과학과 사회

# AI는 세상을 더 공정하게 만들 것인가: AI, 데이터, 그리고 사회 편향

2021년 10월 22일

윤진혁

필자는 매일 저녁 프로야구를 시청한다. 많은 사람이 필자처럼 야구를 보며 하루의 스트레스를 풀기도 하고, 같은 팀을 응원하며 서로 감정을 공유하기도 한다. 코로나19로 인해 언택트 시대로 접어든 이후로는 야구장에서 함께 응원하지 못 하는 아쉬움을 인터넷 포털의 실시간 채팅으로 달랜다. 누구나 팔이 안으로 굽는다. 우리 팀이 손해를 본다고 생각하면 댓글 창은 심판에 대한 성토가 이어진다.

지난해 동안 한국 프로야구의 모든 투수는 합쳐서 220,874개의 공을 던졌다.[1] 이 중 배트에 맞지 않은 모든 공을 심판이 스트라이크 혹은 볼로 판정하게 되어있는데, 매 경기 석연치 않은 판정이 나온다. 팀에 따라서 같은 코스의 공도 서로 다른 판정을 받는 것처럼 보이기도 한다.([그림1]) 오심도 경기 일부라는 야구 격언이 있다. 하지만 다르게 말하자면 심판이 공정하지 않다면 경기의 양상을 크게 바꿀 수도 있다는 뜻이기도 하다. 오심은 승패와 무관하게 팬들과 선수들에게 상처를 남긴다.

## “이럴 거면 AI 심판으로 바꿔라”

오심 논란이 생길 때마다 댓글 창은 AI 심판을 도입하자는 글로 도배가 된다. 한국 프로야구도 이러한 여론의 영향으로 2020년부터 퓨처스리그(2군) 경기에서 로봇 심판을 테스트 중이다. 미국 마이너리그는 올해도 스트라이크-볼 판정을 대리하는 로봇을 테스트 중이다.[3] 몇 년 전 정확한 판정을 위해 비디오 판독 시스템 등 신기술을 도입한 데 이어 경기의 중요한 요소인 공의 판정 또한 기술의 힘을 빌리려고 하는 것이다. 사람들은 AI 기술이 더 공정한 경기를 만들어줄 것이라고 믿는 것이다.


그림1 2021년 9월 7일 롯데 자이언츠 대 삼성 라이온즈 경기의 스트라이크존. 볼 판정이 많아질수록 공격자에게 유리해지고, 스트라이크 판정이 많아질수록 수비자에게 유리해진다. 롯데 자이언츠의 공격보다 삼성 라이온즈의 공격 시 스트라이크존 내의 공이 볼 판정을 받는 비율이 높은 것을 볼 수 있다.[2]

다른 곳에서도 이런 현상을 관찰할 수 있다. 청와대 국민청원에는 주기적으로 AI 판사 도입에 관한 청원이 올라온다. 여론에 따라 대법원 법원행정처는 “사법부에서의 인공지능AI 활용방안”, “손해배상 사건에서의 인공지능AI 활용방안” 등의 연구를 수행하여 AI 판사의 도입 가능성을 타진 중이다. 2016년 미국과 영국의 공동 연구에 따르면 판결문을 통해 학습한 모델이 실제 판사의 판결을 79%의 정확도로 예측 가능했다고 한다.[4] 그로부터 5년 동안 BERT와 GPT-3 등의 대규모 언어 모델이 등장하였고 기술적인 측면에서 자연어 처리의 성능 또한 매우 높아졌으므로 훨씬 더 정확하게 판결을 예측할 수 있을 것이다.[5,6]

2020년 12월 한국리서치 설문 조사에서는 법원 판결에 대해 신뢰한다는 판결이 단 29%에 그쳤고, 86%의 응답자는 법원에서 선고하는 형량이 판사에 따라 달라진다고 답변하였다.[7] 심지어 본인이 재판을 받을 때 선택이 가능하다면 AI 판사를 고르겠다는 답변이 48%에 달했는데, 이는 인간 판사를 선택한 39%보다 유의미하게 높은 비율이다. 단순히 즐기기 위한 스포츠뿐 아니라 본인의 삶과 더 밀접한 곳에서도 사람의 판단보다 AI의 판단이 더 공정할 것이라는 믿음이 퍼져 나가고 있다.

## 과연 AI는 공정할 것인가?

AI가 무엇인지에 대한 생각은 사람마다 다를 수 있지만, 2021년 현재 AI라는 단어는 기계학습을 통해서 모델을 학습하고, 학습한 모델을 통해서 문제를 푸는 과정을 통칭하는 개념으로 사용된다. 내부의 세부적인 모델과 메커니즘은 다를 수 있지만, 대부분의 연구는 모델의 구조적 성능을 개선함과 동시에 모델 인자의 수를 늘리고 입력 데이터의 양을 늘려 정확도를 개선하는 형태로 이루어지고 있다. 이 과정에서 모델의 성능은 원 데이터에서 분리한 평가용 데이터를 얼마나 정확하게 맞추는 지로 평가된다. 좋은 성능의 모델은 현재 가지고 있는 데이터를 가장 잘 반영한 모델이다.

하지만 최근 도덕적인 AI, 인간에게 도움이 되는 AI에 관한 관심이 커지며 이러한 평가 방향성에 대한 비판의 목소리가 커지고 있다. 2019년 6월의 구글 AI 포럼에서는 데이터의 편향성에 따른 다양한 AI의 문제점을 제시하였는데, 기계학습을 통한 성별 분석에서 성별에 따른 특성 이외에 인종적, 사회적, 소득, 종교적 특성이 영향을 끼쳐 잘못된 분석을 제시하는 경우가 많았으며, 반대로 이러한 경우에 찾아낸 편향성을 통해 인간의 편향성을 찾아내는 결과도 도출할 수 있다는 것을 밝혀냈다.

이를 확인할 수 있는 예로 구글의 이미지 검색에서 교수professor 와 교사teacher를 검색한 결과를 살펴보자.([그림2]) 영어는 성별에 따라 단어가 달라지지 않으므로 두 단어에는 직접 성별을 나타내는 정보는 없다. 하지만 검색에 나타난 결과는 매우 다르다. 두 검색어 모두 주로 칠판 앞에 사람이 서 있는 모습이 검색되는데 교수는 주로 남성이, 교사는 주로 여성의 이미지가 검색된다. 더 놀라운 사실은 두 검색어 모두 유럽계 미국인을 제외한 아시아계 및 아프리카계 등의 소수민족은 거의 검색되지 않는다는 것이다. 당연히 이 검색어에는 인종적 정보도 들어가 있지 않다. 이러한 편향은 어디에서 오는 것일까?

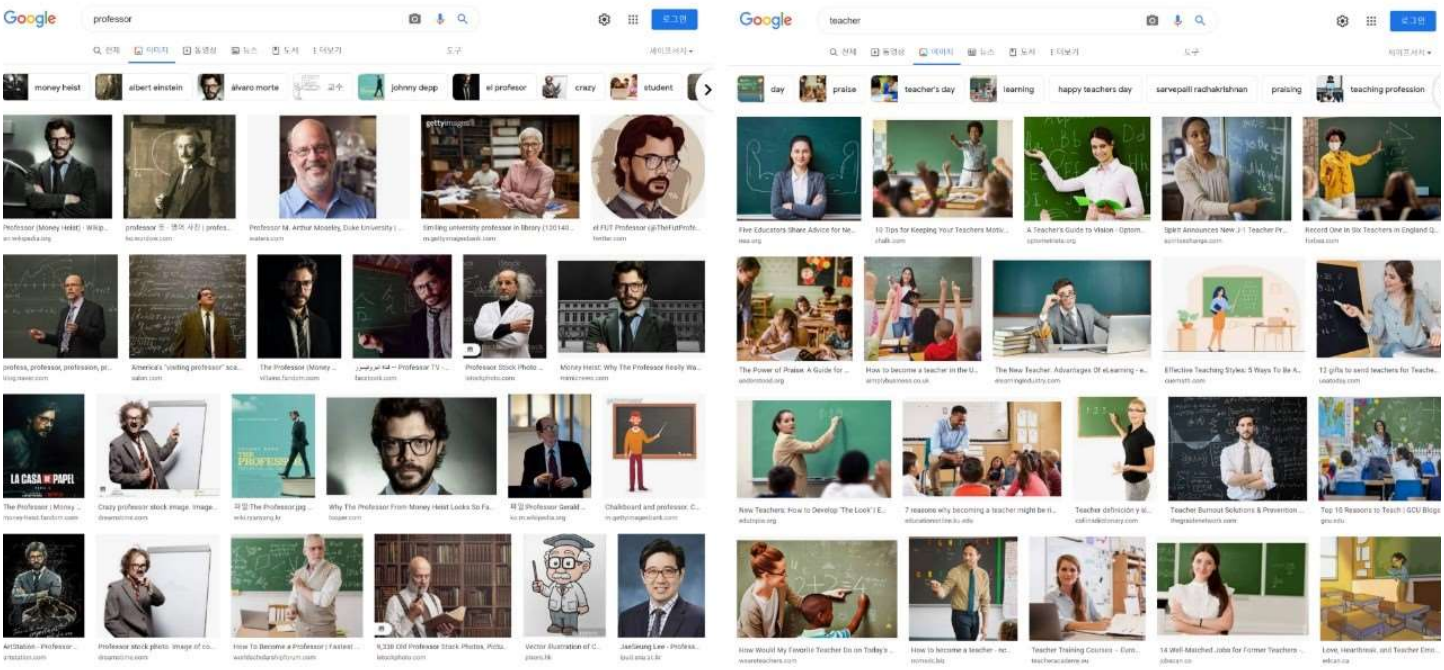

그림2 구글 이미지 검색에서 나타난 성별 편향의 예시. 좌측은 교수Professor를, 우측은 교사Teacher를 검색한 결과이다. 검색어 자체에는 어떠한 성별과 정보도 담고 있지 않지만, 사회의 구조에서 나타나는 직업별 성비에 의하여 교수는 주로 칠판 앞의 남성 이미지가 교사는 주로 칠판 앞의 여성 이미지가 검색된다. 또한 두 가지 경우 모두 유색인종이 거의 나타나지 않는다.

미국교육통계센터National Center for Educational Statistics에 따르면 2018년 기준 미국 전체 정년직 교수의 56.3%는 남성, 43.7%는 여성이었다. 높은 직위로 갈수록 그 편차는 커지는데, 정교수의 경우 66.5%가 남성, 부교수는 54.1%가 남성, 조교수는 47.7%가 남성이었다. 일반적으로 정교수의 경력 기간이 부교수와 조교수에 비해 길기 때문에 더 많은 기

사나 사진 등이 남아 있을 확률이 높을 것이다. 반면 동일 연도 기준 미국 공립 초등학교와 중고등학교의 경우 남성 교사의 비율은 23.5% 여성 교사의 비율은 76.5%였다.[9] 유사하게 공립교사의 경우 79.3%, 교수의 경우 71.4% 정도가 유럽계 미국인이었다. 사회에서 나타나는 직업 선택의 불균형이 데이터에도 그대로 반영된 것이다. 이러한 문제에 대한 고려 없이 모델을 학습하게 된다면, 칠판 앞에 사람이 서 있는 경우 남성은 교수 여성은 교사로 판단하게 될 것이다.

더 큰 데이터라면 어떨까? 많은 연구에서 Google Image, Google News 등의 검색 결과를 통해 학습 데이터를 구축한다. 1,400만 개 이상의 라벨과 이미지로 구성된 ImageNet 또한 이런 편향 문제에 자유롭지 않다.[10] 최근 연구에 따르면 ImageNet 데이터의 약 45%는 미국이 출처라고 한다. 미국의 인구는 전 세계의 단 4%에 불과하다. 반면 세계 인구의 36% 정도가 사는 중국과 인도의 경우 겨우 3% 정도였다. 이러한 편향은 이미지뿐 아니라 번역에서도 동일하게 나타난다. 한국어 문장을 영어로 번역하는 예를 들어보자. “걔는 경찰이야” 라는 문장에서는 성별 정보가 없지만 번역기는 “He is a cop”이라고 번역한다.[11] 반대로 “걔는 간호사야” 라는 문장을 번역기는 “She is a nurse”로 번역한다. 이렇듯 성별을 모르는 상태에서도 사회의 직업에 대한 편견에 의하여 모델이 특정 성별로 추정하는 것은 상당히 흔한 일이다. 특정 성별이 전체적인 집단을 과도하게 대표하는 것이다.

이러한 모델이 표준적으로 사회에 사용되면 어떠한 결과를 초래할까? 미국 컨슈머 리포트에 따르면, 1998년부터 2008년까지 10년간 자동차 사고 기록을 분석한 결과 남성의 사망률 보다 여성의 사망률이 평균적으로 높게 나타나는 것을 발견하였다.[12] 이와 유사하게 2019년 버지니아 대학교University of Virginia의 연구 결과에 따르면 자동차 사고 시 여성이 남성보다 중상을 당할 확률이 73%나 높았다고 한다.[13] 연구진은 이러한 원인이 자동차 실험에 사용하는 더미가 남성의 평균적인 체형에 맞추어 만들어져 있어 안전띠 등의 설계가 남성이 더 안전하도록 만들어져 있기 때문이라 추정하였다. 의료에서도 비슷한 문제가 발생한다. 오랜 기간 백인과 남성의 신체는 의학계 표준처럼 쓰여 왔다. 이로 인해서 의료 장비에 탑재된 소프트웨어에서 피부가 어둡거나 여성의 경우 판정 오류의 확률이 증가할 수 있다는 사실이 보고되었다.[14] 다양한 집단의 광범위한 문제를 풀기에는 현재의 데이터에 기반한 모델이 적절하지 않을 수 있다는 것이다.

## 단순한 숫자로 보이지 않는 숨어있는 데이터 편향성

위의 예시들은 데이터의 비율이나 결과로 나타나는 수치에서 명확하게 성별 혹은 인종 간 차이가 보인다. 사실 이러면 학습 데이터마다 집단의 비율을 조정하는 단순한 방식을 통해서 편향을 줄일 수 있다. 하지만 이렇게 숫자로 보이는 편향이 전부일까? 사람들의 평가가 평가받는 집단의 특성에 따라 달라질까? 연구에 따르면 아프리카계가 느끼는 고통의 정도를 유럽계가 느끼는 정도보다 낮게 진단한다고 한다.[15] 통증에 대한 평가가 부정확하다면 향후 치료에 악영향을 미칠 수 있다. 이러한 예에서 수치상으로 보이지 않는 편향성이 존재할 수 있다는 것을 추정할 수 있다.

AI가 흑인의 재범률을 백인보다 높게 판단해서 가석방 심사와 형량에 부정적인 영향을 주었다는 연구도 있다.[16] 미국 법원에서 사용하고 있는 COMPASCorrectional Offender Management Profiling for Alternative Sanctions는 범죄의 형태와 종류, 개인의 성격, 가족 구성 등의 영역을 종합해 형량을 선고한다. 하지만 이 시스템은 유사한 사회적 상황임에도 불구하고 흑인을 백인보다 2배나 더 많이 재범 위험군으로 예측했다. 이와 유사하게 영국 경찰이 범죄 위험도 예측에 사용했던 HARTHarm Assessment Risk Tool 시스템에서는 인종과 함께 빈부에 대한 편향도 나타났다.[17] 가난한 사람들이

범죄를 더 자주 저지를 확률이 높다고 예측한 것이다. 실제로 저소득층이 범죄를 더 자주 저지를 수도 있다. 하지만 저소득층은 적절한 변호를 받기 더 어렵기 때문에 같은 행동을 해도 유죄 판결을 받을 확률도 높아질 것을 추정해 볼 수 있다. 즉, 범죄의 원인에는 매우 다양한 요인이 있으나 현재의 데이터 수집과 모델의 한계로 특정 집단에 더 불리한 예측을 내 놓을 수 있다는 것이다.

"

위키백과의 편집자들은 토론 끝에 스트리클랜드 교수의 문서를 삭제했다.

그리고 스트리클랜드 교수는 노벨상을 받았다. 노벨상을 타는 것보다 위키백과 문서가 생기는 것이 어려웠던 여성 교수가 있었다. 이 사실은 우연일까?

"

또 다른 예를 들어보자. 2018년 노벨 물리학상은 광학 분야의 세 명의 과학자에게 수여 되었다. 각각 광학 집게Optical Tweezer를 발명한 아서 애쉬킨Arthur Ashkin 박사, 그리고 극초단 레이저 펄스를 높은 출력으로 증폭하는 처프 펄스 증폭Chirped Pulse Amplification을 발명한 제라드 무루Gérard Mourou 교수와 도나 스트리클랜드Donna Strickland 교수이다. 이 중 스트리클랜드 교수만 노벨상 수상 전에는 위키백과에 등재되지 못했었다. 재밌는 사실은 도나 스트리클랜드 교수의 위키백과 문서가 노벨상 전에 만들어진 기록이 있다는 것이다.[18] 하지만 위키백과의 편집자들은 토론 끝에 스트리클랜드 교수가 위키백과에 등재되기에 과학적으로 충분한 성취가 없고 명성도 높지 않다는 결론을 내렸다. 문서를 삭제해버린 것이다. 그리고 스트리클랜드 교수는 노벨상을 받았다. 위키백과 문서가 생기는 것이 노벨상을 타는 것보다 어려웠던 여성 교수가 있었다. 이 사실은 우연일까?

연구 결과에 따르면 이는 우연이 아니다. 연구자의 연구 성취를 평가하는 지표로 h-index라는 것이 있다. 이 방법은 연구자가 발표한 논문과 논문별 인용 수를 이용해서 구하는데, x번 이상 인용된 논문이 x개 있을 때 이 연구자의 h-index를 x라고 정의한다.[19] 다시 말해 h-index가 높은 연구자들은 양과 질적으로 모두 우수한 연구자로 볼 수 있다. 연구자의 명성을 판단하는 또 다른 단순한 방법은 위키백과에 문서가 존재하는지 확인하는 것이다. 2019년 발표된 연구에서는 물리학, 경제학, 철학 세 개 분야 총 15,049명을 대상으로 하여 저자의 h-index에 따라서 위키백과 문서가 있는 저자의 비율을 분석하였는데, 남성보다 여성의 경우 그 비율이 낮았다.[20] 동등한 정도로 성취했다고 추정할 수 있는 저자들이지만, 여성이 더 저평가를 받고 있다는 것이다.

왜 이러한 문제가 발생하는 것일까? 위키백과의 자체 조사에 따르면 위키백과 전체 편집진의 10% 정도만 여성이고, 나머지는 남성 혹은 성별을 밝히기를 원하지 않았다고 한다.[21] 즉 참여자 구성 자체에 편향성이 있었으며 그로 인해서 데이터에 자연스럽게 편향성이 녹아들었다고 볼 수 있다. 꽤 많은 자연어 모델들이 위키백과에서 수집한 문장들을

토대로 데이터를 구축해 학습하고 있다는 점을 생각해보면, 많은 모델이 의도치 않은 편향성을 가지게 되었음을 추측해볼 수 있다.[6,7]

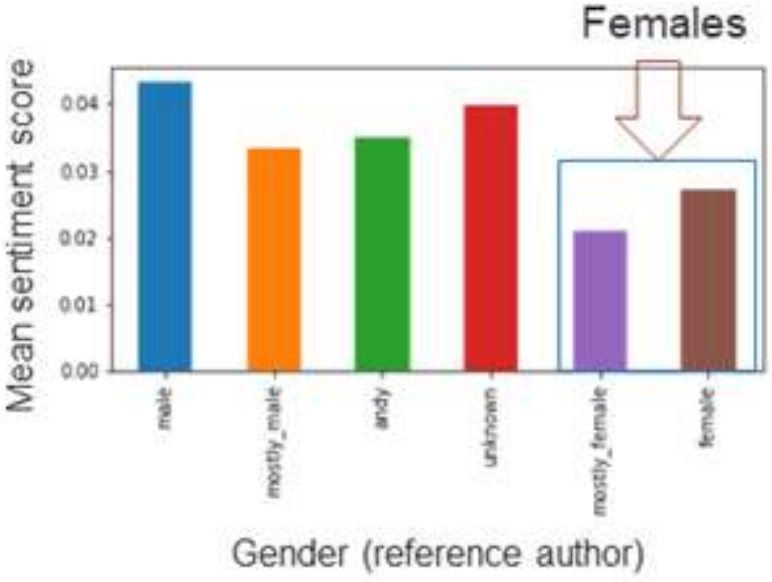


그림3 논문 저자의 성별에 따른 인용 문장의 평균 긍정도. 본 결과는 Microsoft가 제공하는 서지 데이터인 Microsoft Academic Graph의 모든 인용 문구를 토대로 필자가 Vader Sentiment Analyzer를 통해 직접 구하였다. 여기서 y축의 값이 클수록 평균적인 인용이 긍정적인 것을 뜻하고, 값이 작을수록 부정적인 것을 의미한다. 여성 이름을 가진 저자의 경우 남성이나 성별이 불분명한 이름을 가진 저자에 비해서 부정적인 인용을 더 많이 받는 것을 알 수 있다.

마지막으로 필자의 최근의 연구 결과를 이야기해보고자 한다. 이 결과는 아직 출판되지는 않은 결과이다. Microsoft의 서비스 중에 Google Scholar와 유사한 Microsoft Academic이 있다 (훌륭한 서비스지만 안타깝게도 올해 12월에 종료 예정이다). 이 데이터에는 1600년대부터 현재까지 출판된 논문 약 8천만 건과 이 논문 사이의 인용 관계가 수록되어 있다.

필자는 Microsoft Academic에서 약 2억 4천 건의 인용 문맥을 추출하여서 역사적으로 과학기술의 인용이 어떤 맥락에서 이루어졌는지 분석해 보았다. 사실 모든 인용이 동등하게 이루어지지 않는다. 다양한 분류법이 있지만 가장 단순하게는 긍정적, 중립적, 부정적인 인용으로 나눌 수 있다. 부정적인 인용은 저자가 기존 연구를 비판한다고 볼 수 있으며, 긍정적인 인용은 저자가 기존 연구의 중요성을 높게 평가했다고 볼 수 있다. 사실 많은 인용은 가치판단이 없는 중립적 인용이지만 적지 않은 수의 저자가 기존 연구를 긍정하거나 부정하곤 한다. 재미있게도 이름을 통해 피인용 된 논문의 저자 성별을 확인해 본 결과, 여성 저자로 추정되는 논문 인용 문구가 더 부정적인 경향을 보였다.([그림3]) 논문의 질이나 연구자의 분야 등에 대한 추가적인 분석이 더 필요하지만, 전체적인 경향성에서는 여성 저자가 남성 저자에 비해서 평균적으로 부정적인 인용을 많이 받고 있는 것이다.

## 마치며

2016년 3월 알파고 쇼크 이후 5년이 지났고, 짧은 시간 동안 딥 러닝과 대용량 데이터를 기반으로 AI 모델이 급속도로 도입되기 시작했다. IT와 비IT 분야를 가리지 않고 AI에 기반한 의사결정을 도입하려는 시도가 커지고 있으며 심지어는 도입부의 예처럼 스포츠의 심판과 법률의 판단까지 AI에 맡기려는 경향이 나타난다. 이러한 경향은 AI가 인간의 선택보다 더 공정할 것이란 인식에서 기인한다. 다시 말해, 많은 사람이 AI를 통해 현재의 편향된 사회 구조와 다양한 불평등과 불합리가 제거 혹은 완화될 것으로 기대하고 있다.

안타깝게도 현존하는 AI는 사회의 편향성을 그대로 담아내는 거울에 가깝다. 데이터를 만드는 주체인 각 집단이 가진 편향성이 실제 데이터에 반영되는 것을 막아낼 수 없으므로, AI에 대한 의존성이 높아지면 높아질수록 사회적 편향성이 가속화될 위험에 노출된 것이다. 즉, 사회의 기대와 실제 기술의 발전 방향이 다르게 변화하고 있다. 이러한 현상의 개선을 위해서는 사회적 편향성을 제거한 새로운 AI 모델을 제시할 필요가 있다. 모델에 학습된 편향성은 다양한 사회적 문제를 일으킬 가능성이 크다. 2014년 아마존은 이력서 검토에 비밀리에 AI를 도입하였으나 금세 사용을 중단했다. 기존의 이력서와 부서장의 평가를 학습한 모델이 남성 위주의 정보통신업계의 편향성을 그대로 담아내어 지속해서 남성보다 여성에 부정적인 평가를 했기 때문이다.[22]

> 숨어있는 편향성을 제거하기 위해서는 데이터가 만들어지는 사회와 시스템 자체에 대해 더 많은 이해가 필요하다.

오늘도 필자는 야구를 본다. 한국 프로야구 심판의 판정은 세계 최고 수준으로 정확하다고 한다. 하지만 마음에 들지 않는 스트라이크 콜 하나에 AI 심판은 더 나을까 하는 생각은 지울 수가 없다. 한국 프로야구 규정집에는 스트라이크 존이 이렇게 표현되어 있다 “유니폼의 어깨 윗부분과 바지 윗부분 중간의 수평선을 상한선으로 하고, 무릎 아랫부분을 하한선으로 하는 홈 베이스 상공을 말한다”. 타자의 키에 따라서 달라지겠지만 키와 자세가 같다면 스트라이크 존이 달라지지는 않을 것이다. 이런 경우 모델은 규칙만 배우면 된다. AI는 세상 모든 것이 야구처럼 규칙이 명확하다면 편향되지 않은 결과를 보여줄 것이다. 비디오 판독처럼 심판 판정을 보완하는 형태로 쓴다면 프로야구의 재미를 더할 수 있을 것이다.

안타깝게도 사회는 야구 경기의 규칙보다 훨씬 복잡하다. 재판은 수많은 사안을 법만으로 판단할 수 없어 판례를 참조할 수밖에 없다. 그래서 판결에 편향이 남아 있다면 자연스럽게 다시 모델에 녹아들 수밖에 없다. 학습 데이터에서 통계적으로 확인할 수 있는 편향성은 쉽게 제거할 수 있지만 숨어있는 편향성을 제거하기 위해서는 데이터가 만들어지는 사회와 시스템 자체에 대해 더 많은 이해가 필요하다. 우리가 인간과 사회를 조금 더 깊게 이해한다면 반대로 더 정확하고 편향성이 없는 AI를 만들 수 있을 것이다.

이 글에서 필자는 몇 가지 예를 통해 AI와 데이터의 편향성에 관하여 이야기해보았다. 이러한 편향성의 가장 큰 문제는

사회의 취약한 집단일수록 피해자가 되기 쉽다는 것이다. 이러한 편향성에 대한 심도 있는 분석은 AI와 빅데이터 시대에 사회적인 비용을 감소시키기 위해 필수적이며, 앞으로 이루어지는 연구에서 이러한 편향성을 제거하기 위한 더 많은 시도가 있기를 희망한다.

---

## 참고문헌


1. 한국야구위원회 기록실, Accessed:2021.09.08.
2. 스트존, Accessed:2021.09.08.
3. Playing rules to be tested during 2021 MiLB season, Accessed:2021.09.08.
4. Aletras, Nikolaos, et al. "Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective." *PeerJ Computer Science* 2, 2016.
5. DEVLIN, Jacob, et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." *NAACL-HLT* 2019.
6. BROWN, Tom B., et al. "Language models are few-shot learners." *arXiv preprint arXiv:2005.14165.* 2020.
7. 한국리서치, [기획] 판결의 온도차 – 사법부와 국민 법 감정 사이, 2020.12.23.
8. National Center for Education Statistics, Full-time faculty in degree-granting postsecondary institutions, by race/ethnicity, sex, and academic rank: Fall 2015, fall 2017, and fall 2018,
9. National Center for Education Statistics, Number and percentage distribution of teachers in public and private elementary and secondary schools, by selected teacher characteristics: Selected years, 1987-88 through 2017-18.
10. SHANKAR, Shreya, et al. No classification without representation: Assessing geodiversity issues in open data sets for the developing world. *arXiv preprint arXiv:1711.08536*, 2017.
11. CHO, Won Ik, et al. On Measuring Gender Bias in Translation of Gender-neutral Pronouns. *GeBNLP* 2019.
12. BARRY, Keith. "The crash test bias: How male-focused testing puts female drivers at risk." *Consumer Report*, 2019.
13. Forman, Jason, et al. "Automobile injury trends in the contemporary fleet: belted occupants in frontal collisions." *Traffic injury prevention* 20, 2019.
14. ZOU, James; SCHIEBINGER, Londa. AI can be sexist and racist—it's time to make it fair. 2018.
15. HOFFMAN, Kelly M., et al. Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences* 11, 2016.
16. YAPO, Adrienne; WEISS, Joseph. Ethical implications of bias in machine learning. 2018.
17. OSWALD, Marion, et al. Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality. *Information & Communications Technology Law,* 27, 2018.
18. ERHART, Ed. Why didn't Wikipedia have an article on Donna Strickland, winner of a Nobel Prize?. *Wikimedia blog,* 2018.
19. HIRSCH, Jorge E. An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences,* 102, 2005,
20. SCHELLEKENS, Menno H.; HOLSTEGE, Floris; YASSERI, Taha. Female scholars need to achieve more for equal public recognition. *arXiv preprint arXiv:1904.06310,* 2019.
21. KHANNA, Ayush. Nine out of ten Wikipedians continue to be men: Editor Survey. *Wikimedia blog,* 2012.
22. Amazon scrapped 'sexist AI' tool, BBC News, 2018