

Cs231n Lecture 2 Summary

Image Classification

1. Image Classification

2. 이미지 비교 함수

- Nearest Neighbor

- K-NN

3. Hyper parameters

4. Linear Classification

1. Image Classification

이미지 분류는 컴퓨터 비전에서 가장 중요한 작업이다. 이미지 분류란, 컴퓨터가 어떤 사진을 보았을 때 그 사진이 속하는 카테고리(라벨)를 찾는 것이다.

듣기에는 굉장히 간단해 보이지만 컴퓨터에게 이미지는 단지 숫자의 집합으로밖에 보이지 않기 때문에 굉장히 까다로운 작업이다. 이것을 의미론적인 차이(*Semantic gap*)이라고 한다.

그리고 이미지는 사진 속의 구도, 자세, 조명 등에 따라 픽셀이 쉽게 달라진다. 우리가 만들 알고리즘은 이러한 변화에도 불구하고 항상 같은 객체임을 알아내야 한다.

그래서 사람들이 생각해 낸 방법 중 하나는 데이터 중심 접근방법(*Data-Driven Approach*)이다. 직접 규칙을 써내려가는 것 대신 무수히 많은 데이터를 수집한 후 알고리즘을 이용해 사진을 분류하는 것이다.

2. 이미지 비교 함수

첫 번째 함수는 Nearest Neighbor 이다. 즉 '가까운 이웃 찾기' 라는 방법으로, 사진들을 모두 외운 뒤 다른 사진들을 보고 지금까지 알고 있던 사진과 비교하여 가장 비슷한 것을 찾아내는 방식이다. 데이터셋은 CIFAR-10 이라는 데이터셋을 이용한다.

사진이 비슷한지 알아내는 방법으로는 L1 distance 를 사용한다. 테스트 이미지의 숫자에서 train 이미지의 숫자를 뺀 후 절댓값을 씌워 모두 더한 값이다. 하지만 이 알고리즘은 훈련 시간은 빠르지만 예측 시간이 오래 걸린다는 단점이 있다. 그래서 고안된 방법이 K-Nearest Neighbor 이다. 가장 가까운 사진을 찾는 것은 같지만 k 값을 도입하여 주변 k 개의 값을 보았을 때 가장 가까운 것을 정답으로 간주한다. 이때 L2 distance 를 사용하는데, L2 distance 는 Euclidian distance 라고도 불린다.

3. Hyper parameters

하이퍼 파라미터는 우리가 기계에게 직접 정하여 알려줘야 하는 값을 말한다. 이 하이퍼 파라미터를 제대로 맞추기 위해서는 데이터를 test, train, validation 으로 구분해야 한다. Train 셋에서 훈련된 값을 validation 으로 검증하고 test 는 마지막 한 번만 적용하는 것이다. 또는 cross-validation 이라는 방법도 있는데, 이는 데이터를 더 잘게 쪼갠 뒤 여러 값의 평균으로 하이퍼 파라미터를 설정하는 것이다. 하지만 시간이 오래 걸리므로 자주 사용되는 기술은 아니다. Knn 에서는 보통 k 를 7 로 설정하는 것이 가장 효율적이라고 한다.

하지만 지금까지 나왔던 k-nn 은 이미지 분류에서는 사용되지 않는다. 예측 시간이 오래 걸리기도 하고, 사진 사이에서의 distance 는 그리 의미있는 값이 아니기 때문이다. 더 큰 문제는 차원이 늘어질수록 필요한 데이터의 수가 기하급수적으로 많아진다.

4. Linear Classification

Linear Classification 에서는, 특정 사진의 숫자들과 가중치를 어떤 함수에 넣으면 사진을 분류하는 카테고리의 점수가 나오게 한다. 이를 매개 변수적(Paramatic) 접근 방식이라고 한다. Nearest Neighbor 과는 다르게 가중치 값만으로도 예측이 가능해 훨씬 효율적이다.

하지만 Linear Classification 의 단점은 각 카테고리마다 다양한 모습의 사진이 존재하지만 이것을 하나로 묶어 평균을 내기 때문에 한계가 존재한다.