

## 컴퓨터학콜로키움 (COSE-405)

### :: 3주차 임희석 교수님

강의 : 2022-09-21 / 작성 날짜 : 2022-09-23

고려대학교 컴퓨터학과 2017320108

고재영

우리는 살면서 수많은 것을 보고 듣고 경험하며 감각을 느낀다. 인간 뿐만 아니라 무릇 생물이 라면 당연할 것이다. 하지만 그 중에서도 급변하는 사회에서 살아가고 있는 현대인들은 특히나 텍스트 문자로 표현되는 글을 읽는 일이 사소한 일상 행위에서부터 중요한 업무에서까지 매우 높은 빈도를 가지고 있다. 이에 관련하여 인공지능 기술의 메인 응용분야 중에 하나인 자연어 처리에 대해서 빼놓지 않을 수 없을 것이다. 오늘 이 시간에는 임희석 교수님께서 자연어 처리 (NLP)에 관한 최신 연구 동향에 대한 강연을 하셨다.

먼저 임희석 교수님께서 교수님만의 기본적인 인공지능과 관련한 현재 개발자에 대한 높은 수요를 강조하시며 강연을 시작하셨다. 이러한 AI 분야에서 인간의 지능적인 기능을 수행할 수 있는 컴퓨팅 기술을 개발하는 데에 초점을 맞추어, 현재는 인간이 컴퓨터보다 더 잘 수행가능한 항목에 관하여 기계가 수행하기 위한 연구가 진행되고 있다. 특히나, 이번 강연의 초점인 자연어 처리 분야는 말을 이해하고 생성하는 능력에 관심을 가지는 분야라고 할 수 있다. 임희석 교수님이 강연 중에 역설하신 내용의 핵심은 바로 사용자가 원하는 것을 컴퓨터에게 말하는 행위만으로 컴퓨터가 이를 이해하고 적합한 처리 과정을 수행할 수 있다면, 아무리 개선시키더라도 다분히 억지스럽고 불편함을 느낄 수 있는 유저 인터페이스의 한계를 뛰어넘을 수 있을 것이라는 부분이다. 자연어 처리의 핵심적인 내용은 지식에 대한 습득과 표현방법 2가지로 볼 수 있다. 텍스트 문자와 같은 글이라는 모델리티로 표현되어 있는 대부분의 지식을 인간의 경우 문화생활을 영위하면서 자연스럽게 습득한다. 그리고 새로운 지식은 해당하는 분야에 관련한 지식을 기반으로 새로이 학습하곤 한다. AI의 경우도 이와 마찬가지로 특정 도메인과 관련된 정보가 있어야 도메인 특화된 기계학습이 가능하다.

이어서 임희석 교수님께서서는 자연어 처리에 관련한 여러 응용 분야와 실제 교수님의 연구실의 주요 연구 분야에 대해 설명을 하시고 마무리하셨다. 자연어 처리의 주요 응용분야로는 문서 사이의 유사성과 클러스터링에 기반한 문서 분류화, 문법 자동 수정, 텍스트로부터 정형화되어 있는 지식을 뽑아내는 정보 추출, 자동질의응답 시스템, 요약 시스템, 그리고 기계 번역 등이 존재한다. 이 중에서 몇 가지 흥미로웠던 분야를 뽑아보자면 먼저 정보 추출을 뽑아볼 수 있다. 교수님께서 직접 언급하신 예시로는 신문 기사로부터 포함된 정보를 뽑아내는 일이었는데, 기록에 민감한 스포츠에서 특히, 어떤 종목에 대해서, 어떠한 날짜에 어떤 장소에서, 어느 팀 간의 경기로 어떤 스코어를 기록했는지 뿐만 아니라, 득점 기록과 같은 주요 이벤트 및 MVP로 선정된 선수와 같은 세부적인 통계도 추출하는 것이었다. 마지막으로 현재 자연어 처리에 있어서 꽃이라고 볼 수 있는 기계 번역에 있어서, 현재 널리 사용되는 OTT 서비스의 자막 생성에 관련한 자동 번역의 경우 종종 문맥의 앞뒤가 상응하지 않는 마치 환각 상태인 것만 같은 할루시네이션 현상이 일어나는 한계가 있는데, 도메인에 대한 긴밀한 배경지식을 바탕으로 한 대화모델로 common sense에 부합하는 말을 생성하는 데에 크게 주력하는 것을 설명하셨다.

이번 강연을 기회로 자연어 처리에 관련하여 생각해보게 되었다, 필자는 처음에 생각하기에 실제 세상에서 대부분의 지식이 글이라는 모델리티의 형태로 축적되어 있어 이를 기반으로 하는 자연어 처리의 기조로 미루어 보건대, 글 이외에도 구전되어 오는 형태의 지식이 있기 때문에 이를 음성 데이터와도 연관지어 볼 수 있지 않을까 생각해보았다. 아니나다를까 이를 위해 스스로 좀 더 찾아보니, 텍스트 기반 인공지능 모델의 텍스트에 의존하는 성향을 벗어나, 페이스북 AI에서는 '생성적 화자 언어 모델'로 텍스트에 벗어나 오디오 신호에 직접 고성능 처리가 가능한 이른바 GSLM을 개발했다고 한다. 이전의 음성 입력에 관한 자연어 처리모델은 자동 음성 인식(ASR) 시스템의 훈련에 의존적이었는데 여러 오류와 소수의 언어에만 제한적이었던 반면, GSLM은 이러한 ASR 시스템에 벗어나 음성 데이터 입력을 음성 출력에 엔드 투 엔드 방식으로 전환할 수 있는 의의를 가진다. 텍스트 기반은 리소스 집약적인 작업이었기 때문에, 광범위한 데이터셋이 결여된 언어의 경우 적용하기 힘들었던 한계점을 해결할 수 있다는 점이 흥미로웠다.