AI 이노베이션 레포트

8주차 강의에서는 현재 가장 유망한 기술로 뽑히는 GAN, 상용화가 많이 되어 있는 추천 알고리즘과 자연어 처리 기술에 대한 전반적인 백그라운드를 설명해주셨다. GAN같은 경우는 기존에 학습한 데이터를 기반으로 가짜를 만들어 내어 그러한 가짜를 구분하는 구별자와의 상호 작용을 통해 더더욱 결과물을 발전시켜 나가는 알고리즘으로 인간만의 고유의 영역이라고 할 수 있는 창작의 영역에 발을 담글 수 있는 가능성을 제시하였다. 이러한 GAN 기법을 통해 새로운 예술 작품을 창작하기도 하고 상업용으로 사용할 다양한 콘텐츠들을 시간 대비 효율적으로 생성해 전반적인 비용 절감을 야기할 수 있다. 하지만 기술의 정밀도와 진짜와의 차이가 거의 없음은 여러 윤리적인 문제나 기술의 악용도 충분히 낳을 수가 있다. 고로, 제도적인 규율 장치가 같이 공존하면서 나아가야 되는 기술 중 하나로 볼 수 있다. 추천 알고리즘 같은 경우는 이미, 소개하신 유튜브, 넷플릭스 외에도 수 많은 SNS나 여러 플랫폼에 많이 구축이 되어 있다. 크게는 기존에 사용자의 학습 데이터를 가지고 유사한 다른 사용자의 데이터와 비교하여 예측을 통한 추천 서비스를 제공하는 협업 필터링과 선호하는 콘텐츠에 대한 내용과 유사한 다른 콘텐츠를 제공해주는 내용 기반 필터링, 그리고 이 둘을 적절히 병합시켜 후보군을 추려내는 하이브리드 필터링 이렇게 3가지로 추천 알고리즘이 구성이 되어 있다. 유튜브 같은 경우는 이러한 후보군을 선정하는 것과 같이 수 많은 데이터의 양 대비 시간 복잡도를 개선하기 위한 랭킹 모델을 같이 도입함으로서 선호도의 수치로 정렬을 해주어 더 효과적으로 추천을 제공해주는 서비스를 만들어내기도 한다. 마지막으로는 가장 많이 상용화되어 있는 자연어 처리 기술을 소개해 주셨는데 이는 텍스트로나 음성으로 존재하는 수 많은 데이터를 토큰화와 형태소 분석을 통해(pos 태킹 방법 같은) 인간의 언어를 분석하고 이를 이해하는 자연어 이해 과정과 이러한 이해를 바탕으로 수 많은 언어 조합을 학습하고 분석하여 인간과 매끄러운 대화를 이어나갈 수 있도록 하는 자연어 생성 과정으로 구성이 된다. 많이 상용화가 되었다고 하지만 기존의 형식화된 대화가 아닌, 인간과 자연스러운 대화가 가능한 자연어의 모호성을 처리할 수 있는 지의 여부는 자연어 처리 기술이 해결해야 할 과제 중 하나라고 볼 수 있다.

개인적으로 GAN의 기술은 쉽게 이러한 알고리즘을 현재로서는 상용화하는 것은 조금 조심스러운 부분이다. 그 기술이 사람이 직관적으로 분석을 하기 어렵게 되면, 아직 이 기술에 대한 이해가 떨어지는 일반 대중들은 생성된 가짜를 그대로 믿을 수 있게 되기 때문이다. 인증된 몇몇, 자격이 부여된 몇몇의 전문가와 기업을 제외한, 이 GAN 기술은 아직까지 대중에 공시가 되어서는 안된다고 생각한다.