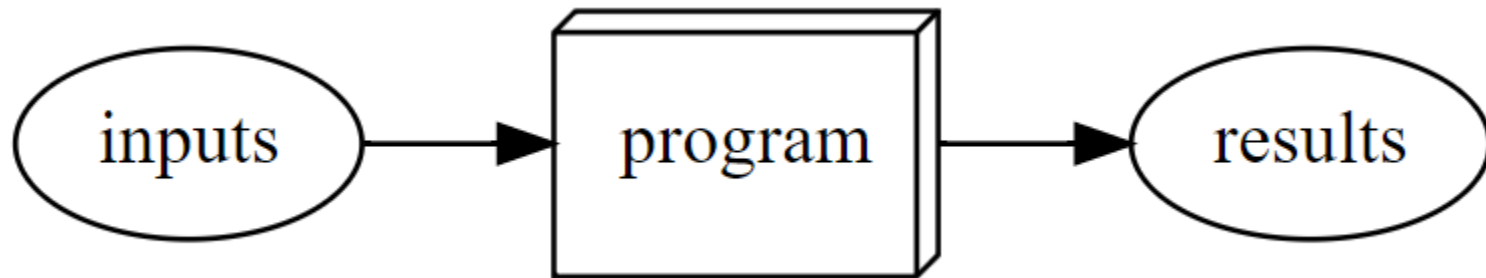


Fast ai: chapter 1.
what is Machine Learning? ~ Limitations Inherent To Machine Learning

정주호

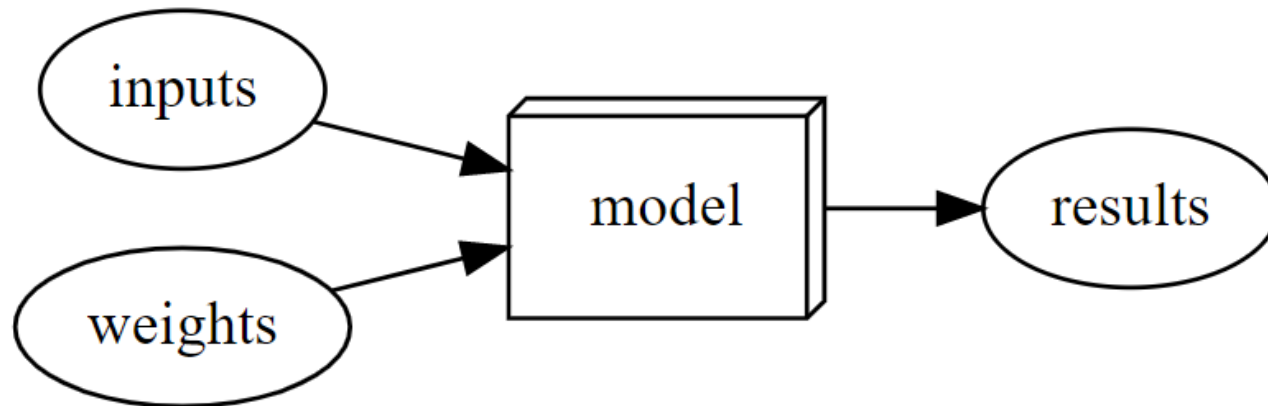
What is Machine Learning?

- 이 섹션에서는 머신러닝이 어떻게 동작하는지 보여주고, 어떤 key concepts 과 원작들에서 어떻게 표현하고자 했는지를 역추적 해보고자 함
- 머신러닝은 일반적인 프로그래밍과 마찬가지로 컴퓨터가 특정 작업을 완료하도록 하는 방법 중 하나이며 각 단계를 기록할 수 있어야 함
- 간단한 과제 같은 경우 우리가 필요로 할 단계들에 대해 생각하고 코드로 변환하여 원하는 결과를 얻을 수 있음

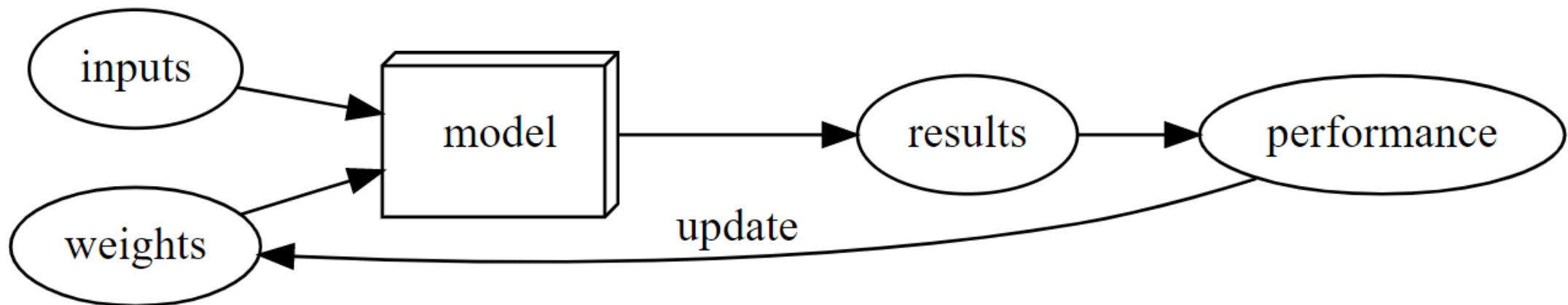


- 하지만, 우리가 시각적으로 물체를 인식하는 것은 뇌에서 어떠한 처리 과정을 통해 인식하는지 확인할 수 없기 때문에 코드로 변환하기 어려움
- 컴퓨팅이 막 시작된 1949년에 IBM의 아서 사무엘(Arthur Samuel)이라는 연구원이 컴퓨터로 작업을 수행하기 위한 다른 방법을 연구하기 시작했는데, 그가 이것을 'Machine Learning' 이라고 불렀음
- 그의 기본적인 생각은 문제를 해결하는 데 필요한 정확한 단계를 컴퓨터에게 입력하는 것 대신에, 풀어야 할 문제의 예를 보여주고, 문제를 해결하는 방법 자체를 알아내도록 하는 것임
- 사무엘의 저서에는 여러가지 powerful한 concepts 들이 포함되어 있음
- weight assignment에 대한 아이디어
- 모든 weight assignment에는 어떤 실제 성능이 존재해야 함
- 그 성능을 자동으로 시험할 수 있는 수단이 있어야 한다는 요구
- weight assignment를 변경하여 성능을 개선하기 위한 메커니즘의 필요성

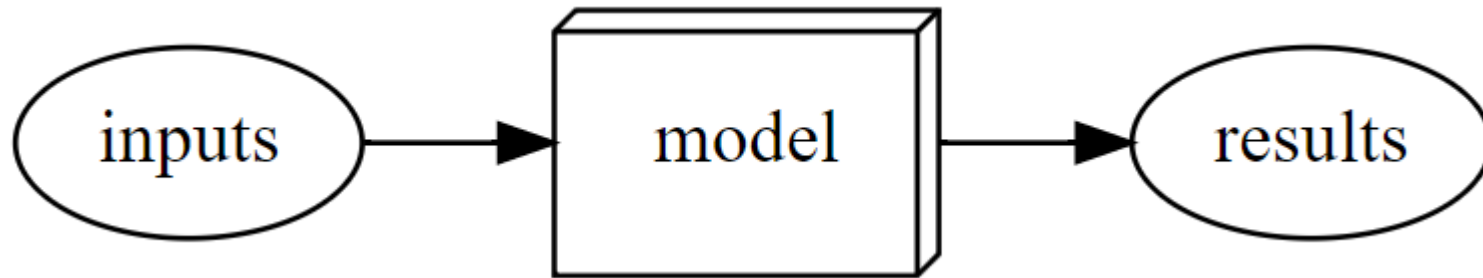
- 각 개념들이 실제로 어떻게 조화를 이루는지 이해하기 위해선 하나씩 확인해봐야 하는데, 우선 weight assignment에 대해 이해할 필요가 있음
- weight는 변수에 불과하며 weight assignment는 해당 변수에 대한 특정 값을 선택하는 것임
- 예를 들어, 프로그램의 입력에서 영상 픽셀을 입력한 후 그 결과를 “개”로 반환하는 등 결과를 생성하기 위해 처리하는 값임
- 다른 종류로의 입력을 넣음으로써 프로그램에 영향을 미치기 때문에 기존의 구조도를 변경할 것임



- 실제 성능면에서 현재 weight assignment에 대한 효과를 테스트할 수 있는 자동적인 수단이 필요하며, 두 모델을 대결하도록 설정하고 어떤 모델이 보통 이기는지를 확인하면 두 모델의 성능을 자동으로 테스트할 수 있음
- 성능을 극대화하기 위해선 weight assignment를 변경할 수 있는 메커니즘이 필요한데, 두 모델의 비교를 통해 성능이 더 좋은 모델과 그렇지 않은 모델의 weight 차이를 확인할 수 있고, 성능이 더 좋은 방향으로 weight를 조금씩 조정할 수 있음
- 이러한 weight 조정도 자동이 되면 학습은 완전히 자동으로 진행될 수 있음
- 수동적으로 weight를 조정하여 모델을 개선하는 대신, 자동화 메커니즘을 통해 성능에 기반한 weight 조정에 의존함



- 최종적으로 모델이 학습되고 가장 좋은 weight assignment가 되면 weight도 model의 일부분이기 때문에 우리가 처음에 언급했던 기초적인 프로그램 단계와 동일함

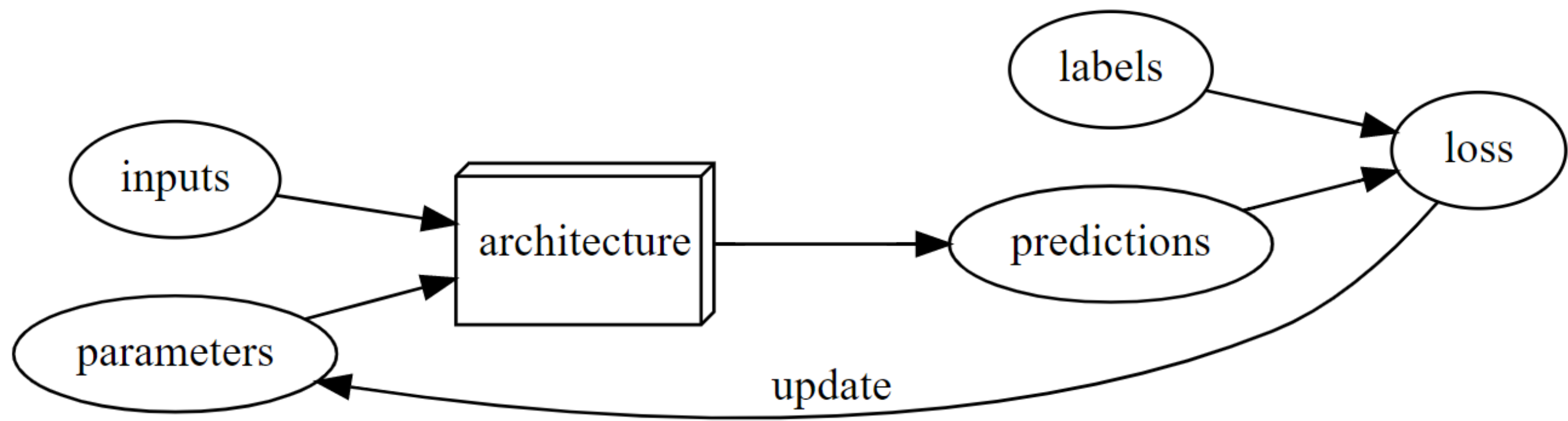


What is a Neural Network?

- 우리가 원하는 것은 어떤 한 종류의 Method를 통해 단순히 weight를 조정함으로써 여러가지의 문제를 유연하게 해결할 수 있는 것을 원함
- 그것이 바로 Neural Network임
- Neural Network를 수학적 함수로 본다면, weight에 따라 극히 유연한 함수임
- 그렇기 때문에, 우리는 Neural Network를 훈련시키는 과정, 즉 좋은 weight assignment를 찾는 과정에 더욱 집중할 수 있을 것임
- 신경망의 가중치를 갱신하고 어떤 작업에서도 개선 시킬 수 있는 방법이 필요로 한데 stochastic gradient descent (SGD)를 통해 간편하게 할 수 있음
- SGD를 통해 weight 값을 자동으로 찾을 수 있음

A Bit of Deep Learning Jargon (딥러닝 전문용어)

- 사무엘은 1960년대에 일을 했기 때문에, 우리가 현재 논의하고 있는 모든 용어는 현대 딥러닝 용어를 기준으로 함
- 모델의 기능적 형태를 architecture라고 함(모델은 아키텍처가 아님)
- weight를 parameters라고 부름
- predictions은 레이블을 포함하지 않는 데이터인 독립적인 변수로부터 계산 됨
- 그 모델의 결과를 predictions라고 함
- 성능의 측정을 loss라고 함
- loss는 predictions뿐만 아니라 정확한 라벨에 따라 달라짐



Limitations Inherent To Machine Learning

- 데이터가 없으면 모델을 만들 수 없음
- 모델은 학습하는데 사용되는 입력 데이터에서 보이는 패턴에서만 작동되는 것을 배울 수 있음
- 이러한 학습 방식들은 actions 을 하는게 아니라, predictions만 함
- 입력 데이터만 필요한 게 아니라, 데이터에 대한 라벨도 필요함
- 일반적으로, 대부분의 사람들은 충분한 데이터가 없다고 말을 하는데, 실제로는 라벨이 있는 데이터가 충분하지 않음을 의미하며 이 책에서는 라벨링 접근법에 대하여 많은 논의가 있을 것임
- 또한, 모델이 환경과 상호작용하는 방법을 고려해야 함
- 긍정적인 데이터를 많이 학습하면 모델은 긍정적인 데이터에 편향되어 결과를 제공하며 해당 주제에 대해선 뒤에 더 자세히 논의 함