

1. 수백만 개의 특성을 가진 훈련 세트에서는 어떤 선형 회귀 알고리즘을 사용할 수 있는가?

: 확률적 경사하강법. 전체 데이터를 한 번에 처리하지 않고 미니 배치 단위로 업데이트하기 때문에 메모리 효율이 높고 고차원 데이터 셋에 적합

2. 훈련 세트에 있는 특성들이 각기 아주 다른 스케일을 가지고 있을 때, 이런 데이터에 잘 작동하지 않는 알고리즘은 무엇인가? 그 이유는 무엇이고, 이 문제를 어떻게 해결할 수 있는가?

: 경사하강법 기반 알고리즘. 비용 등고선이 찌그러진 모양으로 그려지기 때문에 비용함수의 최솟값으로 보다 먼 거리를 지나가서 수렴 속도가 느려질 수 있다. 특성들의 스케일을 통일시키면 비용함수의 최솟값으로 최단 거리로 수렴하기 때문에 (비용 등고선이 원 모양으로 그려짐) 표준화, 정규화 등 특성 스케일 조정이 필요하다.

3. 경사 하강법으로 로지스틱 회귀 모델을 훈련시킬 때 지역 최솟값에 갇힐 가능성이 있는가?

: 로지스틱 회귀의 비용함수는 convex이기 때문에 지역 최솟값에 갇힐 가능성이 없다.

4. 충분히 오랫동안 실행하면 모든 경사 하강법 알고리즘이 같은 모델을 만들 것인가?

: convex라면 대개 동일한 전역 최적점에 도달할 수 있지만 convex가 아닌 경우에는 학습률, 배치 크기 등 여러 요인으로 인해 항상 동일한 모델을 만들지는 않는다.

5. 검증 오차가 상승하면 미니배치 경사 하강법을 즉시 중단하는 것이 좋은가?

: 미니 배치 경사 하강법에서는 일부 데이터만 보고 학습하기 때문에 일시적으로 검증 오차가 상승할 수 있다. 따라서 검증 오차가 상승했다고 즉시 중단하는 것이 아니라 검증 오차가 여러 번 연속 상승하고 개선되지 않을 때 학습을 멈추는 조기종료를 사용하는 것이 좋다.

6. 어떤 경사 하강법 알고리즘이 가장 빠르게 최적 솔루션의 주변에 도달하는가?

: 확률적 경사하강법. 배치 경사하강법은 에포크마다 한 번 그래디언트를 계산하여 파라미터 조정해서 속도가 가장 느리고 가장 빠른 것은 확률적 경사하강법이다.

7. 다항 회귀를 사용했을 때 학습 곡선을 보니 훈련 오차와 검증 오차 사이에 간격이 크다. 이때 이 문제는 무엇인가? 해결 방법에는 무엇이 있는가?

: 데이터의 차수와 모델의 차수가 달라 큰 오차가 생길 수 있다. 모델의 차수가 더 낮은 경우에는 underfitting이 일어나므로 차수를 더 높이고, 모델의 차수가 더 높은 경우에는 overfitting이 일어나므로 차수를 낮춰야한다.

8. 릿지 회귀를 사용했을 때 훈련 오차와 검증 오차가 거의 비슷하고 둘 다 높았습니다. 이 모델에는 높은 편향이 문제인가? 아니면 높은 분산이 문제인가? 이때, 규제 하이퍼파라미터 알파값을 어떻게 해야하는가?

: 높은 편향이 문제이다. 따라서 알파 값을 낮추어 규제 강도를 줄여야한다.

9. 사진을 낮과 밤, 실내와 실외로 분류하려고 할 때, 두 개의 로지스틱 회귀 분류기를 만들어야 하는가? 하나의 소프트맥스 회귀 분류기를 만들어야 하는가?

: 두 개의 로지스틱 회귀 분류기를 만들어야한다. 왜냐하면 사진을 낮/밤, 실내/실외로 나눈다면 두 가지의 독립적인 속성으로 구분하는 것인데 하나의 소프트맥스 회귀 분류기를 만들면 낮/밤/실내/실외로 4가지의 선택지가 하나의 속성이 되기 때문이다.