

그림으로 배우는 AI 인공지능

Chapter1 : AI의 기본

다들 내용은 '약한 AI' = 지능이 있는 것처럼 보이는 AI

기호주의 : 인간의 논리적인 사고구조를 참고한 모델

- 단순, 명쾌

커넥셔니즘 : 인간 뇌의 신경 네트워크를 참고한 모델

- 신경망 (AI의 가능성을 넓힌)
- 딥러닝

Chapter2 : AI 와 프로그램의 기본 구조

리스트 구조 형식의 데이터 탐색

- 선형 탐색 알고리즘 : 위에서부터 순서대로
- 이분 탐색 알고리즘 : 데이터의 중간을 기준으로 나눈다

트리 구조 형식의 데이터 탐색

- 폭 우선 탐색 : 같은 높이의 물건을 옆으로 찾아간다
- 깊이 우선 탐색 : 깊은 곳까지 찾고 나서 돌아간다
- 몬테카를로 트리 탐색 : 정답일 확률이 높은 것부터 차례로 찾는다

★데이터의 종류에 따라 달라지는 탐색 방법

AI나 컴퓨터라고 하더라도 정리되지 않은 데이터를 탐색하는 것은 어려움

- ➔ 소트 알고리즘의 중요성

아키텍처

- ➔ 수집한 데이터를 어디에 모으고, 애플리케이션이나 알고리즘을 어떻게 사용하고, 데이터 센터의 요건, 분석이나 판단에 이용하는 방법, 제안이나 관리 방법 등을 생각해 포함

AI의 사고방식과 의사결정 기법

- Rule(지식) 베이스 : Rule을 정하고 그에 따라 동작하는 것
- State 베이스 : 상태를 결정하고 상태에 따라 동작하는 것이 state 베이스
- Goal 베이스 : 계획을 세우고 계획대로 진행하는 것

- Task 베이스 : 작업을 분할하고 유연하게 진행하는 것 (협조하여 일을 진행)
- Case 베이스 : 과거 사례를 참고하는 것 (Rule화가 어려운 것을 접근하는 법)
- Simulation 베이스 : 미래를 상상하고 재현해서 최적의 것을 찾는 것

Chapter3 : AI에서의 데이터 취급

정보(데이터)

★AI에 따라 취급할 수 있는 데이터의 종류가 다르고 데이터가 무엇을 의미하는지 정보가 생겨난 경위를 이해하고 있어야 한다.

- 구조화 데이터

⇒ 컴퓨터가 보기 쉽게 각각의 요소에 대해 메타 데이터 부여

- 메타 데이터

⇒ 데이터에 대한 데이터 (다른 데이터를 정의하고 기술하는 데이터)

- 비구조화 데이터

⇒ 정리되지 않은 데이터 (인간은 이해할 수 있는 정도로 정리되어 있는 경우가 있음)

데이터 마이닝

⇒ 데이터에서 가치 있는 정보를 찾아내는(마이닝) 기술

⇒ 데이터 관계성에서 유용한 정보를 도출

텍스트 마이닝

⇒ 문장을 분석한다기보다는 단어를 분석하여 가치를 찾는다는 특성이 강해지고 있다

⇒ 단어의 빈도나 관계성으로부터 상품이나 서비스의 평가를 SNS 또는 설문조사로부터 조기 발견

분석방법

1. 회귀 분석

◆ 좌표상에 점을 찍고 점에 따라 선을 그어서 데이터 간의 관계성을 수식으로 표현하는 기법

◆ 수식으로 나타낼 수 있는 것에 강하다

2. 연관성 분석

◆ 다양한 종류의 데이터가 있는 경우 => 어떤 것과 어떤 것이 서로 관계성이 있는지 찾는 것

3. 클러스터링

◆ 다양한 데이터를 그룹화하고 분류하는 기법

◆ 분류 기준을 알 수 없는 잡다한 데이터에 사용할 수 있음

4. 디시전 트리 분석

- ◆ 질문에 답하면서 대상을 분류해 나가는 기법
- ◆ 난해한 분류 대상이라도 분석이 가능해짐

정보가 애매할 때

퍼지 이론 : 애매한 정보에 대해 다룰 수 있도록 하는 것이 퍼지 이론

- 서포트 벡터 머신 : 애매한 것을 명확하게 식별하는 기법
- 베이스 추정
 - 불확정 요소가 많을 경우 유용
 - 과거의 데이터를 참고로 분석을 하는 기법
- 마르코프 추정
 - 과거 데이터나 외적 요인이 무시된 프로세스
 - 확실한 현재의 상태에서부터 미래가 결정

Chapter4 : 기계학습

신경망

- ➔ 가중치의 수치를 조정하는 것으로 학습이 진행되고 해답을 낼 수 있다
- ➔ 입력된 데이터에 대해서 그것을 처리하는 네트워크 간의 파라미터를 조정하는 것으로 적절한 출력을 내는 기법

베이지안 네트워크

- ➔ 데이터의 구조 : 그래프 구조
- ➔ 데이터 간의 연관성에 의해 기술된 네트워크를 상황 예측이나 진단에 사용

지도 학습

- ➔ 모르는 대답은 올바른 대답을 내는 방법
- ➔ 학습을 함으로써 해답이 없어도 문제에서 답을 얻을 수 있게 한다
- ➔ 예측을 세우면서 데이터를 모아 정답에 접근할 수 있다

비지도 학습

- ➔ '특징'을 찾기 위한 기법
- ➔ 답이 없어도 특이한 것을 찾는 작업이나 공통되는 것을 찾는 작업을 학습해나간다
- ➔ 데이터 마이닝, 클러스터링, 온톨로지 영역에서 사용

강화학습

- ➔ 답이라고 확신X 방향성이나 목표를 제시
- ➔ 보상을 주는 형태로 뛰어난 행동을 강화하고 잘못된 행동을 약화시킨다

심층 강화학습 : 다른 기계학습과 친밀도가 높고 딥러닝을 결합

역 강화학습 : 최적의 행동으로부터 최적의 보상을 배운다. 최적의 보상이 있으면 기계도 최적의 행동을 하게 될 것

노 프리 런치 정리 : 문제를 해결하기 위해 그때마다 알맞은 최적의 알고리즘을 선택

국소 최적해 : 가끔 잘 진행되고 있는 행동을 바꾸지 않는다는 것

- ➔ 돌연변이 와 같은 대담한 변화로 국소 최적해 탈출 가능

기계 학습의 효율화

1. 더미 데이터
 - 진짜를 속 빼닮은 더미 데이터를 만듦
 - 질 좋은 더미 데이터는 질 나쁜 진짜 데이터보다 학습효율이 좋다
2. 반지도 학습
 - 지도 학습과 반지도 학습 결합
 - 지도 학습으로 중요한 특징을 배우고 비지도 학습으로 특징을 알아보는 법 배운다
3. 사전 학습
 - 메인 학습 전에 기본사항을 배우는 학습
4. 전이 학습
 - 미리 다른 태스크를 학습한 모델을 전용하는 학습
 - 이를 통해 학습 기간을 대폭 단축 시킬 수 있음
5. 앙상블 학습
 - ✓ 과학습(over-fit)
 - 편견이라고 생각하면 편함
 - 비슷한 데이터를 너무 많이 학습
 - ➔ 훈련할 때는 오차가 적음, 테스트 오차는 커진다

Sol) 앙상블 학습 : 여러 개의 AI를 넣어 합의로 결정하는 학습