1. 인공지능에서 지능에 해당하는 기능은 무엇인가?
- Classfication 과 Regression이 있다.

2. 인공지능의 종류 3가지에 대해서 설명하시오 (지도학습, 반지도학습, 강화학습)
- 인공지능의 종류는 학습방식에 따라 지도학습, 비지도학습, 반지도학습, 강화학습으로 나뉜다.
지도학습은 입력 데이터와 그에 상응하는 레이블 쌍으로 구성된 데이터셋을 학습시키는 방법이다.
비지도 학습은 레이블이나 명시적인 피드백 없이 입력데이터의 구조나 패턴을 발견하고 학습하는 방법으로 주로 데이터 간의 관계 발견 혹은 클러스터링에 사용된다.
위의 두 방식을 섞은 방법이 반지도 학습으로 일부만 레이블이 지정된 데이터를 사용해 학습한다. 레이블이 적은 데이터셋에서 효과적이다.

강화학습은 시행착오를 통해 학습하는 방법으로, 정의된 '에이전트'가 환경과 상호작용을 통해 환경으로부터 보상 혹은 피드백을 받으며 보상을 최대화 하도록 학습한다. 시스템이 특정 작업을 수행하는 최적의 policy를 학습하는 것이 목적이다.

3. 전통적인 프로그래밍 방법과 인공지능 프로그램의 차이는 무엇인가?
- 전통적인 프로그래밍은 사람이 직접 규칙과 조건을 정의해 알고리즘을 작성한다. 이는 일반화 능력을 제한시켜 환경의 변화에 적응이 불가능해서 수정이 필요하다. AI 프로그램은 스스로 패턴과 규칙을 학습하고 이를 통해 알고리즘을 작성한다.
따라서 환경의 변화에도 적응이 빠르다. 즉 일반화 및 적용 능력이 좋다.

4. 머신러닝과 딥러닝의 차이점은 무엇인가?
- 머신러닝은 데이터에서 특징추출이 필요하고 이 특징을 선택하기도 해야한다. 대신 작은 데이터 셋에서도 잘 동작한다.
딥러닝은 데이터 추출을 요구하지 않고, 비교적 큰 규모의 데이터셋에서 높은 성능을 보이며 복잡한 데이터를 다루는데 효과적이지만 많은 데이터와 계산 리소스가 요구된다.

5. Classfication 과 Regression 의 주된 차이점은?
- Classfication은 이산적인 값을, Regression은 연속적인 값을 예측한다. 따라서 Regression은 예측하려는 값의 범위 제한이 없다.

6. 머신러닝에서 차원의 저주(Curse of Demensionality)란?
- 차원의 저주는 머신러닝과 데이터 분석에서 발생하는 현상으로, 고차원에서 데이터가 희소해지고 불규칙해지는 현상이다.
이는 과적합, 계산 비용증가 등의 원인이 된다.

7. Demension Reduction은 왜 필요한가?
- 고차원 데이터는 시각화가 어렵고, 계산 비용과 복잡성이 증가하며 불필요한 데이터 특징이 포함될 수 있다.
Demension Reduction으로 이를 해결하고 필요한 메모리 용량을 줄여 데이터 저장 및 관리의 효율을 높일 수 있다.

8. Ridge 와 Lasso의 공통점과 차이점 (Regulation, 규제, Scaling)
- Ridge와 Lasso는 공통적으로 규제를 사용해 모델의 가중치를 제한, 축소시켜 모델의 해석 가능성을 향상한다.
둘의 차이점은 규제의 방식과 요구사항인데, Ridge는 L2 규제를 사용하며 가중치의 제곱합을 최소화하기때문에 가중치가 0이되지는 않는다. 또, 변수들이 Scaling되지 않아도 잘 작동하므로 변수의 Scale에 신경 쓸 필요가 없다.
그러나 Lasso는 변수들이 Scaling 되어야 하며, L1 규제로 가중치 합을 최소화하므로 가중치가 0이되어 특성선택이 가능하다.

9. over fitting VS Under fitting
Over fitting은 모델이 학습데이터에 너무 맞춰져 학습데이터에 대해 과하게 복잡한 구조를 학습한 경우를 말한다.
이 경우에는 학습 데이터에 한해서만 좋은 성능을 보이고, 새로운 데이터에서는 급격한 성능 저하를 보인다.
규제, 데이터 양 증가 등의 방법으로 해결한다.
반면 Under fitting은 모델이 너무 단순해 학습데이터의 복잡한 패턴을 잡아내리 못하는 경우를 의미한다.
이 경우 모든 데이터에 대해 성능이 저하된다. 해결을 위해 모델의 구조변경, 특성추가 등을 통해 복잡성을 증가시킨다.

10. Feature Engineering 과 Feature Selection의 차이점은?
- Feature Engineering은 기존 데이터를 사용하여 새로운 특성을 만들어 내거나 변환하는 과정이다.
원본 데이터의 정보를 활용하여 새롭고 의미있는 특성 생성 혹은 특성 변형으로 모델의 성능 향상.
Feature Selection은 필요없는 특성들을 제거하거나 중요도가 낮은 특성 제외 등을 통해 가장 중요한 특성들을 선택하는 과정이다.
모델에 유용한 특성들을 선택하여 모델의 복잡성을 줄이고 과적합을 방지, 학습시간 단축에 사용된다.

11. 전처리 (Preprocessing)의 목적과 방법? (노이즈, 결측치, 이상치)
- 전처리는 데이터에서 유용한 정보추출, 모델 성능 향상을 목적으로 데이터 분석 및 모델적용전에 거치는 것을 말한다.
평활화 작업이나 필터링으로 노이즈의 제거가 가능하다. 이상치는 통계적 기법으로 제거하거나 대체, 보정한다.
결측치는 해당 특성의 최빈값 평균값, 중앙값 등으로 대체가능하다.
노이즈: 데이터에 포함된 목적외의 불필요한 정보 결측치: 데이터 포인트에 발생한 값의 누락
이상치: 대부분의 데이터 패턴에서 벗어난 극단적 데이터 포인트

12. EDA (Exploratory Data Analysis) 란? 데이터의 특성 파악 (분포, 상관관계)
- EDA는 데이터를 탐색하고 특징을 파악해 데이터의 패턴, 구조를 이해하는 과정을 말한다.
EDA의 목적은 변수간의 상관관계를 분석하여 변수 간 관련성을 파악하고, 모델링에 유용한 변수를 식별하기, 변수의 분포를 시각화
하여 데이터의 중심 경향과 퍼짐 정도를 파악하기 등이 있다.

13. 회귀에서 절편과 기울기가 의미하는 바는? 딥러닝과 어떻게 연관되는가?
- 절편은 회귀직선이 독립변수(X, feature)와 종속변수(Y, label) 간의 관계를 설명하는데 있어서 독립변수가 0일 때 종속 변수의 예측값이다. 회귀직선이 종속변수를 예측하는데 얼마나 멀리 떨어져 있는지 나타내기도 한다.
기울기는 독립변수의 변화가 종속변수에 미치는 영향을 나타낸다.
딥러닝에서는 기울기와 절편이 곧 가중치와 편향을 나타내는데, 학습과정에서 조정되어 데이터를 잘 예측하는 모델을 만들도록 한다.

14. Activation fuction 함수를 이용하는 이유는? Soft max, Sigmoid 함수의 차이는?
- 활성화 함수는 인공신경망에서 각 뉴런의 출력을 결정하는 함수이다. 활성화함수는 신경망이 비선형 함수를 학습할 수 있게 해주고, 기울기의 폭주 및 소실을 방지한다. Sigmoid 함수는 입력값을 0과 1 사이로 변환하여 주로 이진 분류문제에서 사용된다.
Softmax 함수는 입력 값을 클래스에 대한 확률값으로 변환하고 다중 클래스 문제에서 출력층에 주로 사용한다.
두 함수는 출력 범위가 0에서 1까지인 점은 같지만, Softmax의 모든 확률의 합은 1이 된다는 점이 다르다.

15. Foward propagation, Backward propagation 이란?
- Foward Propagation은 입력 데이터를 신경망을 통해 전달해 출력을 계산하는 과정이다.
Backward Propagation은 Foward Propagation의 결과와 실제 값 사이의 오차를 계산, 이 오차를 각 층과 가중치에 전달하여 가중치를 조정하는 과정이다.

16. 손실함수란 무엇인가? 가장 많이 사용하는 손실함수 4가지는?
- 손실함수는 머신러닝 모델이 예측한 값과 실제 값 사이의 차이을 측정하는 함수이다. 모델의 예측값과 실제 값의 차이를 평가하고 최소 화해 모델을 학습시키는데 사용된다.
가장 많이 사용되는 손실함수는 주로 회귀 문제에서 실제 값과 예측값 사이의 제곱오차 평균을 계산하는 평균 제곱오차 와 실제값과 예측값 사이의 절대오차의 평균을 계산하는 평균 절대오차, 두 함수의 장점을 모은 후버손실 과 이진분류, 다중 클래스 분류에 많이 사용되는 크로스 엔트로피 손실 이렇게 4가지가 있다.

17. 옵티마이저 (optimizer) 란 무엇일까? 옵티마이저와 손실함수의 차이점은?
옵티마이저는 인공신경망 모델에서 학습에 사용되는 알고리즘이다. 모델의 가중치와 편향을 조정해 손실함수를 최소화하고, 최적의 모델 파라미터를 찾는 역할이다. 이를 통해 가중치 업데이트 속도조정, 학습과정 안정화가 이루어진다.
손실함수는 모델 출력과 실제 데이터 사이를 계산하여 모델의 예측이 얼마나 잘못된건지 측정한다.
반면, 옵티마이저는 그 오차를 줄이기 위해 모델의 파라미터의 조정을 결정하여 파라미터를 최적화한다.

18. 경사하강법 의미는? (확률적 경사 하강법, 배치 경사하강법, 미니 배치 경사 하강법)

경사하강법은 최적화 알고리즘의 종류 중 하나이다. 머신러닝에서와 딥러닝에서 비용함수 혹은 손실함수의 값 최소화를 위해 사용된다. 함수의 기울기를 사용해 손실을 최소화하는 가중치를 찾는 과정으로, 손실함수의 최소값에 도달 할 때까지 반복된다.

경사 하강법에는 확률적 경사하강법, 배치 경사하강법, 미니 배치 경사하강법이 있다.

확률적 경사 하강법은 각 반복에서 하나의 샘플을 무작위 선택해 기울기를 계산한다.
계산 비용은 낮고 빠르지만, 불안정하며 매우 큰 데이터 set에 적합하다.

배치 경사 하강법은 전체 훈련 데이터 셋을 사용해 각 반복에서 기울기를 계산한다.
기울기 계산은 정확하지만 데이터가 클수록 계산 비용이 많이들고 속도가 느려지므로 작은 데이터 셋에 적합하다.
미니 배치 경사 하강법은 전체 데이터셋에서 무작위로 선택된 샘플들의 소규모집합 (미니 배치)를 사용해 각 반복에서 기울기를 계산한다. 확률적 경사 하강법과 배치 경사하강법의 장점을 결합하여 균형이 좋고 가장 널리 사용된다.

19. 교차 검증, k-fold 교차검증의 의미와 차이

교차검증은 모델의 성능이 훈련데이터에 과적합되는 것을 막고, 새로운 데이터에서도 일반화 될수 있는지 확인하기 위한 방법이다. 데이터를 훈련 데이터와 검증 데이터로 나눠 사용하며 모델을 평가한다. 이를 통해 모델의 성능을 정확히 알 수 있다.

K-fold 교차검증은 교차 검증의 한 방식으로, 전체 데이터를 k개의 같은 크기의 부분집합으로 나눈다. 그리고 k번 반복을 거치며 매번 다른 집합을 검증에 사용하고 나머지 k-1개의 집합의 데이터를 합쳐 훈련 데이터로 사용한다. 모든 반복의 성능지표는 평균 내어 모델의 성능평가에 활용된다.

둘의 차이점은 검증 세트의 크기와 데이터 활용 방식에서 발견된다.

20. 하이퍼 파라미터 튜닝이란 무엇인가?

하이퍼 파라미터는 모델 학습 과정에 영향을 주지만, 사전에 모델 구축 단계에서 설정되며 데이터에서 자동으로 학습되지는 않는 파라미터이다. 학습률, 정규화 매개변수, 에폭 수, 배치 크기등이 이에 해당된다.
하이퍼 파라미터 튜닝은 머신러닝 모델의 성능을 최적화 하기위해 모델의 하이퍼리터를 조정하는 과정이다.