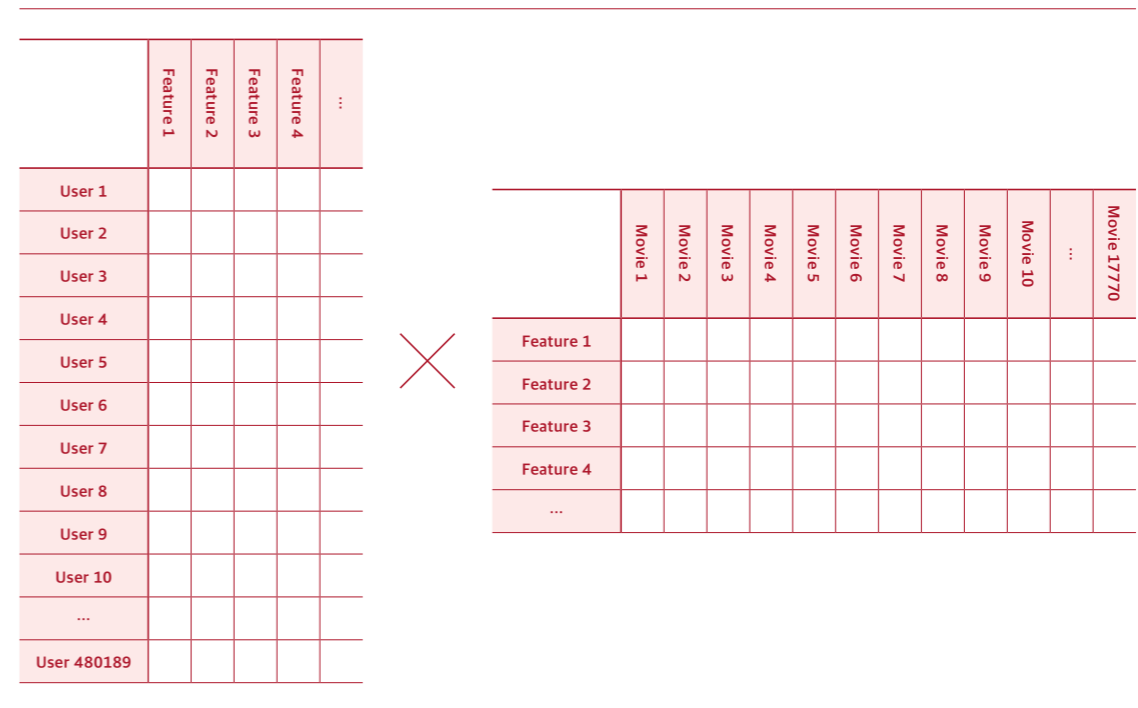
협업 필터링 방식의 발전에 있어 가장 잘 알려진 사례는 넷플릭스의 영화 추천 알고리즘 경진대회이다. 넷플릭스 사는 2006년 10월 자사의 영화 추천 알고리즘을 10% 개선할 경우 100만 달러의 상금을 지급하는 “Netflix Prize” 대회를 개최했다.¹¹⁾ 넷플릭스 사는 이 경진대회를 위해 사용자 480,189명의 영화 17,770 건에 대한 평점(1~5점) 데이터 약 1억 건을 공개했다. 알고리즘의 성능에 대한 평가는 공개되지 않은 테스트 데이터셋 140여만 건을 통해 이루어졌다. Netflix Prize 경진대회의 목적은 테스트 데이터셋에 포함된 실제 사용자 평가 결과와의 오차를 최소화하는 알고리즘을 개발하는 것이었다.

Netflix Prize 경진대회는 대단한 성공을 거두었다. 전세계 186개국에서 41,305팀의 51,051명이 참여하였다. 넷플릭스 사는 당초 대회 기간을 5년으로 정해 두었는데, 3년도 채 되지 않아 목표한 10%의 성능 향상을 이루었다. 이미 넷플릭스 소속 연구자들이 영화 추천 알고리즘의 개선을 위해 오랫동안 상당한 자원을 투입해 왔던 점을 고려하면 놀라운 결과라 할 수 있다.

	Movie 1	Movie 2	Movie 3	Movie 4	Movie 5	Movie 6	Movie 7	Movie 8	Movie 9	Movie 10	...	Movie 17770
User 1			1		2							3
User 2		2		3	3			4				
User 3							5	3		4		
User 4	2				3			2				2
User 5		4				5			3			4
User 6			2									
User 7			2					4	2	3		
User 8	3	4				4						
User 9										3		
User 10			1		2							2
...												
User 480189		4			3			3				

경진대회를 통해 제안된 다양한 알고리즘이 있지만, 이 문제의 가장 일반적인 해결 방법은 행렬 인수분해(matrix factorization)를 이용하는 것이다. 그 개요는 다음과 같다.¹²⁾ 영화 추천 문제는 아래와 같이 방대한 표(행렬)의 빈 칸을 채워 넣는 문제로 생각해 볼 수 있다. 즉, 사용자 480,189명의 영화 17,770건에 대한 영화 평점을 한 개의 표로 그리면 모두 85억 여개(480,189×17,770)의 칸을 가진 거대한 행렬로 표시된다. 넷플릭스 사는 전체 85억 여개의 칸 중에서 약 1억 건에 대한 정보를 학습 데이터로 제공했다. 나머지 84억 개 빈칸을 채워 넣는 것이 풀어야 할 과제이다.

이 문제를 풀기 위해 행렬 인수분해를 이용한다. 이를 간단히 설명하면 더 작은 표(행렬)로 나누어 계산하는 것이다. 흔히 ‘차원 축소(dimension reduction)’라 부르는 전략이다. 만약 우리가 개별 이용자를 100개의 특성으로 표현하고, 각 영화들도 100개의 특성을 표현한다고 해 보자. 그러면 (이용자 48만명)×(영화 17,770건)로 구성된 위 거대한 행렬은 (이용자 48만명)×(이용자 특성 100개) 행렬과 (영화 특성 100개)×(영화 17,770건)의 행렬의 곱으로 인수 분해하여 근사(approximation)될 수 있다.¹³⁾



위와 같이 100 차원으로 줄이는 전략을 적용하면 우리가 찾아야 하는 미지수는 약 5천만 개로 줄어든다(480,189×100 + 17,770×100 = 49,795,900). 5천만이 작은 숫자는 아니지만, 적어도 84억 개의 빈 칸을 채우는 문제보다는 간단하다. 미지수를 찾는 일은 머신러닝 알고리즘을 적용하여 넷플릭스가 제공한 1억 건의 학습 데이터를 통해

11) Kate Greene, "The \$1 Million Netflix Challenge", MIT Technology Review, 2006. 10. 6.
 12) Jeff Howbert, "The Netflix Prize Contest", 2014. 2. 4. http://courses.washington.edu/css581/lecture_slides/09a_Netflix_Prize.pdf
 13) (m×n) 크기의 행렬은 (m×k) 크기 행렬과 (k×n) 크기 행렬의 곱으로 근사될 수 있다.

이루어진다.¹⁴⁾ 이러한 과정이 어려워 보이지만, 그 원리만 두고 보면 아주 복잡하지는 않다. 결국 학습 데이터를 통해 이용자와 영화를 표현하는 특성 값을 도출하고, 이를 통해 이용자가 평가하지 않은 영화의 평점을 예측하는 것이다.

협업 필터링의 장점은 이용자의 나이, 성별, 지역 등 그 인구통계적 데이터가 전혀 없더라도 행동 패턴에만 기반하여 추천 아이템을 찾아낼 수 있다는 것이다.¹⁵⁾ 이는 다수 이용자 간에 피드백의 유사도에 기반하여 추천을 수행하기 때문이다. 또다른 장점은 이용자 프로필에서 벗어나는 아이템을 추천할 가능성이 생긴다는 점이다.¹⁶⁾ 예를 들어 로맨틱 코미디 영화들에 높은 평점을 매긴 이용자들이 공통적으로 선호하는 스릴러 영화가 있다고 해 보자. 콘텐츠 기반 필터링을 통해 얻어진 이용자 프로필이나 영화 특성 정보만으로는 해당 스릴러 영화를 추천하기 어렵다. 하지만 협업 필터링을 통해서 이러한 가능성을 찾아낼 수 있게 된다.

하지만 협업 필터링 시스템의 단점도 있다. 우선 사용자들의 기존 평점 데이터가 충분히 축적되어 있어야 하고, 그 데이터가 이용자 선호를 정확하게 반영할 수 있어야 한다. 특히 신규 이용자와 같이 추천에 활용할 데이터가 존재하지 않는 경우가 문제된다. 이를 보통 '콜드 스타트 문제(cold start problem)'라고 한다.¹⁷⁾ 간단한 해법은 이용자가 서비스를 이용하기에 앞서 어떠한 아이템을 선호하는지 확인하는 것이다. 영화나 음악 스트리밍 앱을 처음 설치하면 선호하는 음악이나 영화를 선택하도록 요구하는 것이 이 때문이다. 또한 전체 이용자의 특성을 평균 내어 신규 이용자에게 적용하거나, 만약 이용자의 성별, 연령, 지역 등을 안다면 그와 유사한 집단의 평균을 적용할 수도 있다.

또한, 이른바 '회색 양(grey sheep)' 문제도 있다. 이는 어떤 이용자가 독특한 선호 체계를 갖고 있어 다른 어떠한 이용자 그룹과도 매칭되지 않는 경우를 일컫는다.¹⁸⁾ 협업 필터링 방식으로는 이러한 '회색 양' 이용자에게 적합한 추천을 하기 어렵게 된다. 이에 비해 콘텐츠 기반 필터링을 통해서 이용자 프로필과 아이템 간의 유사성에 따라 아이템을 추천하는 경우 이러한 이용자에게도 적절한 추천을 할 수 있다. 그래서 많은 경우 콘텐츠 필터링과 협업 필터링 중 하나만을 사용하기보다 양자를 적절히 조합한 하이브리드(hybrid)형 시스템을 사용하게 된다.

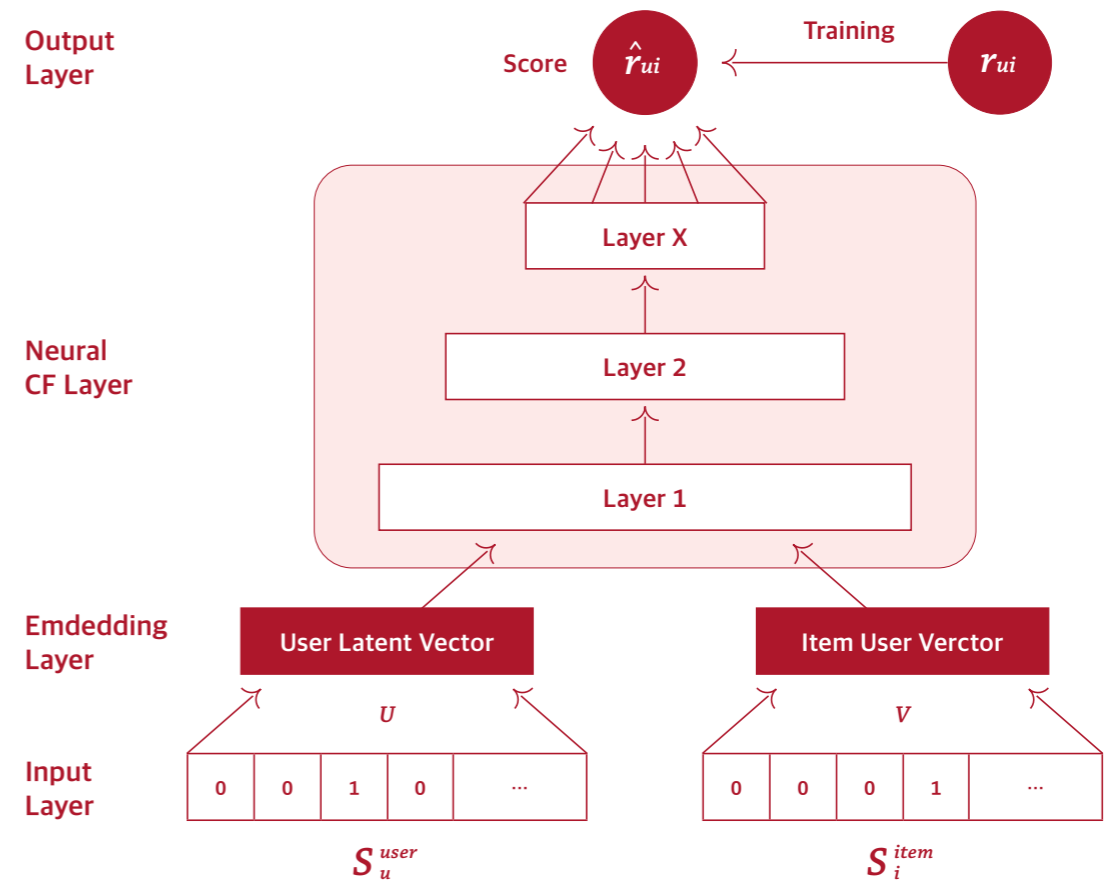
다. 딥러닝 기반 추천 시스템

다른 컴퓨터 공학 분야와 마찬가지로 추천 시스템에 있어서도 딥러닝(deep learning) 기법을 활용하여 성능을 개선하려는 시도가 늘어나고 있다. 유튜브의 비디오 추천, 앱 스토어에서의 앱 추천, 뉴스 추천 등에 있어서 딥러닝 모델을 활용함으로써 전통

적 방식에 비해 상당한 성능 개선을 보이고 있다.¹⁹⁾

딥러닝 방식은 여러 계층으로 구성된 인공신경망을 이용하는 머신러닝 기법을 의미한다. 심층 인공신경망에 포함된 다수의 뉴런들은 입력된 데이터를 추상화하여 표현함으로써 주어진 과제를 더욱 잘 해결할 수 있게 된다. 이처럼 심층 인공신경망을 통해 추상화된 표현을 'deep representation'이라고 부르는데, 이는 이용자와 콘텐츠 특성을 잘 나타내는 것으로 알려져 있다.

예를 들어 협업 필터링을 위한 단순한 인공신경망 구조는 아래 그림과 같다.²⁰⁾ 이용자 정보와 아이템 정보가 인공신경망에 입력되면, 여러 계층의 인공신경망을 통해 정보의 변환 및 추상화가 이루어지게 된다. 최종적으로 마지막 계층에 포함된 인공 뉴런의 출력 값을 이용해서 이용자의 해당 아이템 선호도 점수가 산출되어 나온다. 인공 뉴런 간의 연결 가중치는 이용자의 기존 아이템에 대한 평가 결과를 이용하여 학습된다.



14) 이용자 및 영화 데이터를 몇 차원의 특성으로 축소할 것인지는 반복 실험을 통해 경험적으로 측정할 수 있다. 차원이 커질수록 정확도가 올라갈 것이지만, 그에 따라 필요한 메모리가 증가하고, 계산 시간이 오래 걸리게 된다. Netflix Prize 경진대회 우승자는 50차원부터 1,500차원까지 다양한 차원으로 테스트한 결과를 제시하고 있다. Yehuda Koren, Robert Bell, and Chris Volinsky, "Matrix factorization techniques for recommender systems", Computer 42.8 (2009).
 15) Yehuda Koren, Robert Bell, and Chris Volinsky, 앞의 논문, p. 152.
 16) Yehuda Koren, Robert Bell, and Chris Volinsky, 앞의 논문.
 17) Shah Khusro, Zafar Ali and Irfan Ullah, "Recommender Systems: Issues, Challenges, and Research Opportunities", Information Science and Applications (ICISA) 2016, p. 1182.
 18) Shah Khusro, Zafar Ali and Irfan Ullah, 앞의 논문, p. 1184.

19) Shuai Zhang et al., 앞의 논문.
 20) Shuai Zhang et al., 앞의 논문, p. 10.

현재 사용되는 딥러닝 인공지능망은 이보다 더 복잡한 구조를 띄는 경우가 대부분이다. 대표적인 아키텍처로는 합성곱 신경망(Convolutional Neural Network, CNN), 순환신경망(Recurrent Neural Network, RNN), 오토인코더(Autoencoder, AE), 적대적 네트워크(Adversarial Network), 어텐션 모델(Attention Model), 심층 강화학습(Deep Reinforcement Learning, DRL) 등이 있다.²¹⁾

딥러닝 방식은 전통적 기법에 비하여 여러 장점이 있다. 앞서 설명한 행렬 인수분해 등 고전적 방식은 본질적으로 선형(linear) 모델로서, 입력 변수 상호 간의 복잡한 상호작용을 모델링하는 데 한계가 있다. 이에 비해 인공지능망은 비선형 활성화함수가 사용됨으로써 더 복잡한 패턴을 포착할 수 있게 된다.

또한 딥러닝 방식은 데이터의 추상화된 표현을 학습하는데 있어 우수한 성능을 보인다. 이로써 인간 작업자가 일일이 데이터의 특성을 선택할 필요가 없게 되고,²²⁾ 텍스트, 이미지, 음성, 영상이 복합적으로 포함된 데이터에 대한 추천 시스템을 구축할 수 있게 된다. 뿐만 아니라 딥러닝 아키텍처 중에는 순서를 가지고 입력되는 시퀀스 데이터를 처리하는데 뛰어난 경우도 있으며, 특히 CNN과 RNN 모델이 이러한 작업들에 있어 주요하게 활용되고 있다(기계번역, 음성인식, 챗봇 등). 이러한 모델들을 활용하면 이용자의 이전 행동 시퀀스를 이용하여 다음 행동을 더욱 정교하고 정확하게 예측할 수 있다(이용자가 장바구니에 포함시킬 아이템 추천, 다음에 볼 영상 추천 등).

딥러닝 방식의 현실적인 장점은 딥러닝 개발을 위한 도구가 널리 보급되어 있어 현업의 개발자들이 적용하기에 상대적으로 용이하다는 점이다. 텐서플로우(Tensorflow), 파이토치(PyTorch), 카페(Caffe), 티아노(Theano) 등 다양한 딥러닝 개발 프레임워크가 존재하고, 개발자 커뮤니티가 활성화되어 있으며, 성능 개선도 빠르게 이루어지고 있다. 딥러닝 개발 프레임워크를 이용하여 강력한 딥러닝 모델을 비교적 빠르게 구축하고, 상당한 성능을 달성할 수 있다.

한편, 딥러닝 방식의 주요한 문제점으로는 낮은 설명가능성이 지적된다. 방대한 인공지능망 네트워크의 수많은 인공 뉴런들이 어떤 경우에 활성화되는지를 해석하는 것은 쉽지 않은 과제이다. 그 결과 인공지능망은 블랙박스과 마찬가지로 여겨지는 경향이 있다. 그러나 최근에 제시되고 있는 여러 모델(특히 어텐션 모델 등)은 어느 정도 해석 가능하다. 인공지능망을 이용하여 인간이 이해할 수 있는 설명을 생성해 내는 기법도 활발하게 연구되고 있다.

또한 딥러닝 방식을 활용하려면 방대한 데이터가 필요하다는 문제도 제기된다. 인공지능망의 학습은 데이터를 이용하여 인공 뉴런 간의 연결 강도를 조정하는 방식으로 이루어진다. 인공지능망에 포함된 인공 뉴런의 숫자가 비약적으로 증가하면서, 학습에 필요한 데이터의 용량 역시 증가한다. 하지만 추천 시스템의 경우 다른 딥러닝 시스템에 비해 상당한 양의 데이터를 확보하는 것이 비교적 용이하다. 추천 시스템에 있어 수억 건 이상의 데이터셋을 활용하는 경우는 쉽게 볼 수 있다. 그 결과 추천 시스템에 있어 딥러닝 기법을 적용한 연구가 매우 활발하게 진행되고 있고, 상용 시스템에도 널리 활용되는 추세이다.

21) Shuai Zhang et al., 앞의 논문, pp. 4-5.

22) 인간에 의한 특성 선택 작업을 '특성 엔지니어링(feature engineering)'이라 한다

III. 유튜브 추천 시스템

유튜브는 이제 그저 음악을 듣고 동영상을 보면서 여가 시간을 보내는 사이트가 아니라, 뉴스, 교육, 제품 홍보, 정치적 의사표현 및 여론 형성의 중요한 공간이 되었다. 그와 함께 유튜브의 추천 시스템에 대한 관심도 크게 증가하고 있다. 2018년 1월 보도에 따르면 알고리즘에 의해 추천된 영상이 유튜브 전체 시청 시간의 70%를 차지한다고도 한다.²³⁾ 하지만 유튜브 추천 시스템의 정확한 작동 방식은 공개되어 있지 않다.

다만, 유튜브 웹사이트에는 자신의 추천 시스템에 있어 어떠한 요소가 고려되는지 간략히 소개되어 있기는 하다.²⁴⁾ 유튜브는 이용자 시청 및 검색 이력(사용 설정된 경우), 구독한 채널, 다른 이용자가 동영상을 끝까지 시청했는지(동영상의 품질이 우수하거나 재미있다는 신호) 혹은 동영상을 클릭하여 재생한 후 바로 다른 콘텐츠를 클릭했는지 여부, 무작위 설문조사 결과 등을 고려한다고 한다. 또한 현지 관련 뉴스 등을 표시하는데 있어 국가 및 시간과 같은 이용자 상황도 고려한다고 설명되어 있다. 나아가 유튜브는 "사용자와 관련성 있는 동영상을 추천하기 위해 시스템을 테스트, 학습, 조정을 계속해 나가고 있"고, "시청자의 직접적인 의견을 바탕으로 모든 사용자를 위해 시스템을 조정하고 개선하고 있"다고 하고 있다. 하지만 위와 같이 개별 요소들이 어떠한 방식으로 고려되는지, 각 요소들 간의 경중(輕重)은 어떻게 되는지는 알 수 없다.

한편, 유튜브 추천 시스템의 기술적 작동 원리에 관하여는 구글 연구자들이 2016년 및 2019년 발표한 논문이 있다.²⁵⁾ 유튜브가 실제 위 논문을 통해 공개한 방식을 사용하고 있는지는 확인할 수 없으나, 일응 이를 근거로 유튜브 추천 시스템의 작동 원리를 추측해 볼 수 있다.

2016년 논문 "유튜브 추천을 위한 심층 인공지능망(Deep Neural Networks for YouTube Recommendations)"은 딥러닝 기법을 이용한 유튜브 동영상 추천 방식을 소개하고 있다. 이 논문은 유튜브와 같이 방대한 동영상과 이용자가 존재하고, 새로운 동영상이 지속적으로 업로드되고 있으며, 이용자 선호를 정확하게 측정하기 어려운 환경 하에서는 행렬 인수분해와 같은 고전적 추천 시스템을 적용하기 어렵다고 지적한다. 유튜브는 이 문제를 해결하기 위해 약 10억 개의 파라미터를 갖춘 인공지능망을 구축하고, 수천 억 건의 학습 데이터로 학습시켰다고 한다.

위 논문이 소개하는 인공지능망은 크게 (1) 후보 생성(candidate generation) 단계와 (2) 랭킹(ranking) 단계로 구성되어 있다. (1) 후보 생성 단계는 대략적인 이용자 정보(시청한 영상, 검색어, 인구통계 정보)을 입력받아 수백만 건의 동영상 중에서 수백 건의 후보 동영상을 추출해 낸다. 여기서는 협업 필터링 방식을 이용하여 전반적으로 이용자와 관련성이 있는 동영상을 찾아내는 것을 목적으로 한다. 이 단계에서는 이용자가 선호하지 않

23) Joan E. Solsman, "YouTube's AI is the puppet master over most of what you watch", cnet, 2018. 1. 10. <https://www.cnet.com/news/youtube-cs-2018-neal-mohan/>

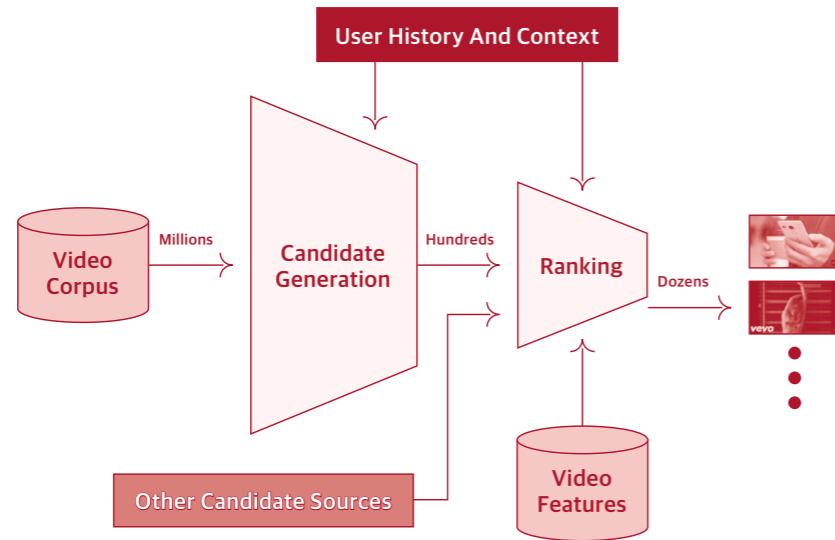
24) https://www.youtube.com/intl/ALL_kr/howyoutubeworks/product-features/recommendations/ 한국어 웹사이트에는 'recommendation'이 '맞춤 동영상'이라고 번역되어 있다.

25) Paul Covington et al., "Deep neural networks for YouTube recommendations", Proceedings of the 10th ACM conference on recommender systems, 2016. 및 Zhe Zhao et al., "Recommending what video to watch next: a multitask ranking system", Proceedings of the 13th ACM Conference on Recommender Systems, 2019.

을 동영상의 잘못 후보로 포함되는 오류(False Positive)를 줄이기 위해 최적화된다.²⁶⁾

다음으로 (2) 랭킹 단계에서는 이용자와 동영상에 관한 정보를 활용하여 후보 동영상마다 점수를 부여하고, 그 점수에 따라 정렬하여 소수의 가장 적합한(best) 추천 영상을 도출한다. 이때 소수의 후보 동영상에 대해서만 점수를 부여하면 되므로, 더욱 풍부하고 세밀한 이용자 정보를 이용하게 된다. 이 단계에서는 이용자가 선호할 동영상이 낮은 점수를 받는 오류(False Negative)를 줄이기 위해 최적화된다.²⁷⁾

위 논문에서는 이 과정을 아래와 같이 도식화하고 있다.



후보 생성 단계에서 흥미로운 점은 동영상이 언제 업로드 되었는지를 중요한 특성으로 고려한다는 점이다. 이용자들은 최신 콘텐츠를 선호하는 경향이 있다. 하지만 종래의 추천 시스템들은 이미 유명해져서 여러 이용자로부터 추천을 받은 동영상을 더 많이 추천하는 한계가 있다. 구글 팀은 이 문제를 해결하기 위해 후보 생성 신경망에서 업로드된 후 시간이 얼마나 경과되었는지('example age')를 입력 받아 활용하고 있다.

한편, 랭킹 단계에서 중요한 것은 순위가 어떤 목표를 최적화하기 위해 결정되는가 하는 문제이다(이를 목적함수(objective function)라 한다). 유튜브는 실시간 이용자 경험 테스트를 통해 순위 결정을 위한 목적함수를 지속적으로 조정한다고 하는데, 여러 변수들 중에서 주로 기대 시청 시간의 영향이 크다고 설명한다. 이때 이용자가 영상을 클릭하는 비율을 주된 기준으로 할 경우에는 낚시성(click-bait) 동영상이 장려되는 경향이 있으므로, 시청 시간이 더 중요한 요소로 고려된다고 한다.

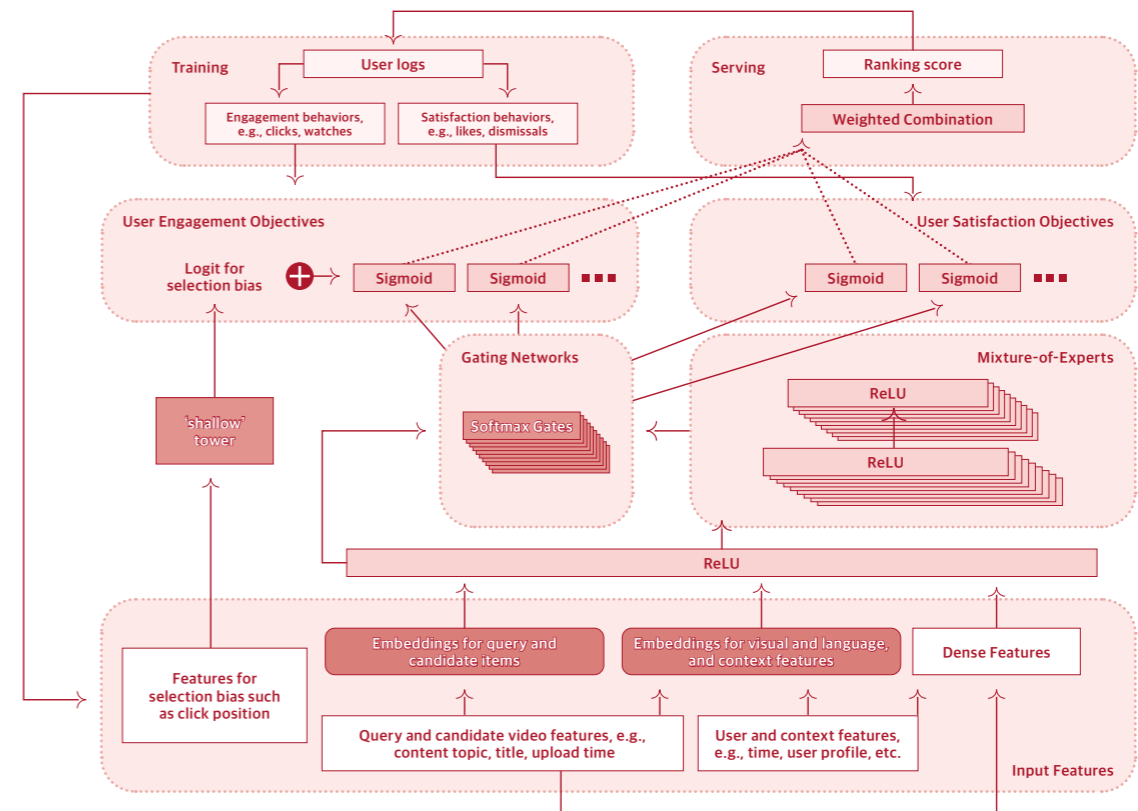
다음으로, 2019년 논문인 “다음에 볼 동영상 추천하기: 다중작업 순위 시스템 (Recommending What Video to Watch Next: A Multitask Ranking System)”은 개선

26) 통계적 용어로 표현하면 정밀도(precision)를 높이기 위해 최적화된다.

27) 통계적 용어로 표현하면 재현율(recall)을 높이기 위해 최적화된다.

된 시스템을 제시하고 있다. 2019년 논문에서 소개된 모델도 2016년 논문과 마찬가지로 후보 생성단계와 랭킹 단계로 구분하고 있다. 그런데 2019년 논문은 그 제목에서 알 수 있는 바와 같이 그 중 랭킹 모델에 보다 초점을 맞추고 있다. 주요한 문제의식은 유튜브와 같은 대규모 비디오 추천 시스템은 순위 결정시 여러 요소를 고려해야 한다는 점이다. 개별 이용자의 클릭이나 시청 시간만이 아니라 ‘좋아요’ 표시, 친구 공유 여부 등과 같은 다양한 요소를 반영해야 한다. 이처럼 다양한 목적을 함께 충족시킬 수 있는 순위 결정 시스템을 구축하는 것이 이 논문의 주된 목표이다.

아래 그림은 2019년 논문이 제시하는 시스템의 개략적 구조이다. 2016년 모델에 비하여 훨씬 복잡한 구조를 취하고 있다는 점을 알 수 있다. 여기서 눈에 띄는 부분은 모델의 목적이 크게 이용자 행동 목적들(User Engagement Objectives)과 이용자 만족 목적들(User Satisfaction Objectives)로 구분되어 있다는 점이다. 이용자 행동 목적에는 이용자 클릭 및 시청시간 등이 포함되고, 이용자 만족 목적에는 ‘좋아요’ 또는 별점 등이 포함된다. 이 논문에서는 위 목적만을 예시로 언급하고 있으나, 위 모델에는 그 이외에도 여러 목적이 추가될 수 있다. 최종적으로는 여러 목적들에 서로 다른 가중치를 적용하여 순위 점수가 도출되는데, 각 목적마다 적용되는 가중치는 “수작업을 통해 조율된다(manually tuned)”고 설명하고 있다.



그런데, 단일한 인공지능경망 모델만으로는 이처럼 다양한 목적에 대해 동시에 최적화하기 쉽지 않다. 그래서 이 논문은 특정 작업(클릭 예측, 시청 시간 예측, 좋아요 예측 등)에 특화된 인공지능경망을 여러 개를 병렬적으로 연결하는 구조를 제안한다. 이처럼 특화된 인공지능경망을 전문가(Expert)라고 부르고, 개별 작업마다 어떠한 전문가 신경망을 사용할 것인지를 판단하는 부분 역시 별도의 인공지능경망을 구축하여 학습시킨다(Gating Network).²⁸⁾ 이러한 과정을 통해 여러 목적을 동시에 충족시킬 수 있는 시스템을 구축할 수 있다는 것이 이 논문의 핵심 요지이다.

나아가, 이 논문은 추가적으로 또다른 문제를 제기한다. 추천 시스템과 사용자 행위 간에 피드백 루프가 존재한다. 즉, 사용자에게 어떤 영상이 추천되면 그 영상을 시청하게 될 가능성이 높아진다. 따라서 학습 데이터 내에는 선택 편향(selection bias)이 존재하게 되고, 그 결과 이용자의 진정한 선호를 파악하기 어렵게 된다. 이 문제를 극복하기 위해서는 모델을 학습하는 단계에서 상위에 랭크되었던 영상에 대해서는 일정한 벌칙을 부가하여 가중치를 의도적으로 낮춘다고 한다.

요컨대 2019년 논문이 시사하는 바는 유튜브 추천 시스템은 여러 목적에 대해 함께 최적화하도록 설계되어 있다는 점이다. 앞서 소개한 넷플릭스의 협업 필터링 방식 추천 시스템은 이용자의 별점을 근거로 하여 이용자 선호를 예측하고자 하였다는 점에서 단일한 목적에 대해 최적화된 것이다. 그러나 유튜브 추천 시스템은 이용자 선호를 파악할 수 있는 다양한 지표(클릭, 시청 시간, 좋아요, 댓글 등)를 목적으로 사용함으로써 추천 시스템의 품질을 개선하고자 하고 있다.

IV. 추천 시스템과 인간 행동

가. 콘텐츠 소비자의 측면

추천 시스템 설계자의 입장에서 보면, 주어진 이용자 행태정보만으로 이용자의 선호를 정확하게 파악하기 어렵다는 문제가 존재한다. 앞서 협업 필터링의 예로 든 넷플릭스의 사례를 들어 보자. 넷플릭스는 2017년 4월 종전의 5개 별점으로 된 평가 시스템을 중단하고, '좋아요(thumbs up)'와 '싫어요(thumbs down)' 체계로 개편하였다. 넷플릭스 임원의 언론 인터뷰에 따르면, 개편 이유는 넷플릭스의 별점 시스템으로는 이용자의 선호를 적절하게 반영하기 어렵기 때문이었다고 한다. 영화에 대한 별점은 영화에 대한 이용자 선호를 충분히 나타내지 못한다. 다른 예로, 중국 음식을 좋아하는 이용자가 어떤 중식당에 낮은 평점을 주었다고 해서 중국 음식 전체를 싫어하게 된 것은 아니다. 마찬가지로 만약 넷플릭스 이용자가 어떤 액션 영화에 낮은 점수를 매겼다고 하더라도, 이는 그 영화가 맘에 들지 않기 때문이지, 액션 영화 장르 전체를 싫어하게 된 것은 아닐 가능성이 크다.

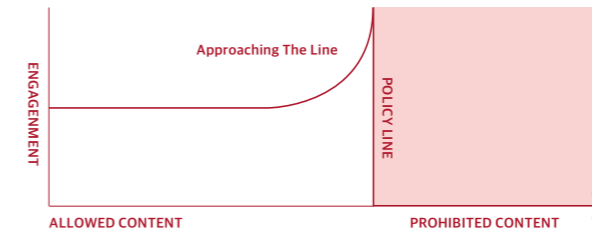
더욱이 넷플릭스 이용자는 별점을 매길 때 자신의 경험을 기준으로 하는 것이 아니라, 마치 영화 평론가가 된 것처럼 행동하는 경향도 있었다고 한다. 즉, 자신이 어떤 영화를 재미있게 보았고 비슷한 영화를 다시 볼 의향이 있더라도, 영화 자체의 완성도가 낮다면 평점을 낮게 주는 경향도 있었던 것이다. 이처럼 우수한 추천 시스템을 구축하기 위해서는 단지 기술적 측면만을 고려하는 것이 아니라 이용자와 서비스 간의 상호 작용에 대

한 주의 깊은 관찰과 고려가 필요하기도 하다.

만약 이용자의 진정한 선호를 파악할 수 있게 된다고 하더라도 문제는 남는다. 우리는 자극적이고 선정적인 콘텐츠에 이끌린다. 그래서 우리의 선호를 정확하게 파악해서 자극적이고 선정적인 콘텐츠를 주로 추천하는 모델을 개발한다면, 이를 좋은 추천 시스템이라고 평가하기 어렵다.

이와 관련하여 페이스북의 마크 주커버그는 아래 그림과 같이 게시물에 허용 경계선에 점차 접근할수록 이용자의 활동이 증가하는 현상을 지적하였다.²⁹⁾

Natural Engagement Pattern



위 그림의 경계선은 법이나 플랫폼 정책으로 정한 콘텐츠의 한계이다. 포르노그라피나 가짜 뉴스 등은 경계선 밖에 존재한다. 하지만 아슬아슬하게 경계 근처에서 위치에서 금지되지 않는 사례도 있다. 포르노그라피에 해당하지는 않지만 선정적인 영상, 또는 가짜 뉴스까지는 아니더라도 음모론을 제기하는 글 등이 그 예이다. 이를 경계 콘텐츠(borderline content)라 한다. 케이블 TV나 타블로이드 신문 등과 같이 경계 콘텐츠 선호 현상을 악용하는 사례는 예전부터 존재해 왔다.

만약 추천 시스템이 이용자 선호를 매우 정확하게 고려한다면 오히려 경계 콘텐츠가 더욱 자주 추천되는 현상이 나타날 수 있다. 이러한 문제는 쾌락의 '질'의 구별하지 못한다는 공리주의에 대한 고전적 비판을 떠올리게 한다. 일반적 추천 시스템은 고급 쾌락과 저급 쾌락을 구분하기 어렵다. 콘텐츠의 '질'에 대한 고려가 이루어지지 않는 이상 경계 콘텐츠가 더욱 자주 추천되는 현상은 피하기 어렵다.

이에 대해 주커버그는 페이스북이 허용 경계선에 가까운 콘텐츠에 대해서는 부정적 가중치를 적용시킴으로써 대응하고자 한다고 설명하였다.³⁰⁾ 이는 결국 콘텐츠의 '질'을 평가하여 허용 경계선에 얼마나 가까운지 평가한다는 의미이다. 유튜브 역시 이용자가 신뢰할 수 있는 권위있는 콘텐츠(authoritative content)의 추천을 늘리고, 경계 콘텐츠 및 허위 정보를 줄이고자 노력하고 있다고 설명해 왔다.³¹⁾ 앞서 소개한 유튜브 추천 시스템이 다양한 목적을 고려할 수 있도록 설계된 점을 고려하면 권위있는 콘텐츠에 높은 순위가 부여되고, 경계 콘텐츠나 허위 정보에는 낮은 순위가 부여될 수 있도록 랭킹 시스템을 구성할 수 있을 것이다.

28) 이 논문은 위와 같이 다중 작업을 수행하기 위해 여러 전문가들을 선택하는 구조를 Multi-gate Mixture-of-Experts (MMoE)라 한다.

29) Mark Zuckerberg, "A Blueprint for Content Governance and Enforcement". <https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/>

30) Mark Zuckerberg, 앞의 글.

31) <https://blog.youtube/inside-youtube/the-four-rs-of-responsibility-raise-and-reduce>

추천 시스템의 또다른 문제로 지적되는 것은 이른바 ‘필터 버블(filter bubble)’ 또는 ‘에코 챔버(echo chamber)’ 현상이다. 앞서 본 바와 같이 추천 시스템은 콘텐츠 기반 필터링이든 협업 필터링이든 이용자의 기존 피드백 이력을 근거로 한다. 그 결과 기존 피드백에 부합하는 콘텐츠만 추천되고, 이용자가 미처 알지 못한 흥미로운 콘텐츠는 추천되지 못하고, 이로써 이용자에게 제시되는 콘텐츠는 다양성이 줄어들 가능성이 있다. 특히 정치적 의견 교환이 이루어지는 공간에서 이러한 현상이 지속되면, 이용자는 다양한 관점의 견해를 접하지 못하고 개인의 선입견이 지속적으로 강화될 수도 있다. 이는 사회 통합을 해하고 더욱 분화된 사회를 조장할 우려가 있다는 것이 이러한 문제제기의 요지이다.

다만, 이러한 문제제기에 대한 실증적 증거는 그다지 명확하지 않다. 뉴스 추천 시스템에 있어서는 이러한 문제가 아직 나타나고 있지 않은 것으로 보인다.³²⁾ 소셜 미디어 플랫폼에 있어서는 추천 시스템으로 인해 이용자의 분화가 이루어지는 효과가 있으나 그 정도는 상대적으로 작은 수준이다.³³⁾ 그럼에도 불구하고 이러한 잠재적 우려에 대한 대응이 아직 이루어질 필요가 없다고 단언하기 어렵다.

컴퓨터 공학에서는 이러한 문제를 탐색(exploration)과 활용(exploitation) 간의 상충관계(trade-off)라 부른다. 점심 식사를 하러 갈 식당을 고르는 문제가 전형적 예이다. 새로운 식당을 시도해 보는 것이 탐색이고, 기존 경험에 따라 가장 최선의 식당을 가는 것이 활용이다. 탐색 시도가 적으면 맛집을 놓칠 수도 있다. 그렇다고 탐색만 하다 보면 전체 효용이 줄어든다. 효용을 최대화하기 위해서는 양자 간의 적절한 타협점을 찾는 전략이 필요하다. 추천 시스템에서도 비슷한 상충관계가 고려된다. 추천 시스템의 주된 목표는 이용자의 기존 이력에 부합하는 콘텐츠를 추천하는 것(즉, 이용자 선호 정보의 활용)이지만, 이용자 선호를 보다 정확하게 파악하기 위해서는 우연성의 요소를 추가하여 다양한 콘텐츠를 추천할 필요도 있다(즉, 이용자 선호의 탐색). 따라서 추천 시스템 설계자들은 이러한 문제를 고려하여 추천 알고리즘을 작성하게 된다.

나. 콘텐츠 생산자의 측면

추천 시스템은 콘텐츠 생산자가 어떠한 콘텐츠를 만들어 낼 것인지에도 큰 영향을 끼친다. 추천 시스템의 결과는 콘텐츠 생산자의 수익과 직결되어 있다. 그래서 생산자들은 추천 시스템의 작동 원리를 보다 정확하게 알아 내고자 애쓴다. 하지만 플랫폼이 추천 시스템의 정확한 내용을 공개하는 경우는 거의 없다. 무엇보다도 생산자들이 콘텐츠의 품질을 높이기 위해 경쟁하는 것이 아니라, 추천 시스템을 악용하여(낚시글 등) 추천 가능성을 높이기 위한 조작(manipulation) 시도를 할 위험이 크기 때문이다.

구글 창립자들이 초기 구글 검색엔진 알고리즘을 공개한 결과 이러한 검색 순위 조작 시도가 다수 이루어진 바 있다. 구글의 ‘PageRank’ 알고리즘은 다른 사이트로부터의 링크의 숫자와 품질을 근거로 검색 순위를 매긴다. 이러한 점이 알려지자, 여러 웹사이트 운영자들은 다양한 방식으로 구글 알고리즘을 속이고자 시도하였다. 예컨대 유상으로 자신의 사이트에 대한 링크를 구매하거나, 상호 간에 링크를 해 주는 링크 교환 등의 현상이 발생하였다. 구글은 이러한 순위 조작 행위를 금지하고 있으나, 구글의 내부 통계에 의하면 여전히 수십만 개의 웹사이트들이 이러한 행동을 하고 있다고 한다.³⁴⁾ 구글은 ‘Pag-

eRank’ 알고리즘에 대한 부정적 경험 이후 새로운 검색 순위 결정 방식에 대해서는 구체적인 세부 사항을 공개하지 않고 있다.

이처럼 추천 시스템의 작동 방식이 구체적으로 공개되지 않고 있는 상황에서 콘텐츠 생산자들은 그 작동 방식을 어림짐작으로 추측하여 행동하게 된다. 유튜브는 추천 시스템이 동영상의 품질과 재미에 따라 추천이 이루어진다고 설명하여도 콘텐츠 생산자들은 조금이라도 경쟁에서 앞설 방법을 찾게 된다. 이미 대규모 사업화된 유튜브들은 업로드할 영상의 제목과 썸네일(thumbnail)을 결정하기 위해 여러 후보를 만들어 놓고, 미리 섭외한 이용자를 대상으로 소규모 실험을 해서 가장 클릭 확률이 높은 것으로 정하고 있다고 한다.³⁵⁾ 종래 광고 대행사에서 중요한 광고 출시 전 여러 개의 후보 영상을 만들고 이용자 사전 평가에 따라 최종본을 선정하던 것과 비슷하다.

현재와 같이 불확실성이 큰 상황 하에서는 마치 심리학자 스키너의 실험에서 비둘기들이 보인 미신적 행동(superstitious behavior)과 비슷한 행태가 초래될 수도 있다.³⁶⁾ 즉, 콘텐츠 생산자들은 자신에게 영향을 주었던 우연한 사건이 추천 시스템의 작동 원리와 관련이 있다고 믿고 그에 따라 행동할 수 있는 것이다.

이렇게 놓고 보면 추천 시스템에 대한 비밀주의가 계속 지속되는 것이 사회적으로 바람직하다고만 평가하기는 어렵다. 이러한 문제의식의 산물로서 유럽연합에서는 추천 시스템에서의 순위 결정 메커니즘을 일정한 수준으로 공개할 것을 의무화하는 법이 제정되기도 하였다. 앞으로 미디어 플랫폼의 중요성이 커지고, 추천 시스템의 역할이 점차 증가하면서 추천 시스템의 원리를 공개하라는 요구는 더욱 커질 것이고, 이에 대한 적절한 사회적 대응이 필요하게 될 것이다.

V. 결론

인터넷은 우리에게 무한대에 가까운 선택지를 제공한다. 더 많은 선택지는 이 시대의 풍요를 상징한다. 하지만 이를 두고 과연 축복이라고만 볼 수 있을까? 우리는 지나치게 많은 선택지에 압도된다. 대형마트 매대에 무수히 많은 제품이 진열되어 있더라도 매번 같은 상품을 사게 되지 않던가. 무한대의 선택지가 주어진 현실에서는 추천 시스템은 반드시 필요하게 된다. 한편으로는 과거 소수에게만 주어지던 값비싼 개인 맞춤형 추천 서비스가 대중화됨으로써 사회적 후생을 증진시킬 수 있는 반면, 다른 한편으로는 여러 윤리적, 법적 우려가 제기되고 있기도 하다. 추천 시스템에 대한 적절한 사회적 규율은 우리에게 주어진 중요한 과제다.

32) Jennifer Cobbe and Jatinder Singh, "Regulating Recommending: Motivations, Considerations, and Principles", *European Journal of Law and Technology*, Vol.10 No.3, 2019.

33) Jennifer Cobbe and Jatinder Singh, 앞의 논문.

34) https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=65694

35) 이는 한 유명 유튜브가 자신의 경험을 소개한 내용이다. <https://www.youtube.com/watch?v=fHsa9Dqmid8>

36) B. F. Skinner, "Superstition" in the pigeon", *Journal of Experimental Psychology* 38.2 (1948). 스키너는 베고픈 비둘기들은 새장에 가두어 두고 일정한 간격으로 시점에 먹이를 제공했다. 비둘기들은 자신의 특정한 행동이 먹이의 제공과 관련이 있다고 믿는 것처럼 특이한 행동을 보였다. 새장을 반시계 방향으로 도는 비둘기도 있었고, 머리를 좌우로 움직이는 경우도 있었다. 다른 비둘기는 바닥을 쪼는 듯한 행동을 계속했다.