

안전보건감독에서 빅데이터와 기계학습의 역할에 관한 유럽 안전보건청(OSHA)의 토론회자료(2)

번역 대외홍보팀

구성

- 서론
- 위험성에 기반하여 근로감독 대상 사업장 선정하기
- 빅데이터와 머신러닝
- 근로감독 대상을 정할 때 빅데이터와 머신러닝 활용하기
- 도전
- 결론

※ '서론'부터 '빅데이터와 머신러닝'까지 연재된 지난 7월호에 이어 8월호에는 '근로감독 대상을 정할 때 빅데이터와 머신러닝 활용하기'부터 '결론'이 연재됩니다.

근로감독 대상을 정할 때 빅데이터와 머신러닝 활용하기

지도학습 알고리즘이든 비지도학습 알고리즘이든 상당한 양의 관찰 데이터와, '해당 기업의 특성'이라고 할 수 있는 그 기업과 관련된 충분히 많은 양의 데이터를 필요로 합니다. 잘 알려져 있듯이, 대부분의 근로감독 기관은 감독 대상 기업이나 근로감독 활동과 관련된 막대한 양의 데이터를 수집하고 저장합니다. 이러한 데이터 중에 유용한 것들은 주로 해당 기업의 고유한 특성이라고 할 수 있는 것으로 종업원 수, 사업수행기간, 소속 산업군(업종), 근로감독 피감 횟수, 근로감독 결과 및 산재 보고 내용 등입니다. 더욱이 이러한 데이터는 근로감독 횟수가 늘어가면서 점점 많아집니다.

이런 내용을 감안할 때, 빅데이터를 활용하여 고위험성 기업들을 선정하는 일은 언뜻 보기에 머

신러닝 알고리즘을 적용하기에 적합한 작업으로 생각됩니다. 그럼에도 불구하고, 실제로 해당 사례를 찾기는 쉽지 않습니다. 하지만 몇 가지 주목할 만한 사례가 있는데, 그 사례들을 보면 근로감독기관들이 위험성 높은 감독대상 기업을 골라내는 과제를 해결하는 데 있어서 빅데이터와 머신러닝 기술이 상당히 큰 역할을 하는 것을 알 수 있습니다.

첫 번째 사례는, 추락사고와 같은 작업장 재해의 사전 예측에 머신러닝 방법론의 사용이 적절했는지를 조사한 연구입니다(Matías 등, 2008). 해당 예측결과가 비교적 정확하였음에도 불구하고, 이 연구사례에는 몇 가지 결함이 있습니다. 머신러닝 알고리즘에 포함된 '기업 특성'들이 근로감독 기관에서 통상적으로 보유한 데이터 형태(해당 기업의 개인정보보호장비 사용 여부나 경영관리 실태 등)가 아니라는 것입니다. 더구나, 추락사고가 근

로감독기관이 관심을 갖는 작업장 위험성 중에서 아주 작은 부분을 차지하는 데 불과하다는 것입니다.

두 번째 사례 또한 연구 관련입니다(Hajakbari and Minaei-Bidgoli, 2014). 해당 연구에 참여한 연구진은 산업재해의 발생 가능성을 예측하는 배점시스템을 개발하였습니다. 한 발 더 나아가서, 해당 연구에서는 기업의 일반적인 몇 가지 특성(해당 기업의 주력 사업, 종업원의 성비, 총 직원 수 등)에 근거하여 그 기업의 각기 다른 형태의 산재사고 발생 가능성을 예측하는 것이 가능하다고 결론지었습니다. 심지어, 그 배점시스템이 주기적인 안전보건점검이 필요한 사업장을 식별하는 데에 유용하게 활용될 수 있다고 하였습니다. 이 연구에 사용된 데이터는 한 근로감독 기관의 데이터베이스에서 가져온 것입니다. 그러나 이 연구의 결함도 이전 사례와 유사한 경우로, 근로감독 기관이 다루는 작업장의 다양한 위험성 중에서 단 한가지의 위험성만을 대상으로 하였다는 점입니다. 더구나, 축소신고의 가능성이 상당히 높은 것으로 알려져 있는 산재 통계 데이터에 의존하고 있다는 문제점을 지니고 있습니다.

세 번째 사례는 노르웨이 노동감독청(NLIA, Norwegian Labour Inspection Authority)에서 위험성이 높은 기업들을 골라내서 근로감독을 수행해야 하는 근로감독관들을 돕기 위해 개발한 프로그램입니다(Dahl 등, 2018). 위험성그룹 예측도구(RGPT, Risk Group Prediction Tool)라고 이름 붙인 이 프로그램에서는 기업들을 예측

되는 위험성에 따라 최하위험성, 하위험성, 상위험성, 최상위험성 등 네 개의 그룹으로 구분합니다. 어떤 특정 기업이 속한 그룹의 위험성이 높으면 높을수록, 이 기업이 향후에 진행될 근로감독에서 관련 산업보건 규정을 상당한 정도로 준수하지 못하고 있는 것이 발견될 가능성이 높아집니다.

각각의 기업이 어떤 위험성 그룹에 속해 있는지는 NLIA의 홈페이지에서 근로감독관들이 확인할 수 있습니다. 따라서 근로감독 대상 기업이 정해지면 그 기업이 속한 위험성 그룹을 근로감독관들이 미리 확인함으로써, 위험성 정보를 감안한 감독 대상 기업 선정이 가능해지는 것입니다. RGPT는 이른바 이항 로지스틱 회귀분석(binary logistic regression analysis)으로 불리는 방법론을 사용한 머신러닝 알고리즘에 의한 예측 모델링을 기반으로 만들어졌습니다. 이러한 회귀모델링에 근거하여, 노르웨이의 모든 기업(대략 23만여 업체)을 네 개의 위험성 그룹으로 구분합니다. 구분 작업은 두 단계로 이루어집니다. 첫 번째 단계에서는, 회귀모델링을 통해 향후 진행될 근로감독에서 안전보건규정의 기준에 크게 못 미치는 경우(기업 또는 사업장)를 식별해 낼 확률을 예측합니다. 두 번째 단계에서는, 예측된 확률값을 사용해서 각각의 기업을 해당되는 리스크 그룹에 구분하여 배정합니다.

RGPT는 최초로 NLIA에서 수행한 대략 35,000건의 안전보건감독을 통해 도출된 등록 자료를 기반으로 개발되었습니다. 이후 안전보건 감독 건수가 늘어나면서 RGPT에 의해서 이루어지는 예측의 정확도는 자동적으로 점점 높아졌습

니다. 이는, 안전보건감독이 실시되고 그 결과가 NLIA의 데이터베이스에 등록되면서 이루어지는 (올바른 예측 또는 빗나간 예측을 통한) 피드백 과정을 통해서, RGPT에서 사용하는 머신러닝 알고리즘이 스스로 예측 오류를 바로 잡아간다는 것을 의미합니다. RGPT는 지도학습 알고리즘의 일종으로 분류되는데, 감독 대상이 되는 기업의 몇몇 특성들에 대한 사전 확인을 통해 안전보건감독이 혹시나 잘못 수행될지 여부를 사전에 예측할 수 있기 때문입니다. RGPT에서 사용하는 데이터는 감독 대상 기업의 일반적인 특성들로 기업의 규모, 소속 산업군(업종), 근로감독 피감 횟수, 근로감독 결과보고서 내용, 업력, 산재 발생장소 및 산재 보고 내용 등입니다.

RGPT의 예측 유효성과 정확성은 매월 점검하게 되는데, 약 18개월 동안 테스트를 해본 결과 높은 위험성 그룹에 속하는 기업들을 골라낼 경우에 머신러닝 알고리즘이 더욱 더 정확하게 작동되는 것을 알 수 있었습니다. 이것은 감독 대상 기업이 속하는 위험성그룹의 상하 여부를 판정할 때 결과가 뒤바뀌는 오류가 거의 없다는 것을 의미합니다.

NLIA에서 개발한 RGPT를 사용하면서 알게 된 것은, 빅데이터와 머신러닝 기술을 효과적으로 활용하면 근로감독 대상을 비교적 정확하게 골라내는 것이 가능하다는 것입니다.

이것과 유사하게 머신러닝 기술을 활용한 접근방법이 두 곳 이상의 유럽 국가에 있는 근로감독기관에서 상당히 유망한 결과값을 도출하면서 시험운용되고 있습니다. 그 중 관련 논문으로 알려진 두 군데의 해당 기관은 스웨덴 노동환경청

(Swedish Work Environment Authority) (Ridemar, 2018)과 네덜란드의 사회복지 고용부인 SZW(Sociale Zaken en Werkgelegenheid) (Jacobusse and Veenman, 2016)입니다. 그런데 데이터 저장방식이나 품질, 데이터 접근방법, 데이터의 구조 등이 다르기 때문에, NLIA에서 개발한 RGPT를 다른 나라의 근로감독기관에 그대로 적용할 수는 없습니다. 뿐만 아니라, RGPT에서 감독 대상 기업을 선정하려면 RGPT 프로그램 내에서 기업의 위험성이 정의되고 구동되는 방식이 RGPT의 내부 알고리즘에서 받아들일 수 있는 방식이어야 합니다.

앞서 언급했듯이 RGPT는, 어떤 기업이 속한 위험성그룹의 등급이 높으면 높을수록 그 기업에 대한 근로감독에서 관련 산업보건 규정을 상당한 정도로 준수하지 못하고 있음을 발견할 확률이 높아진다는 개념이 포함된 위험성의 정의를 기반으로 작동됩니다. 이것은 RGPT가 기업에 내재하는 위험성이 아닌 이른바 관리·통제 위험성과 주로 관련되어 있음을 의미합니다.

관리·통제 위험성은 위험성을 관리하고자 하는(즉, 제반 법·규정을 준수하고자 하는) 능력이나 의사가 있는지 여부로부터 생겨나는 반면에, 기업에 내재된 위험성은 사업장 내의 각종 활동이나 행동(높은 곳으로부터 떨어짐, 화학물질에 대한 노출, 근골격계 부담 동작 등)으로부터 발생합니다.

현실에서는 관리·통제 위험성과 기업에 내재된 위험성이 서로 연관되어 있습니다. 그러나 이렇게 연관되어 있다고는 해도 두 가지 위험성이

서로 떼려야 뗄 수 없는 고도의 실증적인 상호 연관성을 필연적으로 갖는 것은 아닙니다. 따라서, 감독 대상 기업이 두 가지 중 어느 한 가지의 위험성에만 근거를 두었다고 판단할 경우의 RGPT에 맹목적으로 의존하게 되면, 자칫 다른 한 가지의 위험성을 고려하는 것을 빠뜨릴 수가 있습니다.

노르웨이의 규제 법규에서는, 우선 규제 지역이나 위험에 노출된 노동자 집단 또는 고위험성 산업 등을 식별할 때는 내재적 위험성이 강조되는 반면에 감독 대상 특정 기업을 골라낼 때는 관리·통제 위험성이 강조됩니다.

도전 과제

한 쪽 편의 관리·통제 위험성과 다른 한 쪽의 내재적 위험성이 경험적으로 반드시 상호 연관되어 있지는 않다는 사실로부터 우리는 또다른 (아마도 훨씬 규모가 큰) 도전에 직면하게 되는데, 그것은 다름 아닌 빅데이터와 머신러닝 알고리즘의 위험성을 기반으로 한 감독 대상 기업 선정작업에 적용하고자 하는 도전입니다. 이와 관련된 세 가지 사례가 있는데, 그 사례들은 모두 일차원적인 감독 대상 선정 작업의 예로서, 어떤 위험성에 대한 특별한 정의나 조작화(operationalisation)를 기반으로 대상을 선정하는 작업입니다. 그런데 사업장의 위험성은 단순히 한 가지 특정한 형태로만 존재하는 것이 아닙니다. 따라서 감독기관이 법규 위반 단속을 위해 권한을 발동하는 경우는 산재사고, 유해물질 노출, 생물학적 노출, 사회심리학적 위험, 근골격계 위험요인과 소셜 덤핑(개발도상국

이 저임금·저원가로 생산하여 해외에 덤핑 수출하는 일) 등 사업장 내 복수의 위험성과 관련됩니다.

이들 중에는 더욱 하위 유형의 위험성으로 세분되는 경우도 있습니다. 이렇게 다양한 경우에 대응할 수 있는 위험성 모델을 개발하는 것은 서로 다른 위험성들이 반드시 상호 연관성을 갖는 것은 아니기 때문에 굉장히 어려운 작업입니다. 따라서 이렇게 다양한 다수의 위험성에 대응하는 작업은 한 가지 특정한 위험성의 발생 확률을 예측하는 것과는 판이하게 다릅니다(Dahl 등, 2018). 앞의 내용과 관련되어 있기는 하지만, 두 번째 도전과제는 위험성을 기반으로 한 감독 대상 선정 작업을 더 복잡하게 만듭니다. 소위 정치적 함정이라고 하는 것입니다(Black, 2010). 머신러닝 알고리즘은 성공과 실패를 통해 학습하고 적응한다는 면에서 역동적이기는 하지만, 다양한 정치적 견해를 고려하지는 못합니다.

첫 번째, 이해관계자 각자가 처한 정치적 상황은 변화무쌍합니다. 그래서 오늘 우선순위에 놓아야 할 위험성의 형태가 내일이 되면 달라질 수도 있습니다. 두 번째, 정치적 상황은 다면적입니다. 따라서 정치인, 고용주, 피고용인, 언론과 대중 등의 다양한 이해관계자들은 어떤 형태의 위험성을 우선순위에 두어야 하는지에 대해 각기 다른 의견을 갖게 될 수 있습니다. 이는 작업장에서의 위험성이라고 하는 것이 반드시 객관적인 독립체가 아닌 사회적인 구성체라는 것을 의미합니다. 고려할 만한 가치가 있는 세 번째 도전과제는, 근로감독기관들이 감독 대상과 관련된 어마어마한 양의 데이터를 보유하고 있기는 해도,

그 데이터들이 대개 감독 대상 기업의 수준과 연관된다는 사실과 관련됩니다. 그리고 기업의 수준 데이터라는 것이 반드시 고려해야 할 가장 적절한 데이터인 것은 아니라는 사실입니다(참조: Gunningham and Sinclair, 2007). RGPT의 데이터베이스에서, 각각의 기업은 그 기업의 일련 번호와 같은 고유 식별자에 의해 구별됩니다.

머신러닝 알고리즘이 특정 예상 위험성 수치를 특정 회사에 부여하는 능력은 그 기업의 고유 식별자를 통해 이루어집니다. 그렇지만, 모든 잠재적 감독 대상이 고유 식별자에 의해 자동적으로 식별되는 것은 아닙니다. 예를 들어 건설업에서, 감독 대상이 어떤 특정한 건설회사가 아니라 일정 기간 임시로 운영되는 건설 현장인 경우입니다. 이러한 임시 작업장과 관련하여 두 가지 도전과제가 있습니다. 첫 번째, 임시 건설 현장이나 임시 작업장은 고유 식별자로 구별되지 않을 수 있다는 것입니다. 두 번째, 고유 식별자로 구별이 가능하다 해도 임시로 운영되는 현장이나 작업장이 운영을 끝내고 다른 장소로 옮겨갈 경우에 머신러닝 알고리즘이 예측의 성공과 실패를 통해 학습할 기회를 갖지 못할 수가 있기 때문입니다.

결론

위에서 서술한 도전과제를 통해 우리가 알 수 있는 것은, 빅데이터와 머신러닝 기술을 활용한다 해도 고위험성의 감독 대상을 골라내는 작업과 관련한 상당한 어려움이 존재한다는 것입니다. 그렇지만, 어찌되었든 이러한 도전과제가 있다고 해서

위험성 기반의 접근방법에서 빅데이터와 머신러닝 기술의 유용성이 사라지는 것은 아닙니다. 오히려 이러한 도전과제들로부터 우리가 알 수 있는 것은, 위험성 기반의 감독 대상 선정이 머신러닝 알고리즘에 온전히 의존하지 않을 수도 있다는 것입니다.

위에 서술한 NLIA의 사례도 이런 내용을 보여줍니다. 머신러닝 알고리즘을 활용해서 감독 대상을 직접 골라내는 것보다는, 감독기관이 머신러닝 알고리즘이 도출해낸 예측을 기반으로 위험성 정보에 입각한 감독 대상 선정을 하는 것이 나을 것입니다. 이는 인공지능과 인간지능이 상호 보완을 통해서 시너지 효과를 내는 것입니다. 대체적으로, 복잡한 사회적 이벤트와 관련된 예측이 필요할 경우에 인공지능과 인간지능을 결합하여 과업을 수행하는 것이 필수적이라고 여겨집니다. 🐼