6주차 과제

Liam(황하람)

목차

1. 주제
2. 활성화 함수(시그모이드, tanh, ReLU)의 정의와 특징 분석

- 각 활성화 함수의 수학적 정의

- 출력 그래프와 특징

- 딥러닝에서 해당 활성화 함수를 사용하는 이유와 한계

1. MLP모델을 활용한 비선형 데이터셋 분류 문제 해결

- 사용된 데이터셋

- 모델구조

- 학습 결과 분석

- 선형모델과의 비교 분석

1. CNN데이터셋을 활용한 이미지 분류

- 사용된 데이터셋

- 모델 구조

- 학습 결과 분석

- MLP모델과의 비교 분석

1. 활성화 함수(시그모이드, tanh, ReLU)의 정의와 특징 분석

|  |  |
| --- | --- |
| 코드 | [코드링크](https://colab.research.google.com/drive/1p0Z_KNx9sjKg8La9Qe1FtCDX_HEkBp_P#scrollTo=8zb7W_gd4sLQ&line=4&uniqifier=1) |

* 각 활성화 함수의 수학적 정의

|  |  |  |
| --- | --- | --- |
| 활성 함수 | 함수 정의 | 출력범위 |
| 시그모이드 |  | 0 ~ 1 |
| tanh |  | -1 ~ 1 |
| ReLU |  | 0 ~ |

* 출력 그래프와 특징

텍스트, 라인, 그래프, 도표이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

시그모이드

* 음수처리 방식 : 모든 음수 값은 0~0.5 사이의 값으로 변환됨
* 포화영역 : 입력값의 절댓값이 커질수록 0 또는 1에 가까워짐
* 기울기 소실 : 포화영역으로 인해 기울기 소실 발생

tanh

* 음수처리 방식 : 모든 음수값은 -1~0 사이의 값으로 변환됨
* 포화영역 : 입력값의 절댓값이 커질수록 -1 또는 1에 가까워짐
* 기울기 소실 : 포화영역에서 기울기 소실이 발생하나, 시그모이드 보다 큰 변환범위(-1~1)로 인해 시그모이드보다는 기울기 소실측면에서 성능이 좋음

ReLU

* 음수처리 방식 : 모든 음수값을 0으로 변환됨
* 포화영역 : 양수영역에서는 포화영역이 발생하지 않으나, 음수 영역에서는 변환값이 0으로 고정되어 포화영역 발생
* 기울기 소실 : 입력값이 양수이면 기울기 소실이 발생하지 않으나, 음수값이 입력되면 기울기 소실이 발생 함

- 딥러닝에서 해당 활성화 함수를 사용하는 이유와 한계

|  |  |
| --- | --- |
| 활성화 함수를 사용하는 이유 | |
| 딥러닝 연산은 행렬의 곱셈과 덧셈으로 이뤄지기 때문에 선형적인 분류밖에 하지 못한다. 이런 연산에 활성화 함수를 통해 비선형성을 추가함으로써 비선형 문제를 해결하기 위해 사용한다. | |
| 시그모이드 | 출력이 0~1사이기 때문에 이를 확률적으로 해석하기에 유용하다. 하지만 포화영역에서 기울기 소실이 발생하고, 양수의 값만 출력하여 결과가 한쪽방향으로만 치우쳐 질 수 있다. |
| tanh | -1~1사이의 결과를 출력하기 때문에 시그모이드에서 한쪽방향으로 가중치 업데이트가 발생하는 현상을 해결할 수 있지만, 포화영역에서는 여전히 기울기 소실이 발생한다. |
| ReLU | 기울기 소실 문제를 해결하고, 효율적인 연산이 가능하기 때문에 현재 가장 대표적으로 사용되고 있는 활성화 함수이다. |

1. MLP모델을 활용한 비선형 데이터셋 분류 문제 해결

|  |  |
| --- | --- |
| 코드 | [코드링크](https://colab.research.google.com/drive/1p0Z_KNx9sjKg8La9Qe1FtCDX_HEkBp_P#scrollTo=QhGZbVPF42_Q&line=4&uniqifier=1) |

- 사용된 데이터셋

|  |  |
| --- | --- |
| **데이터셋 이름** | make\_moons |
| **전체 샘플 수** | 1000개 |
| **노이즈 정도** | 0.3 |
| **특징 수 / 분류 수** | 2개 / 2개 |
| **랜덤시드 값** | 42 |

텍스트, 스크린샷이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

- 모델구조

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 노드 층 | 입력 | 출력 | 활성화함수 | Bias여부 |
| 은닉층1 | 2 | 8 | ReLU | O |
| 은닉층2 | 8 | 8 | ReLU | O |
| 출력층 | 8 | 1 | X | O |
| 활성함수 ReLU의 선택이유 : 빠른 학습속도, 기울기 소실문제 해결 | | | | |

* 학습 결과 분석 (에포크 20회 수행)

손실감소 그래프

라인, 그래프, 도표, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

정확도 변화 그래프

라인, 그래프, 텍스트, 도표이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

결정경계 시각화

텍스트, 스크린샷이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

- 선형모델과의 비교 분석

선형모델은 직선 또는 평면으로만 데이터를 분류가능하다. 하지만 MLP모델 같은 경우 활성화 함수를 통해 그래프에 비선형성을 추가하고 이를통해 더욱 복잡하고 다양한 분류를 가능하게 한다.

1. CNN데이터셋을 활용한 이미지 분류

|  |  |
| --- | --- |
| 코드 | [코드링크](https://colab.research.google.com/drive/1p0Z_KNx9sjKg8La9Qe1FtCDX_HEkBp_P#scrollTo=Osl5E_4a43HF&line=28&uniqifier=1) |

- 사용된 데이터셋

|  |  |
| --- | --- |
| **데이터셋 이름** | mnist |
| **전체 샘플 수** | 7만개 |
| **특징 수 / 분류 수** | 28\*28 / 0~9 |

- 모델 구조

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 레이어 | 입력채널 | 출력채널 | 커널사이즈 | 패딩 | stride |
| 합성곱층1 | 1 | 16 | 5\*5 | 2 | 1 |
| 풀링층 | 16 | 16 | 2\*2 | X | X |
| 합성곱층2 | 16 | 32 | 5\*5 | 2 | 1 |
| 풀링층 | 32 | 32 | 2\*2 | X | X |
| 선형층1 | 32\*7\*7 | 128 | X | X | X |
| 선형층2 | 128 | 10 | X | X | X |

- 학습 결과 분석(총 에포크 5회 수행)

손실 추세 그래프

라인, 그래프, 텍스트, 도표이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

정확도 추세 그래프

라인, 그래프, 도표, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

- MLP모델과의 비교 분석

1. 이미지의 구조적 특징을 보존하면서 계층적으로 특징을 추출

2. 필터의 크기만큼만 가중치를 학습하면 되므로, 전체 파라미터 감소시켜 모델을 더 빠르게 훈련시키고, 더 적은 데이터로도 일반화 성능을 높임

3. 물체의 위치가 약간 바뀌어도 모델이 일관된 예측을 할 수 있게 되어, 실제 환경의 다양한 데이터를 처리하는데 더 높은 성늘을 가짐