4주차 과제\_2023170855최정연

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

확률적 경사 하강법(SGD), 미니배치 경사 하강법, 또는 희소 데이터를 처리할 수 있는 알고리즘(Lasso 회귀 등)을 사용할 수 있다.

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?

거리 기반 알고리즘(예: KNN, SVM)은 잘 작동하지 않는다. 이유는 스케일 차이가 크면 거리 계산이 왜곡되기 때문이다. 해결 방법은 특성 스케일링(표준화 또는 정규화)이다.

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

아니다. 로지스틱 회귀 비용 함수는 볼록(convex) 함수이므로 지역 최솟값에 갇히지 않는다.

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

아니다. 학습률, 초기화, 배치 구성 등에 따라 다른 경로를 가질 수 있으므로 결과가 다를 수 있다.

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

그렇다. 이는 과대적합을 방지하기 위한 조기 종료(early stopping) 전략이다.

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

모멘텀을 사용하는 경사 하강법(예: Nesterov Accelerated Gradient)이 가장 빠르게 최적점 근처에 도달한다.

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

과대적합 문제이다. 해결 방법은 더 많은 훈련 데이터 확보, 모델 단순화 또는 규제(릿지, 라쏘) 추가이다.

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. 알파 값을 줄여서 규제를 완화해야 한다.

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

하나의 소프트맥스 회귀 분류기를 사용하는 것이 적절하다. 이는 다중 클래스 문제로 볼 수 있기 때문이다.