

241115 딥러닝2

▼ 가중치 업데이트

최적의 가중치값을 찾기 위해 순전파, 역전파 과정을 계속 거치는 것

시그모이드 함수의 성질로 인래 반복할수록 0에 수렴

Gradient값이 0이 되면 w값은 변화 없음

시그모이드 함수:

기울기가 크거나 작아짐에 따라 기울기가 0이 됨

0일 때 기울기 1 (최대)

→ 미분 시 0이 되는 문제 발생

이런 문제로 인해 여러 개선 활성화함수가 존재. 그러나 완전한 기울기 소실 문제는 해결X

Optimization(경사 하강법 연장선):

Momentum

이전 단계의 속도도 함께 고려함. 진동을 줄이고 속도를 높임.

더 개선한 것이 Nesterov Momentum: 좀 더 세밀한 조정 가능

Adagrad (Adaptive Gradient Algorithm)

고정된 학습률이 아닌, 바뀌는 학습률 (유동적)

오래 학습할수록 결국 가중치 변화 줄어들고 기울기 0에 수렴

RMSprop (Root Mean Square Propagation)

현재 기울기에 더 큰 가중치 부여

최근의 기울기 더 반영, 학습률 감소 속도를 억제?(너무 빠르게 감소하지 않게)함

Adagrad 개선책

Adam (Adaptive Moment Estimation)

각 단계마다 학습률 자동 조정을 엄격히 함

초기 설정값이 생각보다 많아 감으로 지정해야 함. 하지만 default에서도 잘 됨.

딥러닝에서 많이 쓰임

▼ CNN (Convolution Neural Network)

Conv 계층 연산

풀링: 필터를 이용해 가로세로 크기를 줄임

가장자리 요소는 단 한 번 밖에 관여할 수 없음 → 제로 패딩으로 해결

최대값 풀링, 평균값 풀링이 있음.

평균값 풀링은 강력한 특징을 약하게 만드는 경향이 있어 잘 사용되지 않음
드롭아웃

임의로 선택한 노드를 off하는 방식

특정 노드에만 의존하는 경향을 줄일 수 있음

과제:

수업 자료 정리

코드 수정:

Conv 계층 7개~

FC 계층 1~3개

BatchNormalization, maxPooling, Dropout 필수 사용

Train 점수와 Test 점수 모두 65점 이상

모델 학습 시간 10분 이하