

번역을 펴내며

이 책에 대하여

Leslie Valiant, *Probably Approximately Correct*:

Nature's Algorithms for Learning and Prospering in a Complex World,

Basic Books, New York, 2013.

밸리언트(Leslie Valiant) 교수의 이 책을 만난 건 2016년 여름이다. 알파고 충격이 그해 3월이었다. 알파고를 만든 기계 학습(machine learning)을 이해할 시각을 찾던 중 만난 책이다. 밸리언트 교수는 기계 학습 이론으로 2010년 컴퓨터과학의 노벨상이라 할 튜링상(Turing Award)을 받았다.

기계 학습 이전까지는 소프트웨어를 만들려면 한 가지 방식뿐이었다. 언어와 논리다. 컴퓨터에게 시키고 싶은 일이 있으면, 그 일을 하는 방법을 우리가 직접 글로 쓰는 것이다. 모든 세밀한 과정을 논리에 맞게 빠뜨림 없이 글로 써서 컴퓨터에 옮기면 컴퓨터는 그 글에 적힌 대로 실행에 옮긴다.

그런데 우리의 언어와 논리로 아직 그 방법을 표현할 수 없는 일들이 있다. 이세돌을 이기고 자동차를 운전하고 통역을 하고 눈치를 채는 일. 그동안 우리는 그런 일을 어떻게 하는 것인지 표현할 수 없었다. “홍시 맛이 나서 홍시 맛이 난다고 한 것이운데 어떻게 그 맛을 알았냐고 물으시면 어찌 답해야 할지 모르겠습니다.” 드라마 <대장금>의 한 장

면이다. 글로, 명확한 논리로 그런 일을 어떻게 하는지, 어찌 답해야 할지 우리는 아직 모른다. 강아지 사진인지를 판단하는 방법을 글로 적어 보시라. 그리고 그 방법대로 판단하는 과정을 밟아보시라. 틀린 답을 내기 일쑤다. 논리와 언어로는 그 방법을 잡아내지 못한다. 우리가 늘 하는 일이니 분명히 가능한 일인데, 그 방법을 일일이 표현하려 들면 난감해진다. 아직 우리가 이해하는 언어로는 구체적으로 작성할 수 없는 소프트웨어가 우리 몸에서 작동하고 있는 것이다.

기계 학습은 그런 소프트웨어를 만드는 방식이다. 단, 우리가 직접 만들지는 못하고 컴퓨터가 만들도록 한다. 컴퓨터가 학습 과정을 밟게 해서 만들어 내도록 한다. 우리가 할 일은 그 학습 방법만을 논리와 언어로 작성해서 컴퓨터에게 전달하는 것이다. 컴퓨터는 쓰인 대로 학습 과정을 진행하고 그 결과로 소프트웨어를 뱉어낸다. 물론 그 소프트웨어도 글로 쓰여 있다. 그러나 어떤 경우(특히 인공신경망 스타일인 경우) 그 소프트웨어를 우리가 이해하기는 아직 불가능하다. 너무 아래 단계의 글이기 때문이다. 마치 분자 구조로만 적혀 있는 장치랄까. 뭐가 무슨 역할을 어떻게 해서 풍미를 느끼는 혀와 코가 된 건지 상위에서 파악하기 힘든. 그런데 이 소프트웨어가 용케 홍시 맛이 있다/없다를 답한다. 우리가 원하는 일을 얼추 그럴듯하게 해내는 소프트웨어가 학습을 통해서 어쨌든 만들어지는 것이다.

기계 학습은 주입식 학습이다. 많은 정답 예시들을 주입한다. 그러면 출력으로 그런 정답을 낼 수 있는 소프트웨어가 만들어진다. 홍시를 넣은 음식을 많이 맛보게 한다. 설명은 없다. 예시들만 주입할 뿐이다. 그러면 홍시 맛이 도는 요리인지를 감별하는 함수가 만들어진다. “2일 때 3”, “-1일 때 0”을 주면 $x+1$ 함수가 만들어진다. 특수에서 보편으로 건

너뛰는 것이다. 어거지이고 거짓일 수 있다. 사실은 x^2-1 함수일 수 있다. 충분히 많은 예시를 주입한다면 어떨까? 서당개 3년이면 풍월을 읊는다는데, 서당개에게 5만 년어치의 글을 들려준다면? 기계 학습을 돌리는 컴퓨터에게 5만 년어치 책을 주입해주면 학습 결과로 얼추 글을 쓰는 서당개를 만들어 준다. 종종 엉터리 글을 쓸 수는 있지만 어느 정도 해낸다.

지금 기계 학습은 전통적인 소프트웨어 제작의 1960년대와 비슷하다. 언어로 직접 짜는 소프트웨어가 시작되던 시절. 멋진 결과들이 나오기 시작하면서 한편에서는 기계적인 계산의 세계(소프트웨어의 세계)를 이해하려는 노력이 시작되었다. 컴퓨터로 어디까지 할 수 있는지, 빨리 풀 수 있는 문제들의 경계는 어디인지. 또, 소프트웨어를 기계에 대한 명령어로만 바라볼 필요가 있는지 등. 컴퓨팅 계산 이론(*computation theory*)과 프로그래밍 언어 이론(*programming language theory*)이 발아하던 시기였다.

지금 이 그때와 유사하다. 그동안 지어내지 못했던 소프트웨어를 기계 학습으로 길어 올리며 북적이는 동요 혹은 격동. 서로들 몸을 던져 들어서는 기계 학습의 우물 혹은 정글.

그런 한편에선 그 정글이 어떤 세계인지를 이해하려고 애쓰는 사람들이 있다. 저건 뭐고 이건 뭐란다. 저건 우리보다 세지만 여기가 급소란다. 이렇게까지 가능해. 그 끝은 이렇고. 이렇게 우리를 안심시키는 이야기, 그런 과학이 있다. 독자들은 그런 과학의 한 줄기를 이 책에서 구경하게 된다.

벨리언트 교수가 그 선두였다. 그의 업적은 기계 학습을 수학으로 명확하게 정의한 것이다. 그가 만든 정의가(혹은 그 변형이) 앞으로 출현할 모든 기계 학습 알고리즘을 포섭할지는 두고 볼 일이지만 그 가능성

은 큰 듯하다.

이런 정의가 왜 중요할까? 정의해야 이해할 수 있기 때문이다. 모든 기계 학습을 아우를 튼튼한 정의가 있어야, 비로소 어떤 것이 기계 학습으로 가능할지를 알 수 있기 때문이다. 물론 정의가 없어도 새로운 기계 학습 기술은 부지런히 만들어질 것이다. 하지만 모든 기술을 아우르는 기계 학습을 정확하게 정의하지 않고는 기계 학습의 능력과 한계의 경계를 알지 못한다. 고안한 기술이 그런 지도에서 어디쯤인지, 그래서 그 쪽으로 달려가다 부닥칠 한계선을 피할 기술이 어디서부터 필요할는지 알지 못한다. 오리무중. 마치 연금술로만 머물며 헛발질한다고 할까.

컴퓨터의 원천 설계도를 선보인 튜링(Alan Turing)의 업적이 사실 그런 것이었다. 튜링은 1930년대에 기계적인 계산이 뭔지를 명확히 정의한다. 그 정의 덕분에 컴퓨터로 하는 온갖 문제 풀이 능력과 한계를 파악할 수 있었다. 그리고 그 정의가 지금까지 깨지지 않고 유지된 덕분에, 파악했던 것들이 지금까지 사실로 유지될 수 있었고, 컴퓨터 하드웨어와 소프트웨어 분야가 서로 신경 쓰지 않고 전속력으로 발전할 수 있었다.

어떤 독자는 궁금해할 것 같다. 이 책을 읽으면 더 좋은 인공지능을 만들 수 있는 걸까? 당장은 그렇지 않다. 계산 이론을 안다고 더 빠른 길 찾기 알고리즘을 만드는 데 도움 되지 않고, 생리학을 안다고 달리기 금메달리스트를 만드는 데 즉각 도움 되지는 않는 것과 같다.

대신에, 이 책으로는 기계 학습에서 하는 일이 결국 뭐다,라는 것을 어느 정도 이해하게 된다. 기계 학습의 정의다. 그래서 기계 학습이 어디까지 나아가게 될지. 지금 이상으로 어떤 것을 상상할 수 있는지. 학습이 만들 인공지능과 인간지능의 한계는 어디까지일지. 현재의 너머까지를 바라보는 긴 안목. 그런 대담함을 자극 받는다. 당장의 정황에 즉각

반응하는 기민함이 익숙한 독자에겐 아마도 소홀했던 시각일 것이다.

이 지점이 이 책을 번역하려던 동기의 시작이었다.

번역 동기

두 가지 정도가 있었다.

하나, 기계 학습(machine learning)을 보다 근본적으로 이해할 수 있는 시점을 소개하기다. 지금 모두가 놀라는 인공지능을 만드는 주인공 기술이 기계 학습이다. 이 기계 학습의 바다로 많은 배들이 나서고 있다. 만선의 깃발을 날리며 돌아오는 배도, 빈손으로 귀항하는 배도, 폭풍우에 뒤집히는 배도 있을 것이다. 모두가 동요하는 이런 야단법석의 세계. 그 아래를 질서 있게 바라보는 과학이 있다. 흥분과 우려로 충혈된 눈빛을 차분히 가라앉혀 줄 과학. 널리 알리는 것이 어떨까 싶었다.

다른 한 동기로는, 학습 과학(learning science)이라는 새로운 과학을 소개하기다. 학습 과학은 기계 학습만이 아니고 전통적인 과학이 다루지 못한 자연 현상을 다룰 새로운 방법이 될 수 있다,는 소식이다. 중요한 흐름으로 맹렬해질 수 있을 터, 지금 소리소문내는 것도 시의적절하지 않을까 싶었다.

그래서 아무쪼록 이 번역서가 다음의 역할로 독자들에게 다가가는 것을 상상한다. 하나는, 일종의 익스트림 투어 코스 입구에 놓인 가이드북이다. '4차 산업혁명'의 엔진이라는 컴퓨터과학이 인공지능과 기계 학습 관련해서 물밑에서 다지고 있는 기초공사의 한구석. 그곳을 구경하는 투어 코스에 함께하는 안내서 정도. 또 하나는, 일종의 저널리즘 역할이다. 학습 과학이 인공지능을 넘어 그 이상을 펼치게 될 분야임을 예측하는 흥미로운 기별, 그 속 내용을 짚어보는 심야 다큐멘터리 정도다. 마지막

으로, 근거 있는 논의를 위한 참고서 역할이다. 인공지능 관련 논의가 신비감에 기대어 부유하지 않도록 과학적인 근거로 내실을 심는 채비다.

번역 방법에 대하여

두 의견에 힘입어 번역에 임했다.

- 번역은 도착어 문화권에만 일방적으로 공헌하는 게 아니다. 번역본과 원서를 모두 합할 때 오리지널 텍스트만 가지고는 도달할 수 없던 수준의 이해를 달성하게 된다.
- 번역은 출발어로 쓰인 텍스트를 도착어로 최대한 정확하게 해석하고 풀어내는 데 머물지 않는다. 번역하는 중에 도착어에는 어휘가 만들어지고 정비되고 문법마저 변형되는 변혁이 일어난다.

이런 생각이 따라 들었다. 직역은 이제 기계가 할 일이다. 사람인 나는 다른 수준의 번역을 해야 한다. 가능해야지 않을까. 원본을 읽는 사람들보다 번역본을 읽는 사람들이 보다 쉽게 잘 이해할 수 있는 번역.

그래서 번역 시작 전에 저자의 다짐을 받아두었다. 번역 중에 조금이라도 의문스러운 부분에 대해서는 자세히 캐물으며 귀찮게 할 듯한데 괜찮겠는지. 나와 저자의 세부 전공이 다르다 보니 의문 가는 대목이 있었다. 저자와 친분이 있을 것 같은 그 동네 동료 교수에게 도움을 청했다. 나를 저자에게 소개시켜 달라. 그렇게 소개받으며 일면식도 없던 저자와 가까워질 수 있었다. 저자는 나의 계획을 기꺼이 환영해 주었다.

번역에는 원 책에는 없던 것이 추가되었다. 원 책이 2013년에 출판된 이후 깊은 신경망 *deep neural net*, 딥뉴럴넷을 이용한 기계 학습 기술은 알파

고 같은 놀라운 능력을 다양한 분야에서 보여주고 있다. 새로운 이 기술들을 저자의 이론으로 어떻게 바라볼 수 있는지 물었다. 그 내용을 받아 9장에 첨가하였다.

이 외에 세 가지를 추가하였다. 요약, 각주, 소제목이다. 각 장마다 맨 앞에 그 장의 요점을 정리해서 넣었다. 전경을 미리 조망하고 숲속으로 들어서도록 돕는 안내판이다. 그리고 종종 해설을 각주로 첨가했다. 저자의 호흡이 가파르다고 판단되면 최대한 쉽게 배경을 해설하는 데 각주를 이용했다. 모든 각주는 번역자의 각주다. 내용이 절로 구분되지 않고 너무 길어진 경우 소제목을 달아 부분 부분의 요점을 미리 드러냈다.

번역은 다음을 유지하려고 했다.

- 전문 용어는 최대한 쉽게 풀어 썼다. 전문 개념의 핵심을 쉽게 전달 하도록, 최대한 쉬운 일상어로 풀어 번역했다. 학술은 학술의 언어들 - 우리로서는 소리로만 읽을 원어나 한문을 - 사용해야만 정확하고 정밀하고 경제적이란 데 동의하지 않는다. 아무리 정교한 전문 지식이라도 쉬운 일상어로 짧고 정밀하게 전달될 수 있다고 믿는다. 시에서 평범한 언어로 밀도 있게 전달되는 정밀한 느낌을 겪으며 믿게 된 바이기도 하다. 전문 용어는 항상 해당 우리말 아래에 이탤릭체 첨자_{script}로 원문을 덧붙였다.
- 가능하면 두괄식 글이 되도록 했다. 각 단락의 요점을 간략하게 정리해서 첫 문장으로 놓는 방식이다. 글을 빨리 수월하게 읽을 수 있게 하는 장치다. 단락의 첫 문장들만 주욱 읽어도 내용을 파악할 수 있다. 두괄식을 위해서 단락에 첫 문장을 추가하기도 했고, 하나의 단락이 하나의 요점만 가지도록 단락들을 쪼개거나 합치고 문장을 재배치하기도 했다.

읽기 안내

차례를 보고 관심 가는 곳으로 곧바로 들어서도 별 문제없을 것이다.

아니면 이런 식도 가능하다. 전체 조망을 반복해서 훑는데, 매번 더 자세히 읽는 방식이다. 우선 저자가 마련한 줄거리 장을 읽는다. 더 궁금하면 각 장의 요약들만 읽는다. 더 궁금한 장이 있으면 그 장으로 들어가서 우선 각 단락마다 첫 문장만 읽는다. 첫 문장들만으론 간지럽다면 단락 안쪽을 모두 읽는다. 그러다 이것 보라 싶으면 책 모든 문장을 읽는다. 그리고도 간지럽이 남으면 해당 참고 자료들로 옮겨간다.

저자의 참고 자료 인용 부호는 각주와 구별하기 위해 꺾쇠 괄호 안에 알파벳^[2]으로 표시하였다. 용어 설명에는 원본 것에 더해서 몇 가지 번역 용어들의 설명을 더했다. 찾아보기는 통상적인 역할뿐 아니라 용어들의 번역 관계를 한눈에 보여주는 역할도 한다. 영어 기준 찾아보기에는 영문 아래 우리말 번역들이 모여있고, 우리말 기준 찾아보기에는 우리말 아래 해당 영문들이 모여 있다.

알고리즘기계 너머 에코리즘기계

디덕기계 너머 인덕기계

확실히 다 맞는 계산 너머 얼추 거의 맞는 계산

이 모든 걸 과학으로 펼치는 지능기계의 언덕

이 책의 슬로건으로 정리해 봤다. 위 슬로건에 대한 궁금증이 수월하게 풀리는 책읽기가 되기를 아무쪼록 빌어본다. 계획한 바, 원서보다 쉬운 번역본이 가능하리라고 나섰는데 결과가 과연 그런지 모르겠다. 독자 여러분의 평가가 남았다.