

인공지능

1. (가)에 해당하는 인공지능 개념은?

- 인공지능이 인간의 지능을 초월하는 시점을 ‘인공지능의 (가)’이라 한다.
- (가)에 도달하면 인공지능은 자신보다 더 좋은 지능을 가진 또 다른 인공지능을 설계하고 생산할 수 있게 된다.
- 인공지능 과학자 겸 미래학자인 레이 커즈와일(Ray Kurzweil)은 그의 저서 『(가)이 온다』에서 2045년 그 시점이 도래할 것으로 예측했다.

- ① 분기점(branch point)
- ② 임계점(critical point)
- ③ 전환점(turning point)
- ④ 특이점(singularity)

2. 비지도 학습으로 분류되는 기법은?

- ① 랜덤 포레스트(random forest)
- ② 서포트 벡터 머신(support vector machine)
- ③ 의사결정 트리(decision tree)
- ④ 주성분 분석(principle component analysis)

3. 퍼셉트론(perceptron)에 대한 설명으로 옳지 않은 것은?

- ① 퍼셉트론의 하이퍼 매개 변수(hyper parameter)를 정하는 것은 사람이 해야 한다.
- ② 단층 퍼셉트론이 해결할 수 있는 문제는 비선형 영역을 포함한다.
- ③ 퍼셉트론은 다수의 신호를 입력으로 받아 하나의 신호를 출력한다.
- ④ 단층 퍼셉트론의 구성 요소 중 임곗값은 바이어스(bias)로 대체할 수 있다.

4. 다음 설명에 해당하는 딥러닝 알고리즘은?

- 마이크로소프트 연구팀에서 개발한 이미지 분류를 위한 딥러닝 모델이다.
- 레지듀얼 블록(residual block)을 도입하여 스킵 연결(skip connection)을 사용한다.
- 레지듀얼 블록의 출력은 입력값과 레지듀얼 함수(residual function)의 결과를 더한 형태로 표현할 수 있다.
- 이러한 구조를 통해 심층 신경망의 학습 효율을 높일 수 있다.

- ① ResNet
- ② DenseNet
- ③ MobileNet
- ④ RCNN

5. 심층 신경망에서 역전파 알고리즘을 사용할 때 발생할 수 있는 기울기 소실 문제(gradient vanishing problem)를 해결하기 위해 개발된 활성화 함수(activation function)는?

- ① 시그모이드(sigmoid)
- ② 계단 함수(step function)
- ③ 지수 선형 유닛(exponential linear unit)
- ④ 하이퍼볼릭 탄젠트(hyperbolic tangent)

6. (가) ~ (다)의 데이터 종류와 이에 대한 예를 바르게 연결한 것은?

- (가) 정형 데이터(structured data)
- (나) 비정형 데이터(unstructured data)
- (다) 반정형 데이터(semi-structured data)

- | | | |
|----------|---------|----------|
| A. SNS | B. HTML | C. VIDEO |
| D. EXCEL | E. JSON | F. RDB |
| G. XML | | |

- | (가) | (나) | (다) |
|-----------|---------|---------|
| ① D, F | A, C | B, E, G |
| ② D, F, G | A, C | B, E |
| ③ D, F | A, C, E | B, G |
| ④ F | B, E, G | A, C, D |

7. 데이터 1,000개를 대상으로 5-폴드 교차검증을 수행할 때, 한 번의 반복에서 사용되는 훈련 데이터와 검증 데이터의 개수는? (단, 각 폴드의 크기는 같다고 가정한다)

- | | 훈련 데이터 개수 | 검증 데이터 개수 |
|---|-----------|-----------|
| ① | 200 | 800 |
| ② | 500 | 500 |
| ③ | 800 | 200 |
| ④ | 1,000 | 1,000 |

8. 인공지능 생성모델에 해당하지 않는 것은?

- ① Diffusion Model
- ② DQN(Deep Q-Network)
- ③ GANs(Generative Adversarial Networks)
- ④ VAEs(Variational Autoencoders)

9. 딥러닝 기반의 인공지능 모델에서 설명 가능성(explainability)이 부족한 이유로 옳지 않은 것은?

- ① 학습 데이터의 부족
- ② 비선형적인 데이터 처리
- ③ 추상적 특징의 학습
- ④ 모델의 복잡성

10. 2017년 발표된 아실로마 인공지능 원칙(Asilomar AI principles)에 대한 설명으로 옳지 않은 것은?

- ① 인공지능 개발의 목적, 윤리, 가치 등에 대해 23가지 원칙을 제시했다.
- ② 인공지능 시스템은 반드시 인간의 직접적인 통제 없이 자율적으로 작동해야 한다.
- ③ 연구 목표는 방향성이 없는 지능이 아니라 인간에게 유용한 지능 개발이어야 한다.
- ④ 인공지능 시스템의 실패 원인을 이해할 수 있도록 투명성과 설명 가능성이 확보되어야 한다.

11. Q-러닝 알고리즘에서 다음 조건을 적용한 Q의 업데이트 값은?

- 현재 Q-값: 5
- 보상: 10
- 학습률: 0.5
- 할인계수: 0.9
- 다음 상태의 Q-값 중 최댓값: 8

- ① 6.0
- ② 8.6
- ③ 11.1
- ④ 17.2

12. 입력 채널 수가 c , 출력 채널 수가 k , 커널 크기가 $p \times p$ 인 2D 합성곱 층(convolutional layer)을 사용할 때, 학습해야 할 파라미터인 가중치의 총 개수는? (단, 바이어스 값은 없다고 가정한다)

- ① $c \times k$
- ② $p^2 \times k$
- ③ $p^2 \times c + k$
- ④ $p^2 \times c \times k$

13. 기호주의(symbolic) AI와 연결주의(connectionism) AI에 대한 설명으로 옳지 않은 것은?

- ① 기호주의 AI는 이미지, 음성 인식을 위한 딥러닝 모델을 개발하기 위해 고안되었고, 연결주의 AI는 전문가 시스템과 같은 지식 기반 시스템을 주로 구축한다.
- ② 기호주의 AI는 해석 가능성이 크다는 장점이 있으나 환경 노이즈에 약하며, 연결주의 AI는 해석 가능성은 작지만 노이즈 강인성과 범용 근사 능력이 높다.
- ③ 기호주의 AI는 논리 연산에 기반해 이산적(discrete) 문제를 해결하며, 연결주의 AI는 주로 미분 가능 함수를 이용한 학습을 통해 문제를 해결한다.
- ④ 기호주의 AI는 지식을 명시적인 생성 규칙, 논리식 등 규칙으로 표현하고, 연결주의 AI는 분산된 가중치로 지식을 암묵적으로 저장한다.

14. 트랜스포머(transformer) 모델에 대한 설명으로 옳지 않은 것은?

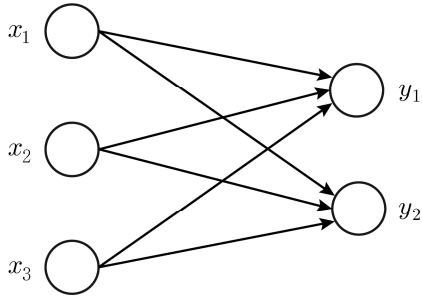
- ① 서열 데이터(sequence data)를 순차적으로 처리하는 RNN 모델과는 다르게 병렬처리가 가능하다.
- ② 위치 임베딩(positional embedding)은 입력 토큰들의 위치 정보를 모델에 제공하기 위해 사용된다.
- ③ 자기 주목(self-attention) 메커니즘은 특정 위치가 입력의 어떤 위치에 주목해야 하는지를 가중치 형태로 수치화하여 표현한다.
- ④ 주목 점수(attention score)는 query 벡터와 value 벡터의 내적을 계산한 후 소프트맥스(softmax) 함수를 적용하여 계산된다.

15. 연합학습(federated learning)의 학습 과정을 순서대로 바르게 나열한 것은?

- (가) 각 디바이스는 로컬 데이터로 모델을 학습시켜 파라미터를 업데이트한다.
- (나) 중앙 서버는 여러 디바이스로부터 받은 파라미터를 집계하여 글로벌 모델을 갱신한다.
- (다) 중앙 서버는 초기 글로벌 모델을 각 디바이스로 배포한다.
- (라) 각 디바이스는 업데이트된 파라미터만 중앙 서버로 전송한다.

- ① (가) → (다) → (라) → (나)
- ② (가) → (라) → (다) → (나)
- ③ (다) → (가) → (라) → (나)
- ④ (다) → (라) → (가) → (나)

16. 다음 퍼셉트론 신경망에서 학습이 가능한 데이터 세트는?



①

x_1	x_2	x_3	y_1	y_2
0	0	0	0	0
0	0	1	0	1
1	1	0	1	0
1	1	1	1	1

②

x_1	x_2	x_3	y_1	y_2
0	0	0	0	0
0	0	1	0	1
1	1	0	1	1
1	1	1	1	0

③

x_1	x_2	x_3	y_1	y_2
0	0	0	0	0
0	0	1	1	0
1	1	0	1	1
1	1	1	0	1

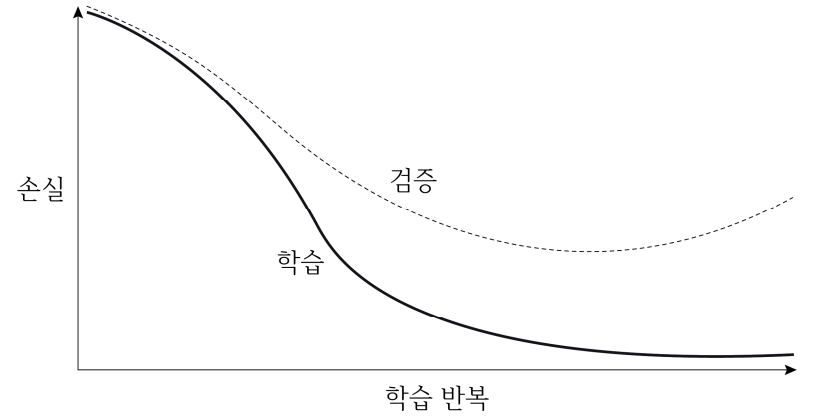
④

x_1	x_2	x_3	y_1	y_2
0	0	0	0	1
0	1	0	1	1
1	0	1	1	0
1	1	1	0	0

17. 어떤 다중 분류 신경망의 출력층 로짓(logit) 벡터가 $[2, 1, 0]$ 이다. 소프트맥스 함수를 적용하여 입력값들에 대한 확률분포를 구할 때, 가장 큰 확률값은? (단, $e \approx 2.7$, $e^2 \approx 7.4$ 이고, 소수점 셋째 자리에서 반올림한다)

- ① 0.25
② 0.33
③ 0.67
④ 0.89

18. 신경망에서 학습의 진행에 따른 학습 데이터의 손실값과 검증 데이터의 손실값이 다음과 같을 때, 이에 대한 설명으로 옳지 않은 것은?



- ① 학습 데이터를 추가하여 검증 손실값을 학습 손실값에 근접시킬 수 있다.
② 신경망 학습 모델이 학습 데이터 세트의 패턴을 적절히 학습하지 못했다는 의미이다.
③ 신경망의 은닉층 개수, 뉴런의 개수 등 학습 매개 변수가 과도하게 많으면 발생할 수 있다.
④ 신경망에서 검증 손실값이 다시 상승하는 현상을 완화하기 위하여 가중치 규제 방법이 사용될 수 있다.

19. 활성화 함수에 대한 설명으로 옳지 않은 것은?

- ① 활성화 함수는 인공신경망에서 각 뉴런이 입력을 받아 출력을 계산하는 함수이다.
② 시그모이드 함수는 S자형 함수로 로지스틱(logistic) 함수가 해당된다.
③ 소프트맥스 함수는 입력받은 값이 0 ~ 1 사이의 값으로 출력되도록 정규화하여 출력의 총합이 1이 되는 특성을 갖는다.
④ 리키렐루(Leaky ReLU) 함수는 입력값이 음수일 경우에 0을 출력하고, 그 외의 경우는 입력값을 그대로 출력한다.

20. 다음 조건에서 어떤 검사자가 진단키트를 한 번 사용하여 양성 반응을 보였을 때, 그 사람이 실제로 A 암에 걸려 있을 확률은?

- 10,000명 중 5명의 확률로 A 암이 발생한다.
○ A 암이 있는 경우 0.99의 확률로 진단키트에 양성 반응이 나오고, A 암이 없는 경우에는 0.95의 확률로 진단키트에 음성 반응이 나온다.

- ① 약 0.01
② 약 0.05
③ 약 0.95
④ 약 0.99

21. (가) ~ (다)의 효율적 파인튜닝(Parameter-Efficient Fine-tuning) 기법과 이에 대한 설명을 바르게 연결한 것은? (단, 각 기법에서는 사전 학습된 모델의 기존 가중치를 고정(freeze)한다)

- (가) LoRA
- (나) Adapter
- (다) Prefix Tuning

- A. 가중치의 변화량을 저랭크 행렬로 표현하여 이를 별도로 학습하는 기법
- B. 각 층의 어텐션 계산에 사용할 학습 가능한 접두 벡터를 추가하여 새로운 작업에 적응시키는 기법
- C. 모델 내부에 작은 추가 모듈을 삽입하여 새로운 작업에 필요한 변화만 별도로 학습하는 기법

	(가)	(나)	(다)
①	A	B	C
②	A	C	B
③	B	A	C
④	C	A	B

22. 다음 문장에서 바이그램(bigram) 언어 모델을 사용할 때, ‘나는 학교에’ 다음에 ‘간다’가 나올 확률을 계산하는 올바른 식은?

나는 학교에 간다.

- ① $P(\text{간다})$
- ② $P(\text{간다} \mid \text{나는})$
- ③ $P(\text{간다} \mid \text{학교에})$
- ④ $P(\text{간다} \mid \text{나는, 학교에})$

23. 오토인코더(autoencoder)에 대한 설명으로 옳지 않은 것은?

- ① 인코더를 통해 입력 데이터의 특징을 추출하고 디코더를 통해 원본 데이터를 재구성하는 학습을 한다.
- ② 잡음제거 오토인코더는 입력데이터와 동일한 데이터가 출력되도록 학습한다.
- ③ 오토인코더는 출력의 결과가 입력값과 동일하도록 비지도 학습을 수행한다.
- ④ 희소 오토인코더는 코딩층에서 출력값이 0이 아닌 노드의 개수를 최소화하는 오토인코더이다.

24. CLIP(Contrastive Language-Image Pretraining)에 대한 설명으로 옳지 않은 것은?

- ① 사전훈련을 통해 텍스트 데이터를 학습해 놓은 모델인 트랜스포머 언어 모델을 기반으로 한다.
- ② 이미지와 텍스트를 각각 임베딩 공간에 표현하고, 서로 대응하는 이미지와 텍스트의 유사도가 높아지도록 대조 학습(contrastive learning)을 수행한다.
- ③ 제로 샷(zero-shot) 예측과는 달리, 클래스를 추가하고 모델을 재학습하여 새로운 입력 이미지를 분류할 수 있다.
- ④ 시각 정보와 언어 정보 사이의 대응 관계를 학습한 멀티 모달 모델로 이미지 검색, 이미지 설명 생성, 시각 질의응답 등 다양한 작업으로 확장할 수 있다.

25. 다음 사례에서 과적합 위험을 최소화하면서 성능을 확보하기 위한 초기 절차로 옳은 것은?

대규모 ImageNet으로 사전 학습된 CNN(Convolutional Neural Network) 모델을 샘플 수가 수천 장 수준인 소규모 의료 영상 분류 과제에 적용하려고 한다.

- ① 사전 학습된 백본(backbone)을 고정하고, 새롭게 초기화한 분류기 레이어만 학습한다.
- ② 모든 레이어를 동일 학습률로 한 번에 동시 미세조정(fine-tune)한다.
- ③ 사전학습 가중치를 전부 고정한 채 드롭아웃(dropout) 비율만 조정한다.
- ④ 모든 레이어를 랜덤값으로 초기화하고 네트워크 전체를 새로 훈련한다.