**필기 테스트**

**이름 :**

1. **머신러닝 알고리즘을 데이터 분석에 적용하는 단계를 순서대로 나열하시오**

|  |
| --- |
| 1. **데이터 수집** 2. **모델 성능 평가** 3. **데이터 준비와 탐구** 4. **모델 성능 향상** 5. **데이터에 대한 모델 훈련** |

1. ⑴-⑵-⑶-⑷-⑸
2. ⑴-⑶-⑸-⑵-⑷
3. ⑴-⑵-⑶-⑸-⑷
4. ⑴-⑸-⑵-⑶-⑷
5. **다음 머신러닝 알고리즘을 분석 목적에 맞게 연결하시오**

|  |  |  |
| --- | --- | --- |
| (1) 결정트리  (2) 선형회귀  (3) 신경망  (4) 클러스터링 | **∙ ∙**  **∙ ∙**  **∙ ∙**  **∙ ∙** | (가) 수치 예측  (나) 다중 용도  (다) 그룹 나눔  (라) 분류 |

1. **최근접 이웃은 k 개의 이웃하는 데이터를 이용하여 분류 모델을 만들 수 있는 알고리즘이다.** ( o , x )
2. **결정 트리는 수치 데이터에만 적용할 수 있는 알고리즘이다.** ( o , x )
3. **서포트 벡터 머신은 블랙박스 모델이 아니다.** ( o , x )
4. **나이브 베이즈 알고리즘은 노이즈와 결측 데이터에 아주 민감하게 수행된다.** ( o , x )
5. **랜덤 포레스트는 결정트리에 앙상블 기법을 적용한 알고리즘이다.** ( o , x )
6. **수치 데이터 간 거리는 유클리디언 거리를 이용하여 계산할 수 있다.** ( o , x )
7. **다음 중 신경망의 레이어의 대한 설명으로 틀린 것은 ?**
8. 레이어의 개수와 정보 진행 방향, 그리고 노드의 개수에 따라 복잡성이 다양해진다.
9. 입력 노드의 수는 입력 데이터의 속성 개수로 결정된다.
10. 출력 노드의 수는 결과의 분류 개수나 모델의 결과 수로 결정된다.
11. 적당한 은닉층의 개수와 노드의 수를 결정할 수 있는 규칙이 있다.
12. **다음 중 머신러닝 알고리즘의 특징에 대한 설명으로 틀린 것은 ?**
13. kNN, 결정나무, 나이브베이즈, 신경망은 분류 모델에 적용할 수 있는 알고리즘이다.
14. K-means, 연관규칙 마이닝은 라벨 데이터가 없는 데이터에 적용할 수 있는 알고리즘이다.
15. 신경망과 SVM 은 대표적인 블랙박스 모델로 결과를 이해하는데 어려움이 있는 알고리즘이다.
16. 앙상블 기법은 동일한 모델 알고리즘에만 적용할 수 있는 기법이다.