

기계학습 및 기본 알고리즘 연구

김동현⁰, 이태호*, 이병준*, 김경태*, 윤희용**

⁰성균관대학교 정보통신대학 전자전기컴퓨터공학과

**성균관대학교 소프트웨어대학 소프트웨어학과

e-mail: {kdh7263⁰, leetaeho*, byungjun*}@skku.edu, kyungtaekim76@gmail.com*, youn7147@skku.edu**

A Study on Machine Learning and Basic Algorithms

Dong-Hyun Kim⁰, Tae-ho Lee*, Byung-Jun Lee*, Kyung-Tae Kim*, Hee-Yong Youn**

⁰Dept. of Electrical and Computer Engineering, Sungkyunkwan University

**Dept. of Software, Sungkyunkwan University

● 요약 ●

본 논문에서는 기계학습 및 기계학습 기법 중에서도 Markov Decision Process (MDP)를 기반으로 하는 강화학습에 대해 알아보고자 한다. 강화학습은 기계학습의 일종으로 주어진 환경 안에서 의사결정자(Agent)는 현재의 상태를 인식하고 가능한 행동 집합 중에서 보상을 극대화할 수 있는 행동을 선택하는 방법이다. 일반적인 기계학습과는 달리 강화학습은 학습에 필요한 사전 지식을 요구하지 않기 때문에 불명확한 환경 속에서도 반복 학습이 가능하다. 본 연구에서는 일반적인 강화학습 및 강화학습 중에서 가장 많이 사용되고 있는 Q-learning 에 대해 간략히 설명한다.

키워드: 기계학습(machine learning), 강화학습(reinforcement learning), 큐러닝(q-learning)

I. Introduction

기계학습이란 사람의 개입 없이도 경험과 추론을 기반으로 컴퓨터 스스로 학습할 수 있는 학습 방법이다. 사람은 컴퓨터에게 스스로 학습할 수 있는 기본적인 환경을 제공하며, 컴퓨터는 학습을 통해 고정된 결과 값을 출력하는 것이 아닌 주어진 환경에 적합한 결과 값을 출력하게 된다. 무선으로 상호 연결된 환경이 증가함에 따라, 이러한 기계학습은 다양한 분야에서 사용되고 있다[1]. 기계학습은 학습 방법에 따라 크게 지도학습 (Supervised learning), 비지도학습 (Unsupervised learning), 그리고 강화학습(Reinforcement learning)의 세 가지로 분류된다.

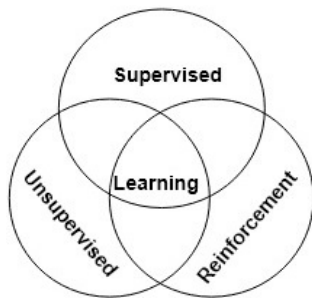


Fig. 1. The type of Machine learning algorithm

II. Preliminaries

1. Related works

1.1 Supervised learning

지도 학습은 컴퓨터에게 Input data와 Labeled data를 제공하고, 컴퓨터는 Labeled data를 이용해 Data를 분류(Classification) 및 회귀(Regression)하는 학습 방법이다. 지도 학습 중 가장 자주 사용되는 알고리즘으로 Support vector machine (SVM)이 있다. SVM은 기존 데이터 집합을 기반으로 새로운 데이터가 들어왔을 때, 두 그룹 사이의 가장 큰 Margin을 가지는 초평면을 통해 어느 그룹에 속하는지를 판단하는 알고리즘이다.

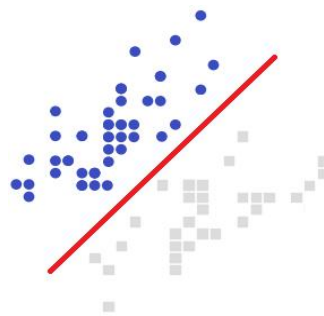


Fig. 2. The example of the SVM

1.2 Unsupervised learning

비 지도학습은 지도학습과는 다르게 데이터에 대해 정확한 정의를 하는 것이 아닌, **Input data**만을 이용하여 비슷한 속성을 지닌 데이터를 군집화(Clustering) 하는 것이다. 대표적인 알고리즘으로 **K-means clustering** 알고리즘이 있다. **K-means** 알고리즘이란 데이터를 **K**개의 Cluster 단위로 묶어, 각 Cluster 사이의 거리를 최소화하는 알고리즘이다.

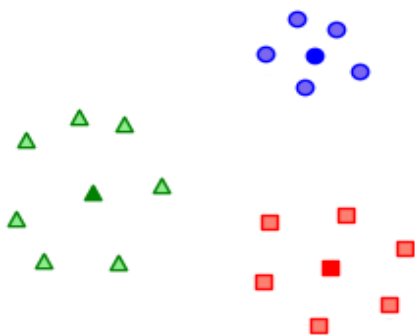


Fig. 3. The example of the K-means clustering

1.3 Reinforcement learning

-Markov process

미래의 상태 s' 가 오직 현재의 상태 s 의 영향만을 받아 전이되는 것을 의미한다. 즉, 과거의 상태는 미래의 상태에 영향을 미치지 않는다.

-Markov Decision process

MDP는 MP를 기반으로 하는 의사 결정 프로세스로, 미래의 상태 s' 는 현재의 상태 s 뿐만 아니라 s 에서 행한 Action의 영향도 받는다. 시간 t 에서 Agent는 현재 상태 s 에서 선택 가능한 Action 중 행동 a 를 취함으로써 새로운 상태 s' 로 전이하며, 이에 따라 Agent는 action a 에 대한 Reward $r(s, a)$ 를 받는다.

-Q-learning

Q-learning은 강화학습의 일종으로, 학습에 사용되는 Scheduling policy는 Greedy하게 향상되고, Q-function을 기반으로 Action 한다.

$$q(s, a) = R(s, a) + \gamma \max_{a'} q(s', a')$$

즉, 가능한 행동 집합 $a \in \{a_1, a_2, \dots, a_n\}$ 중에서 Reward의 합을 극대화 할 수 있는 행동을 함으로써 학습한다.

IV. Conclusions

본 연구에서는 기계학습 및 알고리즘에 대해 서술하였다. 기계학습에 사용되는 알고리즘들은 각기 다른 특성을 지니고 있기 때문에 환경에 맞는 적합한 알고리즘을 사용할 때 가장 기대효과가 클 것으로 예상된다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 정보통신·방송연구 개발 사업(No. 2016-0-00133, 초연결 IoT 노드의 군집 지능화를 통한 Edge Computing 핵심 기술 연구), SW중심대학지원사업(2015-0-00914), 한국연구재단 기초연구사업(No.2016R1A6A3A11931385, 실시간 공공안전 서비스를 위한 소프트웨어 정의 무선 센서 네트워크 핵심기술 연구, 2017R1A2B2009095, 실시간 스트림 데이터 처리 및 Multi-connectivity를 지원하는 SDN 기반 WSN 핵심 기술 연구), 삼성전자, BK21PLUS 사업의 일환으로 수행되었음.

REFERENCES

- [1] L. Cui et al. "A Survey on Application of Machine Learning for Internet of Things," International Journal of Machine Learning and Cybernetics, 2018