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

* 확률적 경사 하강법. 매우 큰 훈련 세트를 다룰 수 있음. 학습 과정이 매우 빠르며 파라미터 조정이 불안정 할 수 있기 때문에 지역 최소값에 상대적으로 덜 민감함.

1. 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 알고리즘. 특성들의 스케일이 다를 경우 비용 등고선이 타원 모양 또는 찌그러진 모양으로 그려지게 되어 비용 함수의 최소값으로 보다 먼 거리를 지나감. 특성들의 스케일을 통일시켜 이 문제를 해결할 수 있음.

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

* 로지스틱 회귀의 비용 함수는 볼록 함수(convex function)이므로, 지역 최솟값에 갇힐 위험이 없다.

1. 충분히 오랫동안 실행하면 모든 경사 하강법 알고리즘이 같은 모델을 만들 것인 가?

* 학습률, 초기화, 배치 방식에 따라 결과가 달라질 수 있기 때문에 아니다.

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

* 즉시 중단하는 것 보다는, 검증셋에 대한 비용 함수의 곡선이 진동이 발생할 수 있기에 검증 손실이 한동안 최솟값보다 높게 유지될 때 훈련을 멈추고 기억해둔 최적의 모델을 사용하는 것이 좋다.

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

* 미니 배치 경사하강법. 배치 크기를 어느 정도 크게 하면 확률적 경사하강법보다 파라미터의 움직임이 덜 불규칙적이게 되며, 배치 경사하강법보다 빠르게 학습한다.

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

* 과대적합이 발생하였다. 훈련 데이터를 추가하거나 릿지회귀, 라쏘회귀 등을 적용한다.

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

* 높은 편향이 문제이다. 알파값을 낮춰서 규제를 완화시켜야 한다.

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

* 낮과 밤, 실내와 실외는 이진 분류이기 때문에 두 개의 로지스틱 회귀 분류기를 만들어야 한다.