|  |
| --- |
| 1, 2장 |
| 1. [[1,0,0,0], [0,1,0,0], [0,0,1,0]] |
| 1. 극소값을 찾아가야 하는데 변곡점에 정류할 가능성이 있음 |
| 1. X |
| 1. Cpu는 직렬연산 GPU는 병렬연산 |
| 1. [[19, 32], [18, 25]] |
| 1. (1) 2 (3) float32 |
| 1. my\_slice = train\_images[ : , -8:-15 , -8:-15 ] |
| 1. 배치, , , 배치 경사하강법 |
|  |
| 1. transpose |
| 1. 지도학습, 비지도학습, 강화학습 |
| 1. 미니배치 경사하강법을 사용 |
| 1. 층의 깊이를 깊게 하는 것 |
|  |
| 1. [[1, 4], [2, 5], [3, 6]] |
| 1. 거짓 |
| 1. 1, 4 |
| 1. 역전파 |
| 3, 4장 |
| 1. 그래프가 변동성이 너무 큰 경우 이전의 값과 새로운 값의 비율을 정하여 경사가 완만해지게 하는방법 |
| 1. 강화학습은 쓸모없어 보이는 행동을 반복함으로써 더 나은 정답을 찾아냄 |
| 1. 높은 |
| 1. Len(sequences) |
| 1. 중간층의 노드수가 내가 구분하고자 하는 클래스의 크기보다 작다면 그만큼 정보를 작은 차원에 압축해야 하기 때문에 불확실함 |
| 1. 드롭아웃, 네트워크 축소, 가중치 규제, 데이터셋 추가 |
| 1. Binary 크로스엔트로피 |
|  |
| 1. 중간층의 노드수가 내가 구분하고자 하는 클래스의 크기보다 작다면 그만큼 정보를 작은 차원에 압축해야 하기 때문에 불확실함 |
| 1. softmax함수 |
| 1. MSE |
| 1. 강화학습은 쓸모없어 보이는 행동을 반복함으로써 더 나은 정답을 찾아냄 |
| 1. 5 |
| 1. 5 |
| 1. 6 x 6 x 4 x 4 = 576 (bias가 없는 기준) |
| 1. (1) O (2) X (3) X |
| 1. 정답과 관측치 사이의 차이? |
| 1. 정규화 |
| 1. 2, 4 |
| 1. 아래거 |
| 5, 6장 |
| 1. 샘플의 크기를 줄이는 것, 풀링 |
| 1. 288 |
| 1. 2 |
| 1. O |
| 1. 1D 컨브넷, RNN |
| 1. 1 |
| 1. 가중치가 너무 많아 과대적합 발생, 풀링층 추가 |
| 1. Train\_generator가 한번 호출 될 때 몇 개의 image를 불러올지 |
| 1. X |
| 1. 연산량 감소, 이전 데이터 기억능력이 떨어짐 |
| 1. Forget\_gate = 과거의 값을 얼마나 잊어버릴지, input\_gate = 새로운 값을 얼마나 추가할지, output\_gate = 어떤값을 출력할지 |
| 1. 단어들의 의미 관계 |
|  |
| 1. 얼마나 많은층을 가져다 쓸건지, 완전연결 분류기를 새로 짤건지, |
|  |
| |  |  |  | | --- | --- | --- | | 0 | 0 | 0 | | 0 | 1 | 2 | | 0 | 4 | 5 |  |  |  |  | | --- | --- | --- | | 0 | 0 | 0 | | 2 | 3 | 0 | | 5 | 6 | 0 |  |  |  |  | | --- | --- | --- | | 0 | 4 | 5 | | 0 | 7 | 8 | | 0 | 0 | 0 |  |  |  |  | | --- | --- | --- | | 5 | 6 | 0 | | 8 | 9 | 0 | | 0 | 0 | 0 | |
| 1. X |
| 1. X |
| 7, 8장 |
|  |
| 1. 2 |
| 1. 5 |
| 1. 낮은 |
| 1. 5 |
| 1. O, O |
| 1. 3 |
| 1. Loss가 최소를 찍은 시점부터 3에포크이내에 더 이상 낮아지지 않는다면 훈련 종료 |
| 1. O |
| 1. 사용자가 직접 설정할 수 있는가 |
| 1. 잠재공간에서 랜덤벡터를 추출하고 이를 가지고 인코딩을 한 뒤, 다시 디코딩 |
| 1. O |
|  |
| 1. 과대적합, |
| 1. 5 |
| 1. Parametric method |
| 1. 4 |
| 1. 큰 |
| 1. 서로에게 |