|  |  |
| --- | --- |
| 교육 제목 | **머신러닝 복습 및 강화학습** |
| 교육 일시 | 2021년 11월 26일 금요일 |
| 교육 장소 | 영우글로벌러닝 2층 |
| **교육 내용** | |
| 오전 | 1. 파이썬 머신러닝 판다스 데이터 분석 교재 리뷰 |
| 오후 | 1. 강화학습 Reinforcement Learning 2. 강화학습의 핵심은 일단 해보는 것입니다    1. 게임 🡪 환경(environment)    2. 게이머 🡪 에이전트(agent)    3. 게임화면 🡪 상태(state)    4. 게이머의 조작 🡪 행동(action)    5. 상과 벌 🡪 보상(reward)    6. 게이머의 판단력 🡪 정책(policy) 3. 강화학습에서는 더 많은 보상을 받을 수 있는 정책을 만드는 것이 핵심 4. Machine Learning은 다음 세가지를 포함하는 개념이다.    1. Supervised Learning(지도학습)    2. Unsupervised Learning(비지도학습)    3. Reinforcement Learning(강화학습) 5. Reinforcement Learning이 가지는 특징    1. There is no supervisor(감시자) : 누구도 답(행동)을 알려주지 않음    2. supervised learning과의 차이점       1. supervisor가 없으므로 optimal solution을 찾는 목적에 적합하다    3. Only a reward signal : 오직 reward(보상) 신호만이 존재 6. Reinforcement Learning은 reward를 maximize(최대화)하는 목적    1. Feedback is delayed, not instaneous : Feedback이 즉각적이지 않고 지연될 수 있다.    2. Action(행동)이 당장의 결과가 아닌 미래의 결과에 영향을 줄 수 있다.    3. Time really matters: 시간이 중요하다!    4. Sequential data: 순서와 시간이 중요한 데이터    5. non i.i.d data: independent identically distributed(독립 동일 분포)가 아니다.    6. i.i.d: 동일한 확률분포로 부터 뽑힌 확률 변수 / 각각은 서로에게 독립적이다.    7. Agent’s actions affect the subsequent data it receives    8. Agent의 Action이 후의 데이터에 영향을 미친다. 7. Major Component of an RL Agent    1. Policy(정책) : Agent가 행동을 결정하는 규칙    2. Value: 지금의 State 혹은 Action이 얼마나 좋은지를 판단하는 척도    3. Model: Model을 통해 Agent가 Environment를 예측한다. 8. Categorizing RL agents 9. Agent의 구성에 따라 크게 3가지로 나눌 수 있다.    1. Value Based    2. No Policy, Only Value Function   i.  B. Policy Based  i. No Value Function, Only Policy  C. Actor Critic  i. Policy와 Value Function 둘다 존재  D. Model Free: Model 없음  i. Model Based: Model 존재 |