<Lec 02>

predicting exam score : regression

|  |  |
| --- | --- |
| x(시간) | y(점수) |
| 10 | 90 |
| 9 | 80 |
| 3 | 50 |
| 2 | 30 |

이 확정된 자료셋을 갖고 학습을 시킴(training) -> if) x=7 집어넣었을 때 y가 얼마일까? regression해서 65정도 ‘예측’ --> linear regression

<training data>

|  |  |
| --- | --- |
| x(관측하는 대상) | y(예측해야 될 대상) |
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |

--> DRW0000411851ef꼴

(Linear) Hypothesis

데이터에 대해 그 데이터에 맞게 linear한 선을 긋는다.

어떤 선이 그 데이터의 경향성을 반영한 선일까?

-> DRW0000411851f1

1. 일차방정식 형태로 나타낼 것이다.
2. 어떤 선이 가장 맞을까?

실제 데이터와 가설한 데이터의 추 선 사이의 거리 측정. -> Cost(Loss) Function

DRW0000411851f3로 계산 X -> DRW0000411851f5차이가 작을 때보다 클 때 페널티를 많이 줌.

DRW0000411851f7(DRW0000411851f9 는 각각 함수의 DRW0000411851fb번째를 의미함. DRW0000411851fd은 데이터 개수)

Cost Function

DRW0000411851ff-> 위의 cost 식에 대입

DRW000041185201

goal : Minimize cost(DRW000041185203 ) cost를 가장 작게 하는 W,b를 구하자.