

인공지능과 알고리즘 윤리

승실대학교 베어드교양대학
 서유희 교수
 yhsuh@ssu.ac.kr

목차

- ▶ 알고리즘과 휴리스틱의 정의
- ▶ 모델링의 개요

지난시간 배운 내용

주	주제	온라인	오프라인
1	인공지능의 과거 현재와 미래	1. 강의 및 교과목 소개(공통, 핵심만) 2. 인공지능의 과거와 현재 3. 인공지능의 미래와 다양한 시선 4. 인공지능 개발환경 구축과 사용법(Anaconda/Colab)	1. 강의 및 교과목 소개(분반별, 자세히) 2. 다양한 인공지능 기술 경험하기 (자연어처리, 시각, 음성) 3. 인공지능 챗봇만들기(IBM 왓슨 어시스턴트)
2	공공데이터를 이용한 사회문제 발견과 해결책 모색	1. 빅데이터의 정의와 가치 2. 공공데이터 수집하기 3. 공공데이터로부터 새로운 인사이트 발견하기 - 행정구역별 인구 데이터와 공공의료기관 현황 데이터 분석	1. 서울시 CCTV설치 현황 분석하기 2. 서울시 범죄발생 현황 분석하기
3	인공지능의 개요 및 머신러닝을 이용한 예측	1. 인공지능의 정의와 분류 2. 인공지능 학습방법 이해하기 3. 인공지능 알고리즘 소개	1. 머신러닝을 이용한 이미지 식별(구글 티쳐블 머신) 2. 머신러닝을 이용한 보스턴 집값 예측
4	인공지능과 데이터 윤리	1. 데이터의 불완전성과 결함에 따른 예측 오류와 차별 2. 데이터 편향성이 예측에 미치는 영향 (구글 티쳐블머신) 3. 지도학습(SVM)을 이용한 타이타닉호 생존자 예측	1. 타이타닉호 생존자 예측 - 데이터 편향성이 예측에 미치는 영향 - 데이터 왜곡에 따른 예측 결과 비교
5	인공지능과 알고리즘 윤리	1. 알고리즘과 모델링의 개요 2. 알고리즘 기반 의사결정 시스템의 한계 3. 윤리가 필요한 인공지능 4. 오렌지3 설치 및 사용법	1. 오렌지3를 이용한 알고리즘에 따른 예측 결과 비교 - 보스턴 집값 예측 - 폐암환자 생존 여부 예측
6	인공지능에 대한 다양한 이슈와 우리의 자세 고찰	1. 인공지능의 윤리적/법적 쟁점 (자율주행자동차, AI로봇, 트랜스 휴먼 등) 2. 인공지능시대 사회, 경제적 불평등 문제 3. 인공지능과 프라이버시 4. 인공지능의 윤리적 대응과 규제	1. 자율주행 자동차의 행동학습 시나리오 경험하기 2. 비윤리적 데이터 생성과 수집(웹 크롤링을 이용한 데이터 수집)
7		기말고사	

학습 목표

- ▶ 알고리즘과 휴리스틱의 차이를 설명할 수 있다.
- ▶ 인공지능에 있어 모델링의 개념과 그 중요성을 설명할 수 있다.

알고리즘과 모델링의 개요

알고리즘과 휴리스틱

6

▶ 알고리즘 (Algorithm)

- ▶ 어떠한 문제를 해결하기 위해 명확히 정의된 일련의 단계적 절차 또는 행동지침
- ▶ 입력과 출력 사이에 가져야 하는 관계를 규정하며 유한 시간내에 종료되어야 함
- ▶ 알고리즘은 정확히 구현되는 경우 모든 정확한 인풋에 대해 정확한 아웃풋을 계산
- ▶ 정해진 정답을 보장하는 문제를 해결하는 방법

▶ 휴리스틱 (Heuristic)

- ▶ 수학적 증명이 없이 경험과 직관을 통해 시행착오를 거치며 발견해 가는 방법
- ▶ 고대그리스어 ‘발견하다’, ‘찾아내다’ 의 뜻인 휴리스케인(heuriskein)에서 유래
- ▶ 모든 조건을 충족시키는 해답을 발견하거나 정답이라는 보장은 없음
- ▶ 인공지능의 많은 알고리즘은 본질적으로 휴리스틱이거나, 휴리스틱 규칙을 사용

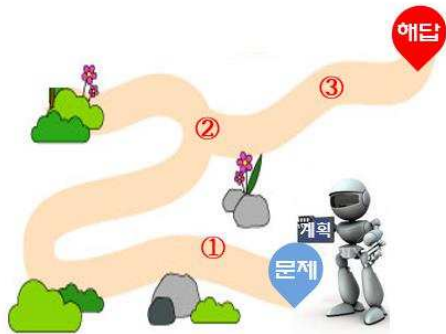
알고리즘과 휴리스틱

7

알고리즘

≠

휴리스틱



알고리즘 사용의 예

8

▶ 정렬 문제

- ▶ 일련의 특성을 가진 분류대상 중 어떤 것이 더 앞이고, 어떤 것이 더 뒤인지 정확히 규정하는 규칙에 따라 정렬
- ▶ 삽입정렬, 버블정렬 등

▶ 정렬 알고리즘 사용의 예

- ▶ 구글이 부여하는 중요도에 따른 웹사이트 정렬
- ▶ 제품 인기도 순 정렬
- ▶ 사진을 평균 색도에 따라 정렬

▶ 알고리즘을 이용해 현실세계의 다양한 문제들을 똑같은 추상적 문제로서 모델링 할 수 있다

- ▶ 알고리즘 자체는 변화시키지 못하나 모델에 따라 결과는 달라진다
 - ▶ 정렬 규칙이 모델의 인풋의 일부임으로

모델링이란?

9

▶ 모델링의 일반적인 정의

- ▶ 모두에게 공통적으로 이해되도록 약속된 방식(수식,그림 등)으로 표현하는 것
- ▶ 복잡한 현실세계를 단순화(추상화) 즉, '추상적 체계적으로 표현하는 작업 또는 방법

▶ 최단경로로 목적지까지 길찾는 네비게이션 만들기 예

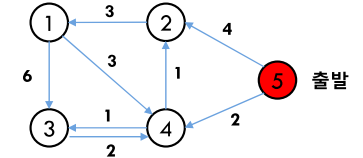
- ▶ 최단경로 알고리즘(다익스트라, Dijkstra)를 사용하여 거리망을 모델링
- ▶ 알고리즘의 인풋 정보 결정
 - ▶ 최단경로의 기준을 측정가능하도록 설정 (킬로미터로 따진 구간 길이 기준)
 - ▶ 일방통행인지, 양방통행인지
- ▶ 알고리즘에 어떤 정보를 인풋으로 줄지 결정하는 것은 모델링의 일환으로 알고리즘 결과를 해석하는데 중요함
- ▶ 인공지능에서는 수학적 특성을 띠지 않는 사회적 개념을 측정가능하게 만들어야 만들어야 운영이 가능
 - ▶ 어떤 소식이 갖는 비중, 우정, 범죄성향, 신용, 사랑

다익스트라 알고리즘

10

• 소스 노드와 다른 모든 노드 간의 최단 경로를 찾기

1. 경로가 가장 짧은 노드 선택을 선택 5번 / 5번 노드의 인접 노드 2번, 4번
2. \min (2번 노드의 현재 최단거리(∞), 5번 노드의 최단거리(0) + 2번-5번 가중치(2)=4) 갱신
3. \min (4번 노드의 현재 최단거리(∞), 5번 노드의 최단거리(0) + 4번-5번 가중치(4)=2) 갱신
4. 나머지 노드 중 경로가 가장 짧은 노드를 선택 4번 / 4번 노드의 인접 노드 2번, 3번
5. \min (2번 노드의 현재 최단거리(4), 4번 노드 최단거리(2) + 4번-2번 가중치(1) =3) 갱신
6. \min (3번 노드의 현재 최단거리(∞), 4번 노드 최단거리(2) + 4번-3번 가중치(1) =3) 갱신
7. 나머지 노드 중 경로가 가장 짧은 노드를 선택 2번 / 2번 노드의 인접 노드 1번
8. \min (1번 노드의 현재 최단거리(∞), 2번 노드의 최단거리(3) + 2번-1번 가중치(3) = 6) 갱신
9. 나머지 노드 중 경로가 가장 짧은 노드를 선택 3번 / 3번 노드의 인접 노드 4번
10. \min (4번 노드의 현재 최단거리(2), 3번 노드의 최단거리(3) + 3번-4번 가중치(2) = 5) 유지
11. 마지막 노드 선택 1번 / 1번 노드의 인접노드 3번, 4번
12. \min (3번 노드의 현재 최단거리(3), 1번 노드의 최단거리(6) + 1번-3번 가중치(6) = 12) 유지
13. \min (4번 노드의 현재 최단거리(2), 1번 노드의 최단거리(6) + 1번-4번 가중치(3) = 9) 유지



5	4	3	2	1
0	∞	∞	∞	∞



5	4	3	2	1
0	∞	∞	∞	∞
0	2	∞	4	∞
0	2	3	3	∞
0	2	3	3	6
0	2	3	3	6

2. 3. 갱신
5. 6. 갱신
8 갱신
10 유지
12. 13 유지, 유지

철도망 모델링의 예

11

▶ 거리망에서 사용한 최단경로 알고리즘을 철도망에서도 그대로 사용할 수 있을까?

- ▶ 기차역은 거리망의 교차로
- ▶ 기차가 직통으로 연이어 통과하는 두 기차역을 하나의 선으로 연결
- ▶ 철도 연결선의 길이를 킬로미터 대신 예상 소요시간으로 표시



만하임역

ICE 770으로는 47분

TGV로는 37분



프랑크푸르트역

철도망 모델링의 예

12

▶ 철도에서 최단노선 문제를 거리망과 똑같이 모델링할 때 나타나는 문제

- ▶ 철도에서는 각 기차의 도착시간 및 출발시간, 갈아탈 때 대기하는 시간 고려해야 함
- ▶ 거리망 모델만으로는 실제적 사용 불가
 - ▶ 각 중간 역에서 다음 역까지 절대적으로 최단시간이 소요되는 기차를 탈수 있을 거라는 조건하에 최단시간 경로를 알려주기 때문
 - ▶ 승강장 대기 시간은 고려하지 않음

“카이저슬라우테른에서 만하임으로 오전 10시에 출발하는 ICE를 타세요.”
“그리고 만하임에서 다음날 아침 7시 50분에 베를린으로 가는 ICE를 타세요.”



“카이저슬라우테른에서 베를린까지 가장 빨리 갈 수 있는 열차경로 추천해줘.”

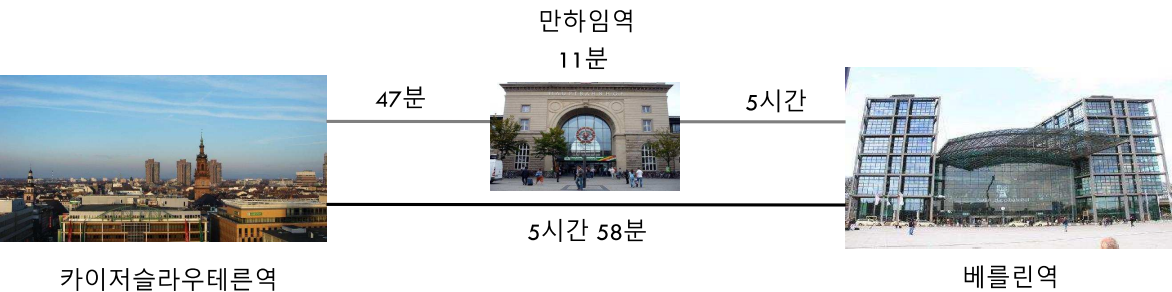


철도망 모델링의 예

13

도이체반(독일철도 주식회사) 사례

- 다양한 최단경로 문제를 해결하기 위해 고전적인 다익스트라 알고리즘을 변형한 알고리즘을 활용

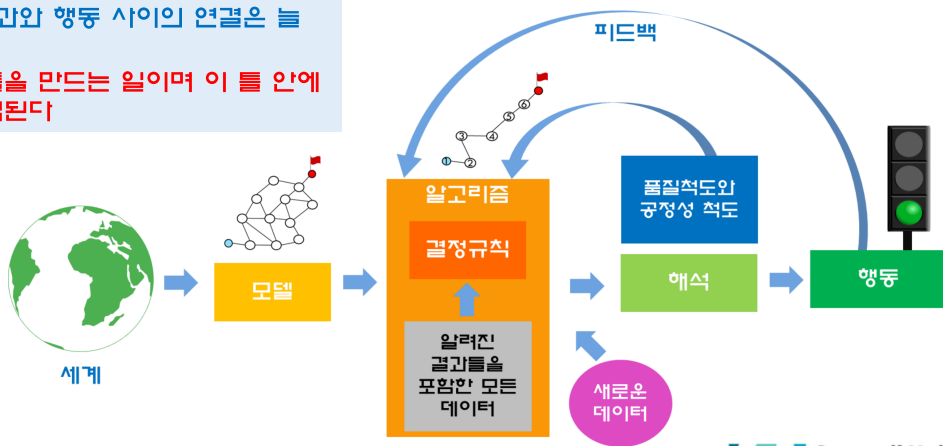


모델링의 중요성

14

알고리즘 기반 의사결정 시스템 개발과 활용의 과정

알고리즘의 결과와 행동 사이의 연결은 늘 분명하지 않다
모델링은 큰 틀을 만드는 일이며 이 틀 안에서 결과가 해석된다



다음시간에 배울 내용

15

주	주제	온라인	오프라인
1	인공지능의 과거 현재와 미래	1. 강의 및 교과목 소개(공통, 핵심만) 2. 인공지능의 과거와 현재 3. 인공지능의 미래와 다양한 시선 4. 인공지능 개발환경 구축과 사용법(Anaconda/Colab)	1. 강의 및 교과목 소개(분반별, 자세히) 2. 다양한 인공지능 기술 경험하기 (자연어처리, 시각, 음성) 3. 인공지능 챗봇만들기(IBM 왓슨 어시스턴트)
2	공공데이터를 이용한 사회문제 발견과 해결책 모색	1. 빅데이터의 정의와 가치 2. 공공데이터 수집하기 3. 공공데이터로부터 새로운 인사이트 발견하기 - 행정구역별 인구 데이터와 공공의료기관 현황 데이터 분석	1. 서울시 CCTV설치 현황 분석하기 2. 서울시 범죄발생 현황 분석하기
3	인공지능의 개요 및 머신러닝을 이용한 예측	1. 인공지능의 정의와 분류 2. 인공지능 학습방법 이해하기 3. 인공지능 알고리즘 소개	1. 머신러닝을 이용한 이미지 식별(구글 티쳐블 머신) 2. 머신러닝을 이용한 보스턴 집값 예측
4	인공지능과 데이터 윤리	1. 데이터의 불완전성과 결함에 따른 예측 오류와 차별 2. 데이터 편향성이 예측에 미치는 영향 (구글티쳐블머신) 3. 지도학습(SVM)을 이용한 타이타닉호 생존자 예측	1. 타이타닉호 생존자 예측 - 데이터 편향성이 예측에 미치는 영향 - 데이터 왜곡에 따른 예측 결과 비교
5	인공지능과 알고리즘 윤리	1. 알고리즘과 모델링의 개요 2. 알고리즘 기반 의사결정 시스템의 한계 3. 윤리가 필요한 인공지능 4. 오렌지3 설치 및 사용법	1. 오렌지3를 이용한 알고리즘에 따른 예측 결과 비교 - 보스턴 집값 예측 - 폐암환자 생존 여부 예측
6	인공지능에 대한 다양한 이슈와 우리의 자세 고찰	1. 인공지능의 윤리적/법적 쟁점 (자율주행자동차, AI로봇, 트랜스 휴먼 등) 2. 인공지능시대 사회, 경제적 불평등, 프라이버시 문제 3. 인공지능과 프라이버시 4. 인공지능의 윤리적 대응과 규제	1. 자율주행 자동차의 행동학습 시나리오 경험하기 2. 비윤리적 데이터 생성과 수집(웹 크롤링을 이용한 데이터 수집)
7		기말고사	