

인공지능기반 정밀의료 및 신약개발

2017320114 컴퓨터학과 최재원

이번 수업에서는 강재우 교수님의 인공지능기반 정밀의료 및 신약개발에 대해서 배울 수 있는 시간을 가졌다. 본격적으로 본론에 들어가기 전에 기본적인 배경 지식에 대해서 다시 짚고 넘어갈 수 있었다. 데이터 마이닝은 데이터의 광산으로부터 value를 찾아내는 것이다. 빅데이터에 기계학습 방법론을 적용을 시켜서 어떠한 의미 있는 knowledge을 추출해 내거나, 어떤 미래에 일어날 일을 예측하는 그런 모델링을 할 때 데이터 마이닝이라는 말을 쓴다. 데이터 마이닝은 여러 곳에 사용이 된다. 예를 들면 Customer relationship management이라는 대부분의 기업에서 도입해서 쓰고 있다. 본인 기업의 고객들의 성향을 clustering 해서 비슷한 성향의 그룹들을 그룹핑 한 다음에 그 그룹이 좋아할 만한 상품을 타겟 마케팅하는 용도로 쓰는 것이 CRM 솔루션이고, 매출을 늘리는데 바로 연결이 되기 때문에 어떻게 분석하는가 등이 중요하다.

Healthcare라는 산업은 굉장히 크고, Inefficiency가 존재한다. 모든 산업들이 inefficiency가 존재하지만, 빅데이터를 통해 계산할 수 있는 비효율성이 존재하고, 그 비효율성 중에 1%만 개선을 해도, 엄청난 cost-saving한 효과를 나타낸다. MD Anderson과 IBM Watson이 조사한 결과에 따르면, 일주일에 한번 정도 72%의 의사들이 불필요한 오더를 내리고, 가중 평균 8.4%의 임상 오류가 발생한다. 이러한 비효율성을 빅데이터나 인공지능 기술을 이용해서 개선을 할 수 있다고 하면 굉장히 큰 이득이다. 암은 단일 약물로서는 지속적인 효과를 보기 어려운 질환이다. 따라서 여러 길목을 동시에 막을 수 있는 다양한 약물이 필요한데, 우리는 아직도 더 많은 약물을 더 빨리 개발해야 된다. 하지만 어려운 점은 FDA에서 승인을 받는 약물의 개수는 해마다 증가를 하지 않고 비슷하거나 감소하는 추세이지만 R&D 비용은 계속 증가하고 있다. 따라서 이렇게 계속 지속이

되면 신약개발이라는 sector가 나중에는 investment에 대해서 return이 없는 sector가 될 수 있다. 그렇기 때문에 기술적인 breakthrough가 필요하고, 그 breakthrough가 인공지능이 기여할 수 있을 것이라는 관점에서 연구를 하고 있다. 전통적인 신약개발은 단백질 수준에서부터 시작을 하는데 길이 너무 멀어서, 연구를 중간에 세포 수준에서부터 시작을 한다.

AstraZeneca-Sanger Challenge는 제공된 데이터를 이용해서 어떤 암 환자에게 어떤 약물의 조합이 가장 시너지가 좋을 지 예측하는 기계학습 모델을 만드는 것이다. 처음에는 기계적으로 컴퓨터만을 이용해서 문제를 풀려고 했으나, 한번 사이클을 도는데 약 20일이 소요가 되어서, 아마존에서 기계 50대를 빌려서 이 기계들이 유기적으로 task를 수행할 수 있도록 하였고, 의미 있는 성적을 거뒀다고 한다. 데이터에 context가 낮설긴 하지만 결국은 기계학습 모델을 학습 시키는 것이기 때문에, abstraction layer를 만들어서 데이터를 abstraction 하고 나면 그 단계 이후로부터는 똑같은 기계학습 패러다임을 적용하면 된다. 따라서 도메인이 달라진다고 하여 너무 지나치게 겁 먹을 필요는 없다. Raw Data를 abstraction해서 중요한 feature들을 골라내는 과정에서만 도메인 전문가들과 소통을 하면서 어떠한 접근 방법이 유의한지를 확인하고 feature를 abstraction layer에 올리고 나면 그 다음부터는 같다.

사실 많은 인공지능 기술들에 대해서 세미나를 들어봤지만, 이번 세미나가 특히나 나에게 있어서는 새로웠던 것 같다. 인공지능이 과연 어떻게 신약개발에 있어서 이용이 될 지 전혀 몰랐기 때문에 비록 깊게는 아니지만 대략적으로 알 수 있어서 좋았다. 또한 처음 단계에서만 협업을 하고 abstraction을 짜고 나면 그 다음부터는 우리가 하던 기계학습이랑 같다는 것도 신기했다. 사실 의료기술이라는 것이 생명을 다루는 것이기 때문에 과연 AI가 이 분야에서까지도 좋은 영향을 끼칠 수 있는지 궁금했는데, 이번 세미나를 통해서 좀 더 연구를 진행하면 충분히 좋은 성과를 낼 수 있을 것 같다.