

최신 ICT 이슈

Ⅲ . 인도 연구팀, 딥러닝을 이용해 시에 ‘풍자’를 이해시키는 연구 진행

◎ 인도 파트나(Patna) 공과대학 학장인 Pushpak Bhattacharyya 는 딥러닝을 통해 인터넷 상에서 반어적(아이러니) 표현의 의견과 악의적 발언을 검출하는 연구를 진행하고 있음

- ▶ 부정적 내용을 긍정적인 표현에 담아 우회적으로 전달하는 풍자(Sarcasm)는 말 그대로의 의미와 속뜻이 다르기 때문에 말하는 사람의 감정이나 상황 맥락을 이해하지 못하는 컴퓨터는 이해할 수 없는 것으로 알려져 왔음
- ▶ 그러나 Bhattacharyya 교수팀은 대량의 텍스트 데이터를 딥러닝을 통해 분석하고, 말과 의미의 관계성에서 문장에 내포된 풍자를 컴퓨터가 파악하게 하는 연구를 진행 중이며, 자신의 연구팀 외에 언어학자, 심리학자와 공동으로 작업을 진행 중
- ▶ Bhattacharyya 교수는 풍자나 아이러니 분석 연구가 국가 원수, 정치인, 유명 연예인, 기업 등 세간의 평판을 신경 써야 하는 사람들에게 도움을 줄 수 있을 것이라고 말하고 있음
- ▶ 이들은 주로 글 쓴 사람의 감정 분석을 통해 트위터와 같은 소셜 미디어에서 자신들의 평판을 체크하고 있는데, Bhattacharyya 에 따르면 풍자야말로 사람의 감정의 움직임을 명확히 보여줌에도 불구하고 기존의 감정 분석 기법은 이를 제대로 이해하지 못하는 문제가 있음

◎ Bhattacharyya 교수에 따르면 딥러닝을 이용해 새로 설계한 자신의 알고리즘이 기존 감정 분석 방식보다 3배 가량 정확하게 풍자를 감지해 낼 수 있었다고 함

- ▶ Bhattacharyya 교수의 조사에 따르면, “배터리가 2 시간이나 버텼네, 멋지군” 등과 같이 풍자나 아이러니를 담고 있는 트윗 중 약 20%는 숫자를 포함하고 있는데, 이런 문장들에서 일반적인 감정 분석 기법은 발화자의 원래 의도를 잘 캐치하지 못한다고 함
- ▶ 본래 비꼬는 말투를 알아차리려면 문맥을 이해할 수 있어야 하는데, Bhattacharyya 교수는

* 본 내용과 관련된 사항은 산업분석팀(☎ 042-612-8296)과 최신 ICT 동향 컬럼리스트 박종훈 집필위원(soma0722@naver.com ☎ 02-576-2600)에게 문의하시기 바랍니다.

** 본 내용은 필자의 주관적인 의견이며 IITP의 공식적인 입장이 아님을 밝힙니다.

“모순된 표현이나 생경한 감정 표현이 풍자의 특징”이라며, 기존의 감정 분석 기법은 일반적 패턴과 다른 표현이 이어지는 것 때문에 본뜻을 이해하지 못하다고 설명

- 문제 해결을 위해 Bhattacharyya 교수는 말이나 표현의 불일치를 검출하는 알고리즘을 설계했으며, 딥러닝을 구성하고 학습시킬 수 있는 라이브러리인 CuDNN 에 고속화된 텐서플로우(TensorFlow) 프레임워크를 도입하고, 엔비디아의 GPU 를 이용한 신경망을 구축하였음
- 그 다음 대량의 트윗, 영화 평론, 1990 년대 인기 시트콤 프렌즈의 대사 등 대량의 텍스트 데이터를 준비해 딥러닝을 이용한 데이터 분석을 실시하였음
- 분석 결과 Bhattacharyya 교수가 설계한 알고리즘은 기존 방식보다 정확하게 풍자나 아이러니를 감지 할 수 있었으며, 특히 숫자를 포함한 트윗에 대해서는 기존 감정 분석 기법보다 약 3 배 가량 높은 80%의 정확도로 풍자를 이해할 수 있었다고 함

(Nvidia, 1. 31. & Gigazine, 2. 1.)