**1,2장**

1. x=np.array([0,1,2])를 to\_categorical() 함수에 적용할 때, num\_classes가 4이면 x의 결과를 적으시오.

2. 확률적 경사하강법의 문제점을 서술하시오.

3. O,X로 답하시오. 그래디언트 부스팅은 랜덤 포레스트와 다르게 트리를 만들 때 고려해야 할 요소들의 무작위성이 없다

4. CPU와 GPU의 연산차이를 간단하게 서술하시오.

5. 다음과 같이 코드가 주어질 때 c의 결과는?

import numpy as np

a=np.array([[2, -3, 2],

[1, 0, 3]])

b=np.array([[3, 7],

[-1,-2],

[5, 6]])

c=np.dot(a,b)

6. 다음과 같이 코드가 주어질 때 (1), (2)의 결과는? (한 문제당 0.5점)

import numpy as np

a=np.array([[3.2, 5, 7, 4],

[10, 23, 54, 77],

[43, 0, -43, -1],

[0, -3, 39, 20]])

(1) : print(a.ndim)

(2) : print(a.dtype)

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

7. (60000, 28, 28)의 크기를 가지고 있는 배열인 train\_images가 있다.

위의 이미지의 부분의 7x7픽셀을 선택하기위해 음수 인덱스를 사용해서 아래 빈칸을 채우시오. (힌트: 음수 인덱스는 현재 축의 끝에서 상대적인 위치를 나타냄)

my\_slice = train\_images[ : , : , : ]

8. 다음 빈칸을 채우시오.

딥러닝 모델은 한 번에 전체 데이터 셋을 처리하는 것이 아닌 단위로 나누어 순차적으로 처리합니다. 이때 나누는 단위를 ( 　)라고 합니다.

데이터를 한 개씩 쓰면 ( ) 너무 헤매고, 전체를 쓰면 정확하지만 너무 ( ).

미니( )은 두가지 방법의 장점을 모두 얻기 위한(서로의 단점을 보완) 타협점입니다

9. 머신러닝 주요 3요소를 쓰시오

10. 아래 넘파이 배열x를 전치하는 코드를 완성하시오

x= np.[ ](x)

11. 머신러닝의 학습 방법 3가지는?

12. 파라미터값이 우리가 원하는 전역 최솟값이 아닌 지역 최솟값에 머무는

현상을 해결하는 방법은?

13. 딥러닝에서 ‘딥’의 의미를 서술하세요.

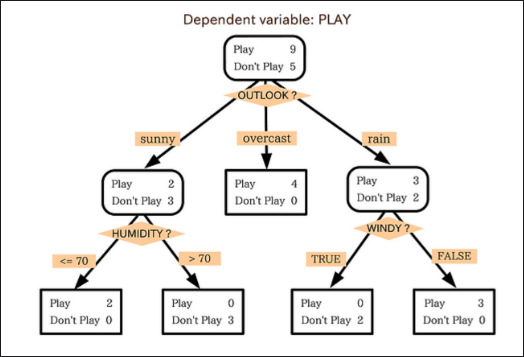
14. CUDA와 CUDA C에 대해 간단히 서술하세요.

15. 다음 코드처럼 x를 전치했을 경우의 결과는?

x = np.array([[1,2,3],[4,5,6]])

print(x.transpose())

16. 다음과 같은 결정트리에서 outlook이 sunny이고 humidity가 >70, windy가 true일 경우 play는 참인가 거짓인가?



17. Adagrad 옵티마이저의 학습이 오래 진행될 경우 기울기의 변화가 미미해져 최적화가 어려워지는 현상을 보완한 옵티마이저 두 개를 고르세요.

① RMSProp

② Momentum

③ SGD(확률적 경사 하강법)

④ AdaDelta

⑤ Adam

18. 빈칸에 알맞은 단어를 적으세요.

경사 하강법을 통해 가중치를 최적화하는 과정을 ( )라고 한다.

**3장, 4장**

1. 지수이동평균에 대하여 설명하시오(난잡해도 상관없음)

2. 강화학습과 비지도학습의 가장 중요한 차이점을 말하시오

3. 예측값이 실제값과 완전히 다를 경우에 크로스엔트로피는 ( )값이 나온다

4. 정수 시퀀스를 이진 행렬로 인코딩하는 코드이다. 빈칸을 채우시오.

import numpy as np

def vetorize\_sequences(sequences, dimension=10000):

results = np.zeros(( , dimension))

for i, sequence in enumerate(sequences):

results[i, sequence] = 1.

return results

5. 충분히 큰 중간층을 두어야 하는 이유는?

6. 과대적합을 방지하기 위한 방법들 4가지는?

7.이진 분류 문제에서 이런 스칼라 시그모이드 출력에 대해 사용할 손실 함수는 ?

8.모델을 정의하는 두가지 방법은?

9.충분히 큰 중간층을 두어야 하는 이유는?

10.단일 레이블, 다중 분류에서는 N개의 클래스에 대한 확률 분포를 출력하기 위해 사용되는 활성화 함수는?

11.회귀에서 사용하는 손실함수는?

12.강화학습과 비지도학습의 차이는?

13.과소적합은 훈련데이터의 특성을 모두 학습하지 못한 경우로 성능이 발전될 여지가 있다는 것을 의미한다. 과소적합을 방지하는 방법으로 옳은 것은?

① 훈련 데이터 모으기

② 네트워크 용량 감소

③ 가중치 규제 추가

④ 드롭아웃 계층 추가

⑤ 학습을 더 수행하기

14.보편적인 머신 러닝 작업의 흐름에서 최초의 모델을 만들기 위해서는 먼저 통계적 검정력을 달성해야 한다. 통계적 검정력을 달성하지 못한 것은?

① IMDB 감성 분류에서 60%의 정확도를 달성하였다.

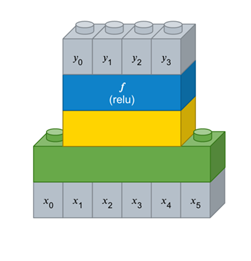
② MNIST 숫자 필기 분류에서 50%의 정확도를 달성하였다.

③ 주사위 숫자 분류에서 40%의 정확도를 달성하였다.

④ 트럼프카드 52장 분류에서 30%의 정확도를 달성하였다.

⑤ 가위바위보 이미지 분류에서 20%의 정확도를 달성하였다.

15. 다음과 같은 층이 갖는 가중치의 개수는?



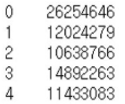
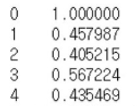
16. 다음 문제를 보고 맞으면 O, 틀리면 X를 표시하세요.

(1) 로지스틱 회귀는 분류 문제이다 ( )

(2) 회귀 문제의 예시에는 기온 예측, 질병의 감염 여부 등이 있다 ( )

(3) MSE는 예측값과 정답값 차이의 절댓값 평균을 구한 것이다 ( )

17. 지도 학습과 비지도 학습에는 이상치 탐지를 위한 모델링 방법이 존재한다. 여기서 이상치에 대한 개념을 적으시오.(키워드 : 관측치)

18. 왼쪽 이미지의 데이터에서 가운데의 코드를 적용하였더니 오른쪽 이미지의 데이터로 EMB000065981b08바뀌었다. 가운데 코드의 역할은 무엇인가?

19. 과대적합을 방지하기 위한 다양한 규제 기법에 대한 설명 중 옳은 것을 모두 고르시오.

① 간단한 모델이 복잡한 모델보다 과대적합이 될 가능성이 낮다.

② 가중치 규제는 가중치 값의 분포가 균일해지도록 하는 기법이다.

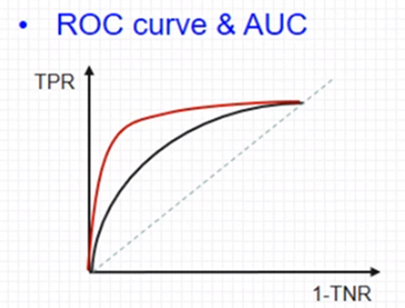
③ 가중치 규제 중 L1 규제는 가중치의 제곱합를 비용함수에 추가하고,

L2 규제는 가중치의 절댓값 합계을 비용함수에 추가한다.

④ 드롭아웃은 해당 층의 노드의 일부를 제거하는 기법이다.

⑤ 드롭아웃을 적용한 층의 출력은 그대로 다음 층으로 전달된다.

20. ROC AUC 성공지표에서 어떤 선이 더 좋은 성능을 나타내는가?



**5장, 6장**

1. 다운샘플링이 무엇이고 대표적인 다운샘플링으로는 무엇이 있을까요?

​

2. model.summary()로 신경망 구조를 살펴보았을때

------------------------------------------------------

Layer(type)                             Output Shape

------------------------------------------------------

max\_pooling2d(MaxPooling2       (None, 3, 3, 32)

------------------------------------------------------

flatten (Flatten)                       (None, ?)

------------------------------------------------------

?에 들어갈 값은?

3. 다음 설명중 틀린 것을 고르시오

​

1)RNN는 시퀀스의 개념이 존재한다

2)L2 노름의 예시로 MAE 가 있다

3)3E-3 은 0.003을 말한다

4)딥러닝의 히트맵은 CAM(class activation map)을 이용해 만든것이다

​

4 .다음 설명을 o,x로 답해주세요

 "기본 RNN구조에서 첫번쨰 타임 스텝에서는 이전 출력이 정의되지 않아

 현재 상태가 없습니다 이때 네트워크는 1인 벡터로 상태를 초기화합니다"

​

5. 다음 괄호 안에 알맞은 답을 쓰세요

CNN과 RNN을 연결하여 긴 시퀀스 처리는 1) (    )의 속도와 경량함을 2) (    )의 순서 감지 능력과 결합하여 사용한다.

​

6. 다음 보기에서 틀린 것을 고르시오

​

1) 1D컨브넷은 시간에 따른 패턴을 잘 처리한다.

​

2) 1D컨브넷은 특정 자연어 처리 같은 일부 문제에 RNN을 대신할 수 있는 빠른 모델이다.

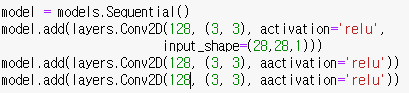
​

3) 1D컨브넷과 2D컨브넷은 전혀 다르게 구성된다.

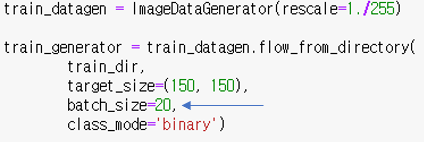
​

4) RNN은 계산비용이 많이 들고 1D컨브넷은 비용이 적게든다.

7. 다음과 같은 모델의 문제점과 해결방법은 무엇인가?



1. ImageDataGenerator를 사용하여 데이터 증식를 증식시킬 때, batch\_size가 의미하는 것은 무엇인가?



9. 다음은 컨브넷 시각화에 대한 설명이다. 다음 중 옳지 않은 것을 모두 고르시오.

1) 중간층 출력을 통한 시각화는 컨브넷과 풀링층등을 통과하면서 도출된 결과를 볼 수 있다.

2) 필터 출력을 통한 시각화는 각 계층의 필터들이 담당하는 패턴을 볼 수 있다.

3) 히트맵 시각화는 이미지처리 모델의 회귀 및 분류에 대해 판단 근거를 제시한다고 할 수 있다.

4) 컨브넷모델에서 시각화는 진행상황에 대한 진단을 위함이다.

5)필터시각화 진행 시, 경사상승법을 하는 이유는 최적화의 원리는 경사하강법과 동일하나, 특히 특이점에 대해 강한 반응성을 보인다는데에 있다.

10. 시계열 데이터 처리모델에 1D-ConvNet을 사용하는 이유와 그럼에도 불구하고,

모델 구축에 1-ConvNet만 사용하기에는 다소 어려움이 있는 이유를 각각 한문장으로 적으시오.

(1)

(2)

11. 순환 신경망 중 LSTM층의 구조를 세가지 게이트로 나눠 설명하시오.

12. 텍스트 벡터화 과정에서 원-핫 인코딩과 달리 단어 임베딩을 사용할 시 벡터에 추가로 어떤 정보를 담을 수 있는지 설명하시오.

13. VGGNet에서 3 x 3 크기의 필터를 사용하는 이유는?

( 다른 사이즈의 필터를 사용하는 것 보다 더 좋은 장점 한 가지를 쓰시오. )

14. 사전 훈련된 네트워크를 선택할 때 고려해야하는 특징 3가지는?

15. RNN은 입력과 출력의 길이를 다르게 설계할 수 있으므로 다양한 용도로 사용할 수 있다. 입력과 출력의 길이에 따라서 달라지는 RNN의 다양한 형태 3가지는?

16. 3x3 크기의 입력에 패딩을 추가하고 스트라이드 2를 사용한 3x3크기의 윈도우로 합성곱하여 추출한 패치를 보여라 (제로패딩을 사용했다)

(힌트 : 총 4개의 nxn표를 만들어서 표현해라)

입력

|  |  |  |
| --- | --- | --- |
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 | 9 |

17. 1D 컨브넷의 속도와 경량함을 RNN의 순서 감지 능력과 결합하는 한 가지 전략은 RNN을 1D 컨브넷 이전에 전처리 단계로 사용하는 것이다 (O/X)

18. 1D 컨브넷은 RNN보다 속도는 빠르지만 계산비용이 더 비싸다 (O/X)

**7,8장**

1. 다중 입력 모델, 다중 출력 모델의 훈련 방식에는 두가지가 있다. 두가지가 무엇인지 쓰시오

2.

다음 중 옳은 것을 고르시오?

ㄱ. ModelCheckpoint는 훈련하는 동안 어떤 지점에서 모델의 현재 가중치를 저장하는 것이다.

ㄴ. ReduceLROnplateau 콜백에서 학습률을 작게 하거나 크게 하면 훈련 도중 지역 최솟값에서 효과적으로 빠져나올 수 있다.

ㄷ. Early stopping 은 검증 손실이 더 이상 향상되지 않을 때 훈련을 중지하는 것이다.

① ㄱ ② ㄱ, ㄴ ③ ㄱ, ㄷ ④ ㄴ, ㄷ ⑤ ㄱ, ㄴ, ㄷ

3. 문제

모델의 성능을 최대로 끌어올리는 방법을 모두 고른 것은?

ㄱ. 깊이별 분리 합성곱

ㄴ. 하이퍼 파라미터 최적화

ㄷ. 모델 앙상블

① ㄱ ② ㄱ, ㄴ ③ ㄱ, ㄷ ④ ㄴ, ㄷ ⑤ ㄱ, ㄴ, ㄷ

4. 맞는 단어를 찾아서 쓰시오

(낮은/높은) 온도일수록 반복적이고 예상되는 텍스트를 만든다.

답:

5.뉴럴 스타일 트랜스퍼의 핵심 개념은 모든 딥러닝 알고리즘의 핵심과 동일하며 손실 함수를 정의하고 이 손실을 최소화하는 것으로 구현된다. 손실 함수에 포함되는 것을 모두 고른 것은?

ㄱ. 변위 손실

ㄴ. 콘텐츠 손실

ㄷ. 스타일 손실

① ㄱ ② ㄴ ③ ㄱ, ㄴ ④ ㄱ, ㄷ ⑤ ㄱ, ㄴ, ㄷ

6. 맞으면 o, 틀리면 x를 하시오

VAE는 실제 같은 단일 이미지를 생성할 수 있지만 구조적이고 연속적인 잠재 공간을 만들지 못한다.

답:

VAE의 파라미터는 2개의 손실 함수로 훈련한다.

답:

7.적대적 생성 신경망에 사용된 특징으로 틀린 것은?

ㄱ. 생성자가 만든 이미지는 진짜 이미지이다.

ㄴ. LeakyReLU 활성화 함수 대신 ReLu 층 사용

ㄷ. tanh 함수 대신 sigmoid 사용

ㄹ. 생성자에 드롭아웃을 사용하거나 생성자를 위해 레이블에 랜덤 노이즈를 추가한다.

① ㄱ ② ㄴ ③ ㄱ, ㄴ, ㄷ ④ ㄱ, ㄷ, ㄹ ⑤ ㄱ, ㄴ, ㄷ, ㄹ

8. 케라스 콜백의 EarlyStopping 인수중 patience=3 이라고 설정하면 나타나는 효과는?

9. 확률적 샘플링이 탐욕적 샘플링보다 엔트로피가 높다고 할 수 있다 (o,x로 답하시오)

10. 파라미터와 하이퍼 파라미터의 핵심적인 차이

11. 오토인코더의 구조를 간단히 설명하시오.

12.컨텐츠 손실에서 타깃 이미지와 생성된 이미지를 사전 훈련된 컨브넷에 주입하여 상위층의 활성화를 계산하는데 이 두 값 사이의 L2노름이 사용하기에 좋다. ( o, x로 답하시오)

13.스타일 손실에서 층의 활성화 출력으로 어떤 행렬을 사용 하는가?

14. 인셉션 모듈로 구성된 GoogleNet의 중간층에 연결된 Softmax 층은 망이 깊어질 때 발생하는 ( ) 문제를 방지하기 위해 중간 층에서도 ( )를 수행하여 가중치를 갱신하도록 해준다.

15. 일정 에포크 동안 모델이 나아지지 않으면 가중치를 저장하고 모델을 멈추고 싶다면 어떠한 콜백을 써야하는지 모두 고르시오.

1. ModelCheckpoint

2. RemoteMonitor

3. ReduceLROnPlateau

4. BaseLogger

5. EarlyStopping

16. 하이퍼파라미터는 Parametric method와 Nonparametric method 중 어떤 방법인가?

17. 뉴럴 스타일 트랜스퍼 기법에서는 타겟 이미지와 생성이미지, 그리고 스타일 참조 이미지와 생성 이미지 사이의 손실을 계산한다. 이 두 손실을 올바르게 적은 것은?

1. 이미지 손실, 스타일 손실

2. 스타일 손실, 노이즈 손실

3. 콘텐츠 손실, 노이즈 손실

4. 콘텐츠 손실, 스타일 손실

5. 콘텐츠 손실, 근손실

18. 오토인코더에서 일반적으로 중요한 특징을 추출하는 은닉층의 노드 개수는 입력 노드보다 작기 때문에 ( ) 효과를 얻을 수 있습니다.

19. GAN에서 생성자와 판별자는 각자 어디서 피드백을 받게 되는가?