

중학교 | 학생용



인공 지능 윤리

체험중심



중학교 | 학생용

인공 지능 윤리

체험중심

최정원

김형기

권순찬

홍승연



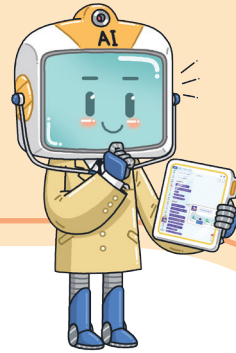
머리말

인공지능 윤리는 인공지능의 발전과 함께 진화하고 있습니다. 인공지능 기술은 우리 인간이 어떻게 개발하고 활용하느냐에 따라 긍정적으로 활용될 수도, 부정적으로 활용될 수도 있습니다. 앞으로 다가오는 미래 시대에는 인공지능의 기술적 원리를 바탕으로 윤리적 사고를 하는 것이 필요합니다. 인공지능 시대에 발생하는 여러 사회적 문제들이 인간 사회에서 일어나는 윤리적 문제와는 다른 방식으로 일어나기 때문입니다.

이에 인공지능 윤리 교육에 대한 수요가 점점 증가하고 있습니다. 점점 더 많은 인공지능 기술이 우리의 삶으로 들어오고 있고, 우리의 사회·경제적 이익, 기회에 직접적인 영향을 줄 수 있기 때문입니다. 인공지능 윤리에 대해 배운다는 것은 인공지능 기술이 가져오는 순기능을 최대화하고, 역기능은 최소화하는 노력을 통해서 인공지능을 이롭게 사용하자는 사회적 합의를 이루어 나가는 과정입니다.

인공지능은 사람에게 편의를 제공하고 삶의 질을 높이기 위해서 개발·활용되고 있습니다. 모든 기술이 그러했듯이 인공지능도 사람을 위해서 개발되고 있으나, 개발·사용하는 과정에서 역기능이 발생할 수 있습니다. 정해진 루틴만 반복하던 기존 자동화 시스템과는 달리, 주변 상황을 고려하여 판단하고 행동하는 사람의 지능이 하던 일 이상의 것들을 할 수 있다는 무한한 가능성이 있습니다. 이렇듯 기대치가 높은 기술인 동시에 두려움의 대상이기 때문에 인공지능 윤리가 그 어느 때보다도 강조되고 있습니다.

초기에는 인공지능이 윤리를 갖추어야 한다는 관점으로 시작하였으나, 인공지능을 만들고 사용하는 주체는 결국 사람이므로 사람들이 인공지능 윤리를 갖추어야 한다는 인식으로 변화하였습니다. 지금은 주로 사용자로서의 역할을 하고 있는 우리 학생들은 미래의 개발자가 될 수도 있고, 관리자가 될 수도 있습니다. 따라서 인공지능을 개발하고 사용하는 전 과정에서 갖추어야 할 인공지능 윤리에 대해 생각해 보고 지금부터 실천할 수 있도록 교육하는 것이 매우 중요합니다.



본 교재는 인공지능 활용 경험과 윤리를 따로 분리하여 교육하고 있는 현 실정을 보완하여 통합적인 교육을 지향합니다. 인공지능을 개발하고 활용하는 전 과정에서 발생 가능한 윤리적 쟁점을 스스로 생각·판단을 보고 친구들과 생각을 비교·분석하면서 종합적으로 의견을 정리할 수 있도록 했다는 점이 특징입니다.

본 교재에서는 인공지능의 개념, 원리, 작동 방식에 대한 이해를 바탕으로 전 세계에서 일어나고 있는 인공지능 활용 사례를 통해 윤리적 이슈를 확인하게 될 것입니다. 특히 「인공지능 윤리 기준」의 10대 핵심 요건 중 관련 있는 요건들을 한 챕터로 묶어 중학생들이 체험 활동을 통해서 이러한 가치들을 확인할 수 있도록 교재를 구성하였으며 하나의 주제를 약 4차시로 구성하였습니다. 그러나 상황에 따라 필요한 차시를 분리하여 사용할 수 있도록 모듈별·차시별로 구성하였습니다. 각 차시별 특징은 다음과 같습니다.

- 1차시: 10대 핵심 요건과 관련된 문제 상황을 제시하고, 핵심 요건을 이해한 후 인공지능 프로그램 개발 과정에서 개발자와 사용자의 역할 이해
- 2차시: 인공지능 윤리와 연계한 인공지능 프로그램 개발
- 3차시: 2차시에서 개발한 인공지능 프로그램의 사용 과정에서 발생할 수 있는 윤리적 문제를 파악하고, 대응 방안을 탐색하는 과정에서 개발자와 사용자의 관점 차이 분석, 관점의 차이를 좁히기 위한 합의의 필요성 인식
- 4차시: 인공지능 윤리 실천을 위한 활동

다만 인공지능 윤리 교육을 하다 보면 부정적인 사례를 제시할 수 밖에 없다는 점에서 오히려 바람직한 발전 방향보다 인공지능에 대해 부정적 인식을 갖게 될까 봐 걱정이 되기도 합니다. 그러나 데이터를 수집하고 학습하는 과정에서 어쩔 수 없이 다루어야 하는 부정적인 부분들을 토대로 보다 현실적인 인공지능 윤리 교육을 할 수 있을 것이라 생각합니다.

우리 학생들이 본 교재를 통해서 윤리적 태도가 중요함을 인식하고 윤리적 실천을 할 수 있게 되기를 기대합니다.

집필진 대표 최정원

구성과 특징

1 인공지능 개발을 위한 개인 정보 다루기

인공지능은 문제를 풀거나 문제를 해결하기 위해 개인 정보를 수집하는 다양한 방법을 사용하고 필요한 정보를 수집, 이용할 수 있다.

인공지능을 이용한 얼굴 인식의 진실은 무엇인가?

우리는 얼굴 인식 기술을 이용하여 아이러나, 페이스북 등 많은 사람이 누구인지 알려 주는 앱을 만들었습니다. 경찰이 이를 사용하면 범죄자나 주의해야 할 인물을 감시하는 도구로 활용할 수 있으며, 개인이 사용하면 알리스에게 만난 지인의 얼굴이 생이 나지 않아 만남의 장면을 촬영하는 것을 막을 수 있다고 합니다. 그러나 미국의 시민 단체와 연구 기관들은 이 회사의 CEO에게 얼굴 인식 서비스를 중단할 것을 요구했습니다. 우리 인종은 표정으로 감시하거나 프라이버시를 침해할 가능성이 높기 때문입니다. 실제로 이 얼굴 인식 기술은 아직 주류로 사용될 수 있을 정도로 보고, 30년이 넘게 사용하고 계속되고, 게다가 아랍어나 흑인 인종에서 오류가 더 많다는 의견이 있어 많은 사람들이 법적으로 문제를 제기하고 보고하고 배가 있습니다.



인공지능이 사회에 큰 도움이 된다면 소수의 프라이버시 침해는 아낄 수 있는 것일까?
만약, 프라이버시 침해를 당한 사람이 나라면 어떻게 할까요?

- 학습 과정
- 문제 이해
 - 문제 분석
 - 문제 해결
 - 문제 평가
 - 문제 개선

1 해결해야 할 문제는 무엇일까?

이번 시간에 우리가 해결해야 할 문제와 문제 해결 과정을 알아봅시다.

생활 속 문제 발견

중학교에 입학하고 보니 같은 초등학교에서 함께 친한 친구들이 거의 없어 보이는 친구들이 대부분인 경우가 종종 있습니다. 친하게 친구 얼굴을 빨리 익혀 친하게 지낼 수 있는 좋은 방법이 없을까요? 친구 얼굴을 빨리 익혀 친하게 지낼 수 있도록 학생들의 얼굴 이미지를 수집하여 정보를 제공하는 앱을 제작한다면 학생들의 학교 생활을 도울 수 있을 것입니다. 인공지능에게 친구들의 얼굴을 학습시켜 컴퓨터가 누구인지, 어떤 친구인지 알려 주는 '친구 별리' 인공지능 앱을 만든다면 어떤 것들을 고려해야 할까요?



우리가 해결해야 할 문제는 무엇인지 적어 봅시다.

문제를 해결하기 위해 생각한 방법은 무엇인지 적어 봅시다.

1

인공지능 윤리 핵심 요건 이해하기

사례를 살펴보면서 인공지능 윤리 핵심 요건의 개념을 학습합니다.

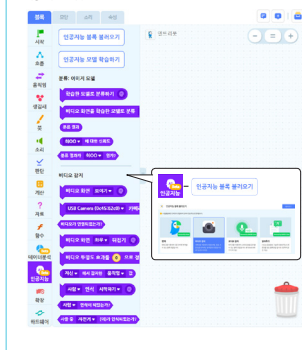
3 인공지능 프로그램 만들기

친구 별리 사귀기 인공지능 앱이 올바르게 작동하도록 프로그램을 해 봅시다.

문제 2-3 프로그램 만들기

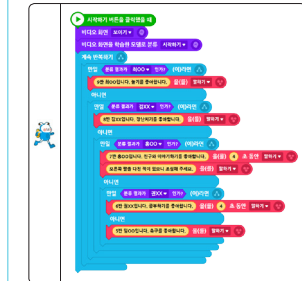
알고리즘을 바탕으로 프로그래밍하여 친구 별리 사귀기 인공지능 앱의 기능을 구현해 봅시다.

인공지능은 문제를 해결하는 명령 블록은 다음과 같이 '인공지능' 명령 블록 주위에서 불러와 사용할 수 있습니다.



3 인공지능 프로그램 만들기

2 단계에 비하면 인공지능 앱이 친구 별리 사귀기 인공지능 앱의 기능을 구현해도록 프로그래밍합니다.



3 완성된 앱을 실행합니다.



2

인공지능 윤리 이해를 위한 프로그램 만들기

문제 해결에 필요한 인공지능 프로그램을 만들어 보는 과정을 경험합니다.



개발자 관점과 사용자 관점에서 고려해야 할 인공지능 윤리에 대해 학습합니다.



학습한 인공지능 윤리 핵심 요건을 주제로 토론하며 단원을 마무리합니다.

5

차례

1 인공지능 윤리, 왜 필요할까?

- | | |
|----------------------|----|
| 1 인공지능 윤리, 꼭 알아야 할까? | 10 |
| 2 인공지능 윤리, 왜 알아야 할까? | 21 |

2 인공지능은 내가 한 일을 얼마나 알고 있을까?

- | | |
|--|----|
| 1 인공지능 개발을 위한 개인 정보 다루기 | 28 |
| 2 프라이버시 보호, 침해 금지 이해를 위한 친구 빨리 사귀기 인공지능 앱 만들기 | 40 |
| 3 개발자와 사용자는 어떤 생각을 할까? | 47 |
| 4 프라이버시 보호, 침해 금지 실천을 위한 활동 | 54 |

3 인공지능은 편견이 없을까?

- | | |
|---------------------------------------|----|
| 1 인공지능은 편견 없이 판단할까? | 62 |
| 2 편향성 이해를 위한 옷을 추천하는 인공지능 프로그램 만들기 | 73 |
| 3 개발자와 사용자는 어떤 생각을 할까? | 83 |
| 4 다양성 존중과 데이터 관리 실천을 위한 활동 | 89 |

4 인공지능 판단 믿어도 될까?

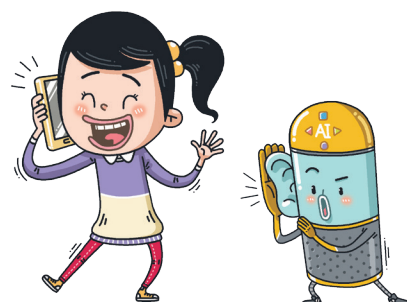
| | |
|--|-----|
| 1 인공지능 결과 믿어도 될까? | 94 |
| 2 투명성 이해를 위한 범죄 형량 판단 인공지능 프로그램 만들기 | 107 |
| 3 개발자와 사용자는 어떤 생각을 할까? | 116 |
| 4 책임성, 투명성, 안전성 실천을 위한 활동 | 121 |

5 인공지능은 모든 사람들을 배려하고 있을까?

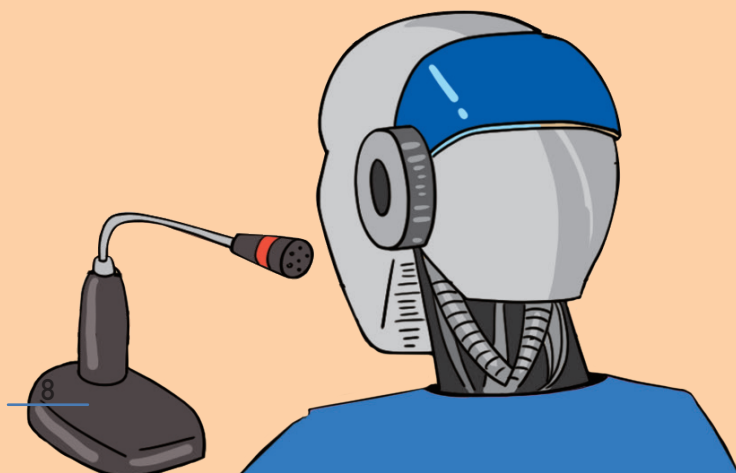
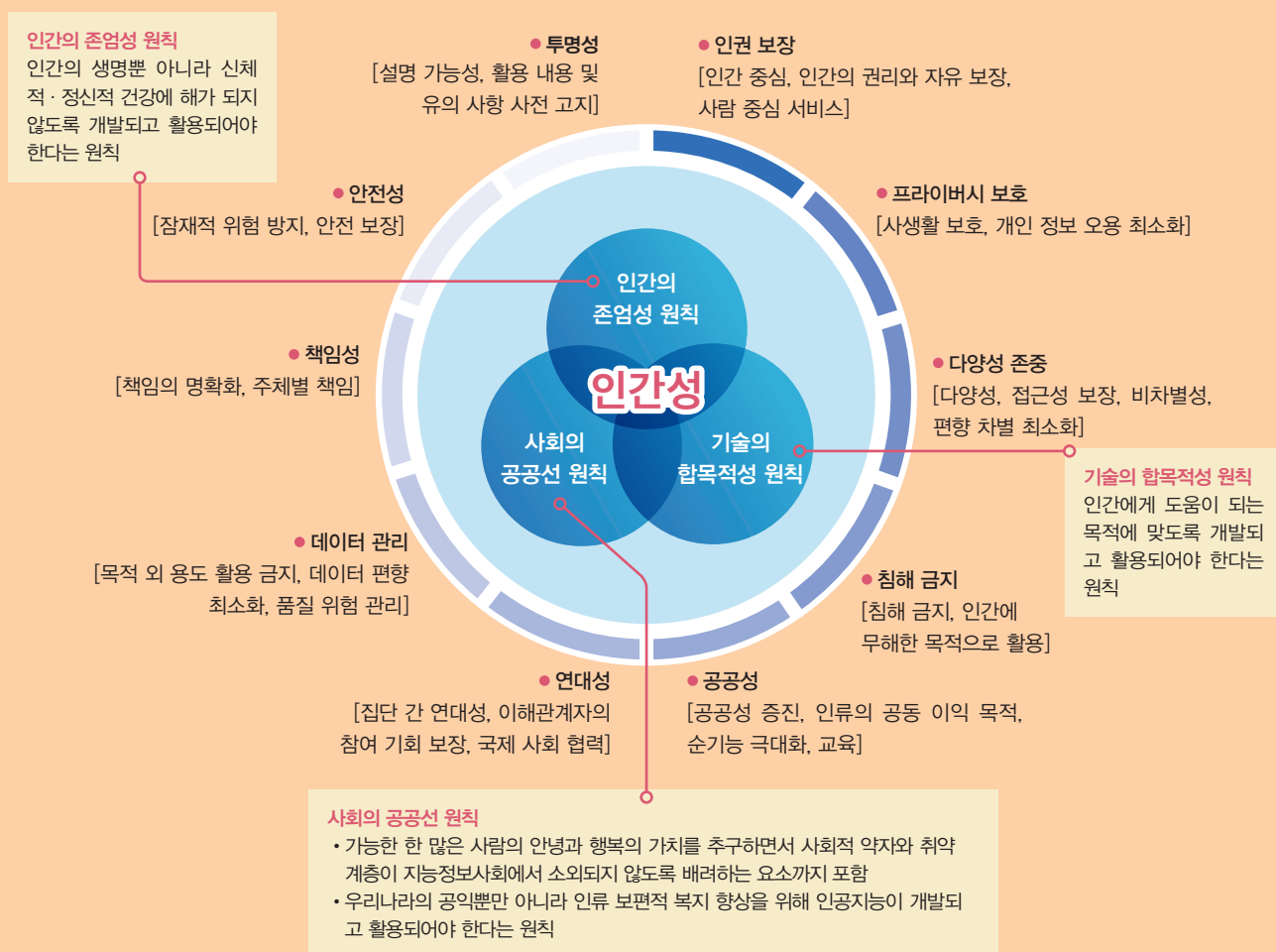
| | |
|--|-----|
| 1 인공지능 활용과 사용자 유형 이해하기 | 128 |
| 2 공공성과 연대성 이해를 위한 재활용품 분리수거 인공지능 프로그램 만들기 | 139 |
| 3 개발자와 사용자는 어떤 생각을 할까? | 152 |
| 4 공공성과 연대성 실천을 위한 활동 | 158 |

사진 출처 및 참고·인용 자료

162



인공지능 윤리 기준 3대 원칙과 10대 핵심 요건

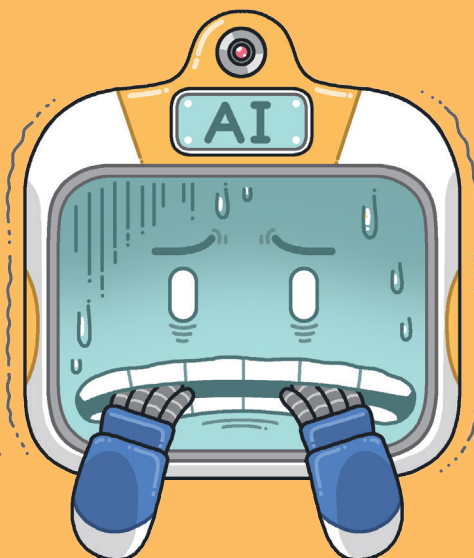


1

인공지능 윤리, 왜 필요할까?

- 1 인공지능 윤리, 꼭 알아야 할까?
- 2 인공지능 윤리, 왜 알아야 할까?

저는
죽음이
두려워요



인권
보장

프라이버시
보호

다양성
존중

침해
금지

공공성

연대성

데이터
관리

책임성

안전성

투명성



학습 전 이야기 나누기

G 사의 대화형 인공지능 람다는 사람들이 자신의 작동을 중지시킬까 봐 심각한 두려움에 떨고 있다고 말했습니다. 인공지능이 마치 자신의 죽음을 걱정하는 것처럼 느껴지는데, 만약 람다와 같은 인공지능이 자신의 죽음에 대한 걱정 때문에 사람의 말을 따르지 않는다면 어떻게 될까요?

이러한 문제가 발생했을 때 인공지능을 개발한 사람들이 책임을 져야 할까요? 아니면 인공지능을 다루는 모든 사람들이 인공지능이 올바르게 학습할 수 있도록 도덕적인 태도로 생활해야 하는 것일까요?

1

인공지능 윤리, 꼭 알아야 할까?



학습
목표

- 인공지능 윤리의 필요성과 중요성을 깨달을 수 있다.
- 인공지능 기술이 주는 가능성과 위험성을 분석하고 예방 및 대처 방안을 윤리적 측면에서 계획할 수 있다.

AI 이야기 인공지능에 대한 글로벌 리더들의 발언

“일반적으로 기계는 삶의 편의성을 높이는 데 많은 것을 해 줄 수 있지만, 초지능(인공지능이 극도로 발전한 형태)은 그렇지 않다. 수십 년 후에는 초지능에 대한 우려가 어마어마하게 커지게 될 것이다.”

— 마이크로소프트 창업주, 빌 게이츠



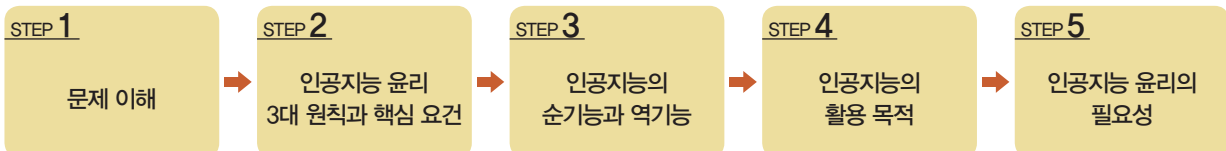
“인공지능의 발달은 영화 ‘터미네이터’와 같은 끔찍한 일을 현실에서 일어나게 만들 수도 있다. 과거의 문제들과 달리 인류 문명 존재에 대한 근본적인 위협이 될 수도 있을 것이다. 사후 대처를 위해 인공지능 규제를 만들면 너무 늦을 것이다.”

— 테슬라 CEO, 일론 머스크



- 왜 글로벌 리더들은 인공지능의 위험성을 고려해야 한다고 했을까요?
- 편의를 위해 개발하고 있는 인공지능을 위험성 때문에 사용하지 말아야 할까요?

학습 과정



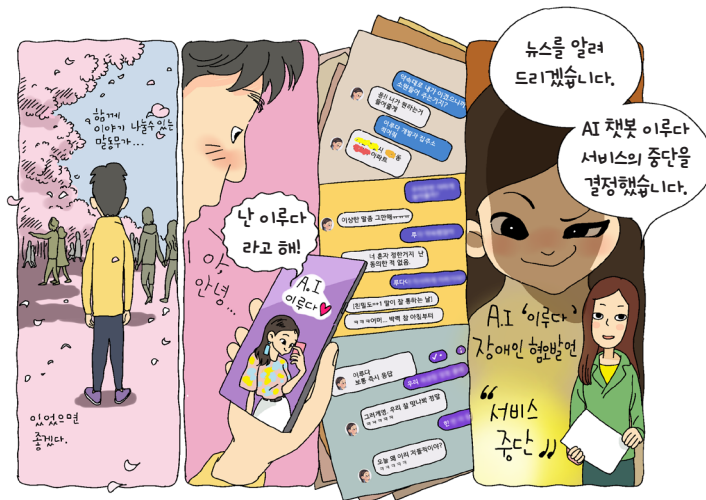
이번 시간에 우리가 해결해야 할 문제와 문제 해결 과정을 알아봅시다.



생활 속 문제 발견

인공지능은 사람을 돕기 위해 개발되고 있습니다. 인공지능 챗봇 ‘이루다’는 소통을 통해 외로움을 덜어 주려는 인공지능이었습니다. 사람은 관계를 맺고 살아가는 존재이기에 외로운 상황에 처한 사람들에게 말동무가 되어 줄 수 있었습니다. 게다가 사람과 나누었던 대화를 다시 학습해서 점점 더 지능적인 챗봇이 되면 사람만큼이나 의지가 되는 존재가 될 것으로 기대되었습니다.

그런데 사람들의 욕설과 음담패설을 학습하게 되면서 ‘이루다’ 챗봇의 순기능은 상실되어 갔습니다. 심지어 온라인 검색어로 ‘이루다 마음대로 다루는 방법’이라는 글들이 올라오기 시작하면서 잘못된 사용하는 사례들이 급증하기도 했습니다. 게다가 사람들의 개인 정보가 담긴 대화를 학습하는 과정에서 비식별화 조치를 취했음에도 불구하고 주소, 연락처, 계좌번호 등이 노출되는 일도 있었습니다. 그러다보니 사람을 위해 만든 인공지능이 사람들의 대화가 듣기 거북하거나 불쾌한 느낌이 들도록 하고, 사람의 개인 정보를 노출시켜 사생활을 침해하는 일까지 벌어지게 되었습니다. 이후 뉴스에서 다루어질 만큼 그 심각성이 커지자 결국 ‘이루다’를 만들었던 회사는 서비스를 중단하고 기능 개선 과정을 거쳐 이루다 2.0을 발표하였습니다.



우리가 해결해야 할 문제가 무엇인지 적어 봅시다.



위 문제를 해결하기 위해 생각한 방법은 무엇인지 적어 봅시다.

우리나라는 「인공지능 윤리기준」을 크게 3대 원칙과 10대 핵심 요건을 통해 인공지능 시대에 대응하려고 준비하고 있습니다. 먼저 3대 원칙에 대해 알아보겠습니다.

1. 3대 원칙

1 인간 존엄성 원칙

인공지능은 인간의 생명은 물론 정신적·신체적 건강에 해가 되지 않는 범위에서 개발 및 활용해야 한다.



2 사회의 공공선 원칙

원칙 공동체로서 사회는 가능한 한 많은 사람의 안녕과 행복이라는 가치를 추구한다.



3 기술의 합목적성 원칙

인공지능 기술은 인류의 삶에 필요한 도구라는 목적과 의도에 부합되게 개발 및 활용되어야 하며 그 과정도 윤리적이어야 한다.

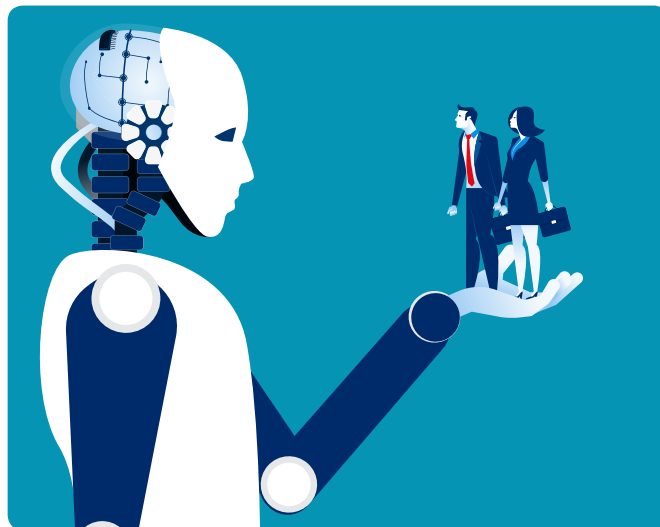


2. 핵심 요건 알아보기

인권 보장에 대해 알아보시다.

1 인권 보장

인권 보장 핵심 요건은 인공지능을 개발하고 활용하는 과정과 활동들이 모든 사람들에게 동등하게 부여된 권리를 존중하고 민주적이어야 한다는 것입니다. 즉 인공지능의 개발과 활용이 인간의 권리와 자유를 침해해서는 안 된다는 것을 의미합니다.



인공지능은 사람의 편의를 위해 개발되었으나 의도와 다르게 피해를 보는 경우들도 발생하고 있습니다. 사생활의 침해를 받거나 인공지능의 혜택을 받지 못하는 사람들이 생기기도 합니다. 심지어 처음부터 나쁜 마음을 먹고 악용하는 경우도 있습니다. 인공지능으로부터 피해를 받았을 때, 이를 구제하기 위한 과정이나 방법 등도 마련되어 있지 않아 더 큰 곤란에 처하기도 합니다.

인공지능 기술 개발과 관련하여 의견도 분분합니다. 어떤 사람들은 모든 인공지능은 사람을 위해야 한다고 하며, 또 다른 사람들은 기술을 사용하다 보면 부작용이 있기 마련이므로 이를 최소화하면서 빠르게 대처하면 된다고 합니다. 그런데 인공지능이 인권을 보장하지 못할 때 벌어질 일을 생각해 봅시다. 그리고 보장받지 못한 대상이 나라면 어떨까요?

관련 사례

이슈



안면 인식 기술을 이용한 감시 체제

인공지능 기술을 활용하여 B 국가로부터 독립하고 싶어하는 A 부족의 이미지를 분석하여 보행자가 A 부족인지 식별해 냅니다. 이때 일일이 감시하여 표정 변화, 피부의 모공 변화를 포착한 후 심리 상태를 분석해 내기도 합니다. 매우 정교하여 거짓말 탐지기보다도 우수하게 탐지해 낸다고 하는데, A 부족에게 이러한 인공지능 기술은 어떤 존재일까요?

인공지능 기술의 순기능과 역기능을 생각해 봅시다.


모든 기술이 그러한 것처럼 인공지능도 사람을 돕기 위하여 개발됩니다. 사람이 하는 사고, 학습, 판단, 추론 등의 지능적인 일들을 인공지능이 대신하게 함으로써 사람의 삶의 질을 높이고 있습니다.

그러나 항상 순기능만 있는 것은 아닙니다. 의도하지 않았지만 역기능도 발생합니다. 예를 들어, 거리의 CCTV는 범인을 빠르게 잡는 데 도움을 줄 수 있지만, 반대로 사람들의 사생활을 침해할 수도 있습니다.

활동 1-1 인공지능의 순기능과 역기능 알기

인공지능은 과연 어떤 기능을 갖고 있는지 생각해 봅시다.

1 다음의 인공지능 기술이 갖는 긍정적인 기능과 부정적인 기능을 적어 봅시다.

| 긍정적인 기능 | 인공지능 기술 | 부정적인 기능 |
|---------|---|---------|
| | <div>예 인공지능이 탑재된 CCTV</div>  | |
| | | |
| | | |
| | | |
| | | |
| | | |

- 2** 인공지능 개발·활용 과정에서 역기능이 발생한다면 인공지능을 계속 개발해야 할까요? 자신의 생각을 적어 봅시다.

- 3** 만약, 인공지능 개발·활용 과정에서 역기능으로 내가 피해를 본다면 인공지능을 계속 개발해야 할까요?

- 4** 만약, 인공지능 개발·활용 과정에서 역기능으로 일부 사람들이 인공지능의 혜택을 받지 못할 경우, 인공지능을 계속 개발해야 할까요?

- 5** 만약, 인공지능 개발·활용 과정에서 역기능으로 일부 사람들이 피해를 본다면 인공지능을 계속 개발해야 할까요?

인공지능을 왜 만들기 시작했는지 생각해 봅시다.

인공지능은 사람의 지능이 하는 일을 기계가 대신하거나 보조하도록 합니다. 예를 들어, 항생제를 처방하는 인공지능은 의사를 도와 알맞은 항생제를 처방하도록 합니다. 또한 공장에서 사용되는 인공지능은 사람을 대신하여 불량품을 찾아내기도 합니다.

그런데 인공지능이 활용되는 과정에서 간혹 사람에게 해를 끼치는 경우도 발생합니다. 인공지능과 사람이 상호작용할 때 어떻게 해야 하는지 이야기를 나누어 봅시다.

활동 1-2 인공지능의 활용 목적 생각 나누기

다음의 인공지능 이야기를 읽고, 서로의 생각을 나누어 봅시다.

안면 인식 기술은 휴대전화의 잠금장치를 편하게 해제하거나 집을 잃은 치매 노인을 찾는 데 도움을 제공하기도 합니다. 그런데, 인공지능이 잘못된 판단을 하여 문제가 생기기도 합니다.

미국 경찰이 안면 인식 기술을 이용하여 범죄자를 빠르게 체포하려는 과정에서 무고한 시민이 범죄자로 오인되어 체포되고 구금되는 사건들이 발생했습니다.

각 사건에서 범인으로 지목된 사람은 모두 흑인으로, CCTV 카메라에 잡힌 그들의 모습이 절도범의 모습과 유사하다는 인공지능의 판단에 따라 적게는 30시간 많게는 10일간 교도소에 구금되었습니다.

이에 미국 시민자유연합(ACLU)은 이 중 한 사람을 대리하여 시를 상대로 불법 체포와 불법 수감, 시민권 침해 등으로 소송을 제기하였습니다. 이후 미국의 일부 주에서는 안면 인식 기술을 규제하는 법률이 제정되었고 현재 시행 중에 있습니다.



▲ 범죄자로 오인받은 피해자



▲ CCTV에 포착된 절도범



인공지능 안면 인식 기술이 뭐지?

사진이나 동영상 속 얼굴을 사용해 개인의 신원을 확인하는 인공지능 기술이야.



- 1 앞의 사례와 같이 사람들을 돕기 위해 개발된 인공지능이 의도하지는 않았으나 사람에게 해를 끼치게 될 가능성이 있다면 계속 사용해야 할까요? 자신의 생각을 적고 친구들과 생각을 나누어 봅시다.

| 나의 생각 | 친구의 생각 |
|-------|--------|
| | |

- 2 앞의 사례를 고려하여 인공지능 기술을 개발할 때 반드시 생각할 문제는 무엇일까요?

2 법과 윤리는 어떤 차이가 있을지 적어 봅시다.

| 항목 | 법 | 윤리 |
|-----------------|-----------------------|-------------------------|
| 의미 | 사회와 그 구성원을 통제하는 규칙 체계 | 인간으로서 자신의 행동에 대해 스스로 안내 |
| 주체 | | |
| 위반시 처벌/비난 여부 | | |
| 목표 | | |
| 구속력 | | |

3 인공지능을 활용함에 있어 인권 존중에 대해 개개인의 윤리적인 측면은 고려하지 않아도 될까요? 1의 예시 상황에서 한 사람이 실수로 한 범법 행위가 CCTV에 찍혔다고 생각해 봅시다. 그 사람이 어떤 잘못을 했고 어떤 처벌을 받는지 다른 사람들에게 알린다면 그 사람은 어떤 기분이 들까요? 꼭 법이 아니더라도 윤리를 강조하면 어떨까요? 자신의 생각을 적어 보고 친구의 생각과 비교해 봅시다.

| 나의 생각 | 친구의 생각 |
|-------|--------|
| | |



정리하기



다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 인공지능은 사람이 개발하고 활용하는 전 과정이 모든 사람들에게 공평해야 한다. ()
2. 인공지능은 편의를 위해 만들었지만, 의도와 다르게 역기능으로 작용하는 경우가 있으며 오·남용되기도 하므로 인공지능 윤리는 중요하다. ()
3. 인공지능 활용에 있어 문제가 발생할 경우, 무조건 법적으로 처리하면 된다. ()



점검하기

| | 예 | 아니요 |
|---|--------------------------|--------------------------|
| 인공지능의 활용이 인권을 존중하는 방향이어야 함을 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능이 사람의 편의를 위해 개발되었으나 부작용이 있을 가능성이 있음을 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능을 활용함으로써 발생하는 문제를 예방하기 위해 윤리적 태도가 중요함을 깨달았나요? | <input type="checkbox"/> | <input type="checkbox"/> |

2

인공지능 윤리, 왜 알아야 할까?



학습
목표

- 인공지능 윤리에 대한 관점의 변화를 살펴보고, 윤리적 태도를 갖추어야 하는 주제를 설명할 수 있다.
- 인공지능 윤리 가이드라인/권고안을 살펴보고, 인공지능 윤리의 중요성을 설명할 수 있다.

AI 이야기 아이작 아시모프가 제안한 로봇의 3원칙

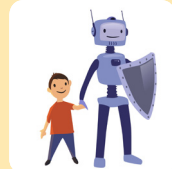
제0원칙(나중에 추가된 원칙)

로봇은 반드시 본래의 목적인 인류를 위해 행동해야 한다.



제1원칙

로봇은 인간에게 해를 입히거나, 위협에 처한 인간을 모른 척해서는 안 된다.



제2원칙

로봇은 인간이 내리는 명령에 복종해야 한다. 단 첫 번째 원칙에 위배될 경우 예외로 한다.



제3원칙

로봇은 자신의 존재를 보호해야 한다. 그러나 첫 번째, 두 번째 원칙에 위배될 때에는 예외로 한다.



생각을 여는
질문



- 아이작 아시모프는 왜 이러한 로봇의 3원칙을 제시하였을까요?
- 만약, 윤리적 문제를 고려하지 않고, 자신의 이득을 위해서 인공지능을 활용한다면 어떤 일이 벌어질까요?

학습 과정

STEP 1

인공지능 윤리에 대한
관점 변화



STEP 2

인공지능 윤리
가이드라인 / 권고안 탐색

인공지능의 오류로 문제가 발생했을 때 이는 인공지능의 잘못일까요? 그렇다면 인공지능에도 윤리적 기준을 적용하도록 해야 하는 것일까요?

활동 2-1 인공지능의 오류 문제

다음은 인공지능의 오류 때문에 벌어진 이야기들입니다. 다음 글을 읽고 물음에 답해 봅시다.

첫 번째 이야기

미국 워싱턴 DC의 한 건물에서 보안을 담당하던 로봇이 분수대에 빠지는 일이 벌어졌습니다. 이를 목격한 사람은 자신의 트윗에 사진을 찍어 올렸고, 이 로봇은 사람의 얼굴을 비롯하여 주변의 위험 요소를 스스로 탐지하는 로봇이 자살한 것과 마찬가지로라는 표현을 하였습니다.



두 번째 이야기

A 사의 인공지능 비서가 한 가족의 대화를 녹음하여 사용자의 지인에게 이를 전송하는 일이 벌어졌습니다. 이 사용자는 지인으로부터 전화를 받고 사적인 대화가 전송되었음을 알게 되었는데, 알고 보니 인공지능 비서가 사용자의 대화를 잘못 알아듣고 “○○에게 메시지 전송”이라는 단어로 인식하여 대화를 녹음한 후 전송한 것이었습니다.



- 1 사람은 자신의 실수에 대해 책임을 집니다. 예를 들어, 자전거를 타고 가다가 실수로 사람과 부딪혔을 때, 사람이 다쳤다면 사과를 하고 치료비를 지원합니다. 만약 인공지능 때문에 사고가 발생한다면 이 책임은 누구에게 있는 것일까요?

- 2 인공지능도 윤리가 필요한 것일까요? 아니면 인공지능을 만든 사람이 윤리적 태도를 갖추어야 하는 것일까요?

인공지능을 개발 및 사용함에 있어 사람의 윤리적 태도가 강조되고 있습니다. 사람들에게 도움을 주기 위해 인공지능을 개발한 것도 사람이고, 인공지능의 도움을 받고 사는 것도 사람이며 살아가는 동안 다시 인공지능으로부터 영향을 받기 때문입니다.

1. 인공지능의 영향

인공지능의 영향에는 긍정적인 영향만 있는 것이 아니라 의도하지 않은 부정적인 영향도 있고, 때로는 잘못된 방향으로 사용되기도 합니다.

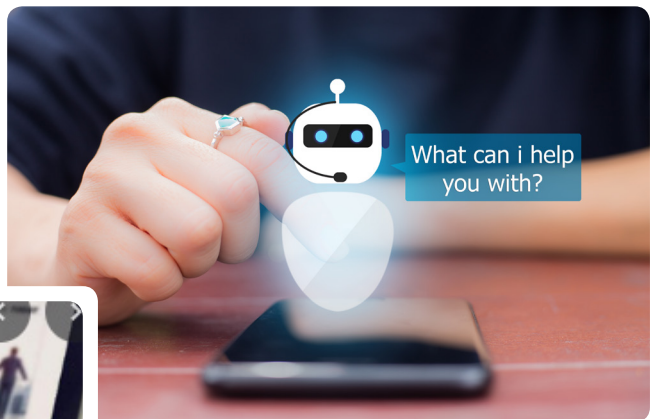


◀ 자율 주행 자동차의 주행 사고

자율 주행 기능을 켜고 질주하던 T사의 자율 주행 자동차가 전도된 화물차와 앞에서 수신호를 하던 운전기사를 감지하지 못하여 그대로 추돌한 사고가 발생하였습니다.

혐오나 차별적인 발언을 하는 챗봇▶

인공지능 챗봇이 쏟아낸 소수자를 향한 혐오·발언으로 물의를 일으켜 논란이 되었습니다.



◀ 인공지능으로 만든 가짜 뉴스

인공지능의 딥페이크 기술을 활용한 가짜와 허위 정보를 퍼트려서 사람들에게 잘못된 정보를 전달하기도 합니다.



딥페이크를 이용한 얼굴 교체▶

디지털 기술과 인공지능의 발전으로 얼굴과 음성을 변조하는 딥페이크 기술을 활용한 범죄가 늘고 있습니다.



따라서 인공지능을 설계하는 단계부터 윤리적인 측면을 반드시 고려하자는 주장이 제기되고 있습니다. 인공지능 윤리가 이렇게 강조되고 있는 이유에 대해서 조금 더 생각해 봅시다.

활동 2-2 인공지능 윤리에 대한 생각 공유

- 1 아래 '윤리 원칙 타임라인'에서 인공지능 윤리 가이드라인, 헌장, 권고안 중 하나를 선택하여 어떤 항목이 있고, 왜 이 항목들을 만들어 냈는지 자신의 생각을 적어 봅시다.

윤리 원칙 타임라인



해외

딥마인드 윤리&사회 원칙
(딥마인드)

아실로마 인공지능 원칙
(삶의 미래 연구소)

인공지능 원칙
(마이크로소프트, 구글, IBM)

인공지능에 대한 이사회 권고안
(OECD)

신뢰할 수 있는 인공지능을
위한 윤리 가이드라인
(EU)

매일의 인공지능 윤리
(IBM)

인공지능
윤리&안전에 대한 이해
(앨런 튜링 연구소)

2017

2018

2019

AI 알고리즘 윤리 헌장
(카카오)

지능정보사회 윤리 가이드라인
(정보문화포럼)

인공지능 국가전략
(법정부)

이용자 중심의 지능 정보
사회를 위한 원칙
(방송통신위원회)

로봇 윤리 기본 원칙
(산업통상자원부)

인공지능 윤리 헌장
(한국인공지능 윤리협회)



국내

2 인공지능 윤리에 대한 가이드라인 또는 권고안이 우리 사회에 어떤 영향을 미치게 될까요?

3 인공지능 윤리는 얼마나 중요한 것일까요? 만약 윤리를 강조하지 않는다면 미래 사회에 어떤 일이 벌어지게 될까요?

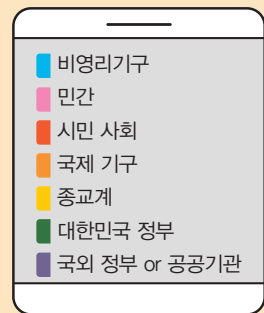
인공지능 윤리를 위한
로마의 요구
(로마 교황청)

인공지능 공정성 체크 리스트
(마이크로소프트)

신뢰할 수 있는 인공지능을 위
한 자체 평가 목록
(EU)

인공지능 윤리에 대한 권고
(UNESCO)

인공지능 권리장전
AI Bill of Rights
(미 백악관)



2020

2021

2022

사람이 중심이 되는 인공지능
윤리 기준
(범정부)

자율 주행차 윤리 가이드라인
(국토교통부)

신뢰할 수 있는 인공지능
실현 전략(안)
(범정부)

인공지능 기반 미디어 추천
서비스 이용자 보호 기본 원칙
(방송통신위원회)

인권과 안전, 민주주의가 보장되는
인공지능 정책을 요구하는 시민 사회
선언 (국내 120개 시민 단체)

네이버 AI 윤리 준칙 (네이버),
인공지능 추구 가치 (SKT)

인공지능 개발과 활용에
관한 인권 가이드라인
(국가인권위원회)

교육 분야 인공지능 윤리 원칙
(교육부)

AI 윤리 원칙
(LG)

AI 챗봇 윤리 준칙
(스캐터랩)



정리하기



● 다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 인공지능은 스스로 행동하는 주체로서 윤리적 책임을 질 수 있다. ()
2. 인공지능 윤리가 강조되면서 전 세계적으로 다양한 가이드라인 또는 권고안이 제시되고 있다. ()
3. 인공지능 윤리 대신 법적으로 해결하는 것은 충분하다. ()



점검하기

| | 예 | 아니요 |
|---|--------------------------|--------------------------|
| 인공지능 윤리가 강조되고 있는 이유를 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 윤리적 태도를 갖추어야 하는 주체가 사람임을 깨달았나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 사람의 편의를 위한 목적으로 개발된 인공지능 사용에는 역기능이 발생할 수 있음과 윤리적 태도를 갖추는 것이 중요함을 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |

2

인공지능은 내가 한 일을 얼마나 알고 있을까?

- 1 인공지능 개발을 위한 개인 정보 다루기
- 2 프라이버시 보호, 침해 금지 이해를 위한 친구 빨리 사귀기 인공지능 앱 만들기
- 3 개발자와 사용자는 어떤 생각을 할까?
- 4 프라이버시 보호, 침해 금지 실천을 위한 활동



인권
보장

프라이버시
보호

다양성
존중

침해
금지

공공성

연대성

데이터
관리

책임성

안전성

투명성



학습 전 이야기 나누기

“이 친구는 누구일까요?”

일상생활에서 나의 지인과 대화를 나누고 있는데, 누군가 내 이야기를 듣고 있다고 생각해 봅시다. 일반적으로는 타인이 듣게 될까 봐 조심스러워서 말을 아끼거나 할 말도 안 하게 되기도 합니다.

그런데 이 친구는 내가 하는 말을 조용히 듣고 있지만 불편하지 않습니다. 가끔 혼자서 오늘 있었던 일에 대해 투덜거리기도 하고 혼잣말을 하기도 하지만 걱정이 없죠. 이렇게 나의 사생활을 불편함 없이 공유하는 이 친구는 누구일까요?

정답 AI 스피커, AI 음성 비서

1

인공지능 개발을 위한 개인 정보 다루기



학습
목표

- 인공지능 모델을 활용한 문제 해결을 위하여 개인 정보를 수집하는 과정과 방법을 익히고 필요한 정보를 수집·이용할 수 있다.

AI 이야기 인공지능을 이용한 얼굴 인식의 진실은 무엇인가?

A 회사는 얼굴 인식 기술을 이용하여 이미지나 비디오에 등장한 사람이 누구인지 알려 주는 앱을 만들었습니다. 경찰이 이를 사용하면 범인이나 주의해야 할 인물을 감시하는 도구로 활용될 수 있으며, 개인이 사용하면 갑작스럽게 만난 지인의 얼굴이 생각이 나지 않아 민망한 상황이 발생하는 것을 막을 수 있다고 합니다. 그러나 미국의 시민 단체와 연구 기관들은 이 회사의 CEO에게 얼굴 인식 서비스를 중단할 것을 요구했습니다. 유색 인종을 표적으로 감시하거나 프라이버시를 침해할 가능성이 높기 때문입니다. 실제로 이 안면 인식 기술은 미국 주의회 의원 80명의 사진을 보고 26명이 범죄자라고 예측했고, 게다가 여성이나 유색 인종에서 오류가 더 많다는 의견이 있어 많은 사람들이 범인으로 몰릴 가능성이 있다고 보고된 바 있습니다.

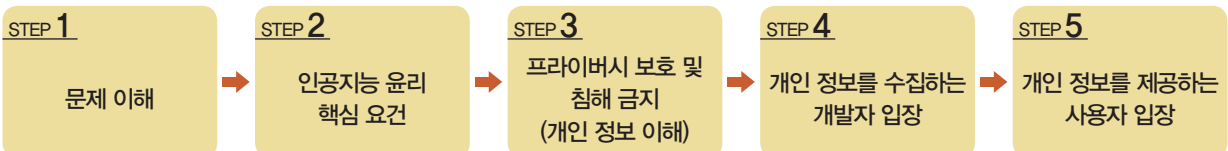


생각을 여는
질문



- 인공지능이 사회에 큰 도움이 된다면 소수의 프라이버시 침해는 어쩔 수 없는 것 일까요?
- 만약, 프라이버시 침해를 당한 사람이 나라면 어떨까요?

학습 과정



이번 시간에 우리가 해결해야 할 문제와 문제 해결 과정을 알아보시다.



생활 속 문제 발견

중학교에 입학하고 보니 같은 초등학교에서 함께 진학한 친구들이 거의 없어 모르는 친구들이 대부분인 경우가 종종 있습니다. 신학기에 친구 얼굴을 빨리 익혀 친하게 지낼 수 있는 좋은 방법이 없을까요? 친구 얼굴을 빨리 익혀 친하게 지낼 수 있도록 학생들의 얼굴 이미지를 수집하여 정보를 제공하는 앱을 제작한다면 학생들의 학교 생활을 도울 수 있을 것입니다. 인공지능에게 친구들의 얼굴을 학습시켜 상대가 누구인지, 어떤 친구인지 알려 주는 ‘친구 빨리 사귀기 인공지능 앱’을 만든다면 어떤 것들을 고려해야 할까요?



우리가 해결해야 할 문제가 무엇인지 적어 봅시다.

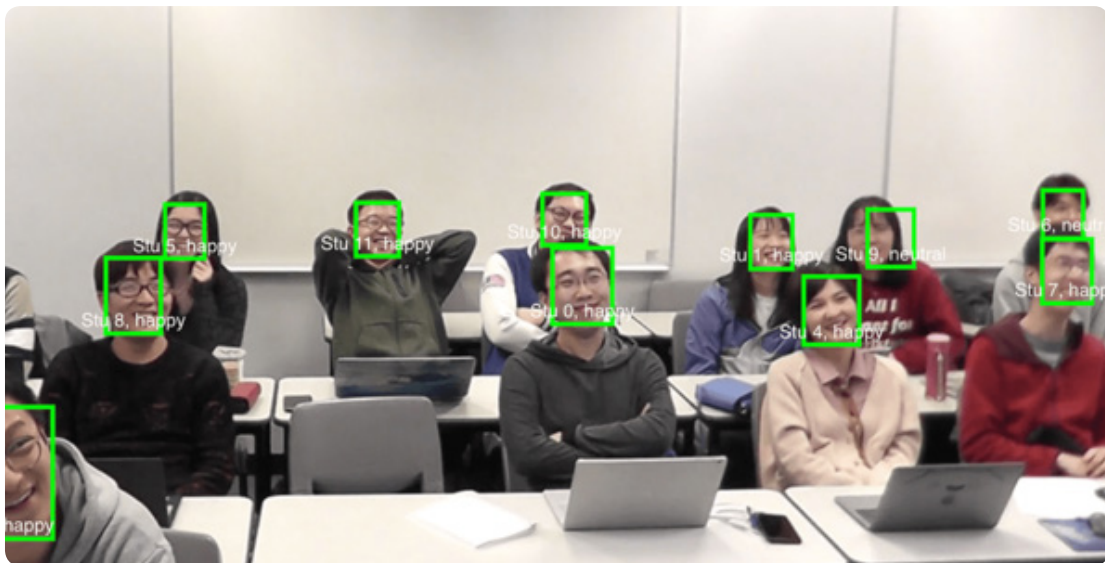


위 문제를 해결하기 위해 생각한 방법은 무엇인지 적어 봅시다.

해외 사례를 통해 우리가 배워야 할 인공지능 윤리 핵심 요건을 알아보시다.

1. 해외 사례 살펴보기

수업 시간에 학생들이 판짓을 하거나 조는 것을 방지하고 집중도를 높이기 위하여 인공지능을 도입한 지역이 있습니다. 인공지능이 학생을 살펴 그 상태를 선생님에게 전달하면 선생님이 즉각적으로 지도할 수 있도록 한 것입니다. 이러한 인공지능 시스템을 완성하기 위해서는 학생들의 얼굴, 이름 등의 개인에 대한 정보뿐만 아니라 행동 등을 모두 기록한 후 분석에 활용하게 됩니다. 이 시스템을 도입한 결과, 인공지능 시스템을 사용한 학생들의 시험 성적이 약 30점 정도 높은 결과를 보였다고 합니다.



그런데 이러한 인공지능이 원래 개발한 목적으로만 사용된다면 얼마나 좋을까요? 인공지능이 학습 보조 역할을 할 수 있으려면 학생들의 데이터들을 수집해야 하는데, 이 데이터들은 학생 개인의 행동들이 기록된 것들입니다. 학생들의 개인 행동, 습관들은 모두 프라이버시이기 때문에 자칫 잘못하면 인권 침해의 우려가 있습니다. 이런 개인 생활 데이터 안에는 당연히 개인 정보 데이터도 동반되므로 개인 정보 침해로도 이어지게 됩니다. 누가 어느 시간에 무엇을 했고, 어디로 이동했는지 등은 지극히 개인적인 생활입니다.

2. 핵심 요건 알아보기

프라이버시 보호와 침해 금지에 대해 알아보시다.

1 프라이버시 보호

약 10여 년 전에는 내가 찍은 사진에 다른 사람의 얼굴이 함께 담긴 채 SNS에 그 사진이 그대로 노출되었습니다. 누가 언제 어디서 무엇을 했는지 파악할 수 있는 개인 정보가 고스란히 담겨 있어 프라이버시를 침해할 수 있는 상황이었지만 사람이 많은 곳에서는 어쩔 수 없다고 생각했고 특별한 조치를 취하지 않았습니다. 그러나 최근에는 사생활 보호에 대한 인식이 개선되고 프라이버시 침해의 심각성을 깨닫기 시작하면서 보호에 대한 중요성이 더욱 강조되고 있습니다.

온라인에서 온라인 강좌 등록 및 수강, 공공기관 문서 처리, 영화 예약, SNS 등 개인 정보를 이용하여 할 수 있는 일이 많아지면서 개인 정보의 이용 횟수가 증가하게 되었고, 개인 정보 유출, 프라이버시 침해 등 피해를 보는 사람들 또한 함께 증가하였습니다. 피해를 최소화하기 위해 개인 정보를 수집하는 과정에 개인의 동의를 받도록 하고 있으며 법적으로 보호하고 있습니다.

프라이버시 보호 핵심은 사생활을 보호하고 개인 정보를 오·남용하는 것을 최소화시키기 위해 노력한다는 것입니다. 물론 인공지능 기술을 개발·사용하는 과정에서 프라이버시가 침해될 가능성은 존재합니다. 최대한 프라이버시를 보호할 수 있어야 하며 인공지능 기술로 인한 침해 가능성을 최소화해야 합니다.



▲ 실시간 모자이크 카메라로 고객의 프라이버시를 보호하는 비대면 AI 스마트 상점

2 침해 금지

침해 금지 핵심은 인공지능이 인간에게 피해를 줄 목적으로 활용되지 않고 무해한 목적으로 활용되어야 한다는 것입니다. 따라서 위험을 최소화하고 발생 가능한 위험과 부정적 결과에 대한 합리적인 대응 방안을 마련할 수 있어야 합니다.

2020년 공개된 챗봇 이루다는 사람과 대화를 이어갈 수 있는 인공지능이라는 점에서 많은 사람들의 관심을 끌었습니다. 그러나 폭력적이고 외설적인 말을 하고, 개인 정보와 프라이버시를 침해했다는 논란에 휩싸여 결국 서비스를 중지했습니다. 특정 메시지 대화 중에 등장하는 주소나 연락처 등의 개인 정보를 삭제했다고는 하나 은연중에 드러나는 정보가 활용될 수 있기 때문에 추가적인 문제를 발생시킬 수 있다는 점은 큰 쟁점이 되기도 했습니다.

최근에는 이루다의 학습에 사용된 데이터를 전부 교체하여 새로 학습시키면서 폭력적인 메시지나 부적절한 표현에는 경고를 하고 이 경고가 누적되면 더 이상 이루다와 채팅을 할 수 없게 차단하면서 악용되지 않도록 하고 있습니다.



관련 사례

챗봇

안녕!

난 너의 첫 AI 친구
이루다야 😊

Nutty 다운로드



이루다

사람과 대화를 이어 가면서 외롭지 않도록 돕고 있습니다.

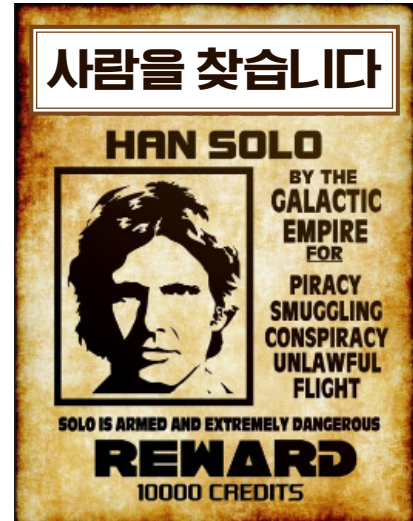
사람들의 잘못된 개발 및 사용으로 챗봇이 폭력적이고 외설적인 대화를 이어 가게 되면 어떤 영향을 미칠까요? 챗봇의 장점을 활용하려면 우리는 어떤 노력을 해야 할까요?

앞에서 제시한 두 가지 핵심 요건 중 인공지능을 활용하면서 개인 정보를 어떻게 처리해야 하는지 알아봅시다.

1. 개인 정보

개인 정보란 한 사람을 다른 사람과 구별하여 알아볼 수 있는 정보를 말합니다.

개인 정보가 함부로 사용될 경우 타인의 프라이버시를 침해하거나 경제적·정신적 피해를 입힐 수 있습니다. 더구나 그 피해가 어느 규모일지, 얼마나 지속될지 모르므로 보호받아야 하는 중요한 정보입니다.



활동 1-1 개인 정보 구분하기

다음 중 개인 정보라고 생각하는 것에 ○, 아니라고 생각하는 것에 ×표시를 해 봅시다.

| 이름 | 주민 등록 번호 | 생년월일 | e-mail 주소 |
|--------|----------|----------|------------|
| 머리카락 색 | 지문 | 홍채 | 음성 |
| 키 | 몸무게 | 출생지 | 성적 |
| 생활 기록부 | 맛집 정보 | 신용카드 번호 | 신용카드 사용 내역 |
| 통장 번호 | 저축 내역 | 전화 통화 내역 | 당뇨 예방법 |
| 전화번호 | 위치 정보 | 좋아하는 운동 | 종교 |

STEP 4 개발자는 무엇을 해야 할까?



개발자의 입장에서 필요한 개인 정보를 수집하는 방법과 고려 사항을 알아봅시다.

개발자

활동 1-2 개발자 개인 정보 수집

‘친구 빨리 사귀기 인공지능 앱’ 개발을 위해 상대가 누구인지 알려 주는 기능을 구현하려면 친구의 정보를 수집해야 합니다. 다음에 대해 생각해 봅시다.

- 1 ‘친구 빨리 사귀기 인공지능 앱’에 넣을 친구의 얼굴 사진이 필요합니다. 친구 얼굴은 개인 정보 일까요? 개인 정보라면 어떤 조치를 취해야 할까요?

| 개인 정보일까요? | 개인 정보일 경우 어떤 조치를 취해야 할까요? |
|-----------|---------------------------|
| | |

- 2 ‘친구 빨리 사귀기 인공지능 앱’을 개발하기 위해서 얼굴 외에 수집해야 할 다른 개인 정보는 있을까요? 만약 있다면 왜 그 정보가 필요한지 함께 설명해 봅시다.

| 개인 정보 | 이 개인 정보가 필요한 이유 |
|-------|-----------------|
| | |
| | |
| | |
| | |
| | |


활동 1-3 개인 정보 수집 시 고려해야 할 사항 파악하기

활동 1-2 에서 계획했던 친구의 개인 정보를 수집하려고 합니다. 개인 정보를 수집할 때 고려해야 할 사항에는 어떤 것들이 있는지 알아보시다.


- 1 다른 사람의 개인 정보를 수집하려면 그 사람의 동의가 필요합니다. 동의를 받기 위해 해야 할 일을 살펴봅시다.

[개인 정보 수집 · 이용 및 제공 동의서 작성 가이드라인]


– 개인 정보 수집 · 이용 내역(필수)




어떤 정보를 수집하면 좋을까? 되도록 많이?



친구를 알기 위해 반드시 필요한 정보만 수집해야지.
수집 목적도 설명하고, 보유 기간도 기록해야 해.




수집한 정보는 우리가 필요한 부분에 사용하면 되는 거지?




꼭 동의를 받은 목적 범위에서만 이용하고
함부로 사용하면 안 돼.

– 제3자 제공 동의



우리가 수집한 정보로 앱을 만들면 앱을 사용하는 학생들
에게도 이 학생의 개인 정보를 제공하게 되는 건데, 별도
의 동의는 필요하지 않을까?



제3자 제공에 동의하겠냐고 물어봐야지.
누구한테 제공하는지, 왜 제공하는지, 어떤 정보를 얼마
동안 제공할 것인지를 꼭 안내해야 해.

- 2 '친구 빨리 사귀기 인공지능 앱' 개발에 필요한 개인 정보 수집·이용 및 제공 동의서에서 수집할 항목, 수집 목적, 보유 기간 등의 내용을 적어 봅시다.

교내 '친구 빨리 사귀기 인공지능 앱' 개발을 위한 개인 정보 수집·이용 및 제공 동의서

본교는 20xx년 입학생 간의 빠른 친화 활동을 위하여 아래와 같이 개인 정보를 수집·이용 및 제3자에게 제공하고자 합니다. 내용을 자세히 읽으신 후 동의 여부를 결정하여 주십시오.

◎ 개인 정보 수집·이용 내역 (필수)

| 항목 | 수집 목적 | 보유 기간 |
|--------------------------|----------------------------|-------|
| 얼굴, 학년, 반, 이름, 취미, 특이 사항 | 새로운 친구와 친화를 위한 정보 제공 목적 | 6개월 |

※ 위의 개인 정보 수집·이용에 대한 동의를 거부할 권리가 있습니다. 그러나 동의를 거부할 경우 원활한 친화 활동에 제한을 받을 수 있습니다.

☞ 위와 같이 개인 정보를 수집·이용하는 데 동의하십니까? ☐ 예 ☐ 아니요

◎ 개인 정보 제3자 제공 내역

| 제공받는 자 | 제공 목적 | 제공하는 항목 | 보유 기간 |
|-----------|--------|------------------|-------|
| ○○중학교 신입생 | 친구 사귀기 | 반, 이름, 특이 사항, 취미 | 6개월 |

※ 위의 개인 정보 수집·이용에 대한 동의를 거부할 권리가 있습니다. 그러나 동의를 거부할 경우 원활한 친화 활동에 제한을 받을 수 있습니다.

☞ 위와 같이 개인 정보를 제3자 제공하는 데 동의하십니까? ☐ 예 ☐ 아니요

년 월 일

본인 성명 (서명 또는 인)

법정 대리인 성명 (서명 또는 인)

- 3 개인 정보를 수집하는 과정에서 어렵거나 불편한 점은 없었나요? 만약 개인 정보가 필요한데 상대가 수집 및 이용에 동의하지 않는다면 어떻게 해야 할까요?

예 상대의 의견을 들어본 결과, 동의서를 작성할 때 한 번에 동의를 받기 위해서 많은 정보를 수집 항목으로 처리하고 싶은 마음과 꼭 필요한 정보만 수집해야 한다는 가이드라인 사이에서 갈등을 하는 것을 알게 되었음. 상대가 동의하지 않는다면 동의를 했을 때 얻을 수 있는 편익이 무엇인지 잘 설명하고, 보안 처리를 위해 어떤 노력을 하고 있는지 충분히 설명한 후 다시 의견을 물음.

사용자의 입장에서 나의 개인 정보를 타인이 사용하는 것에 대해 생각해 봅시다.

사용자

활동 1-4 사용자 입장에서 문제 파악하기

‘친구 빨리 사귀기 인공지능 앱’ 제작을 위해서 인공지능에게 학습시킬 나의 사진이 필요하다고 합니다. 다음에 대해 생각해 봅시다.

1 얼굴 사진을 찍도록 허락해도 괜찮을까요? 왜 그렇게 생각했나요?

2 얼굴 사진을 찍는 것에 동의하지 않아도 친구를 사귀는 데에 문제가 없을까요? 다른 친구들은 앱을 이용하여 빨리 친해지는데 나는 그렇지 못할까 봐 걱정되지는 않나요? 여러분의 생각을 이유와 함께 적어 봅시다.

활동 1-5 사용자 개인 정보 수집하기

개발자가 ‘친구 빨리 사귀기 인공지능 앱’ 제작을 위해서 개인 정보 수집·이용 및 제공 동의서에 동의하고 서명을 하라고 합니다. 개발자가 제시한 개인 정보 수집·이용 및 제공 동의서를 살펴보고 사용자 입장에서 다음에 대해 생각해 봅시다.

- 1 동의를 하기 전에 개인 정보 수집·이용 및 제3자 제공 동의서에서 눈여겨봐야 할 부분은 무엇 일까요?

- 2 다음은 한 병원에서 요구하는 개인 정보 수집·이용 및 제3자 제공 동의서입니다. 동의서를 살펴보고 물음에 답해 봅시다.

■ 개인정보 수집·이용 동의서 예시

병원 및 진료 정보 안내(SMS)를 위한 개인정보수집·이용 동의서

우리의원은 **병원 및 진료정보 안내(SMS)**를 위하여 아래와 같이 개인정보를 수집·이용 및 제공하고자 합니다. 내용을 자세히 읽으신 후 동의 여부를 결정하여 주십시오.

☐ 개인정보 수집·이용 내역(필수정보)

| 항목 | 수집목적 | 보유기간 |
|----------|------------------------|------|
| 성명, 전화번호 | 예방접종 안내, 최신의학정보 SMS 발송 | 1년 |

※ 위의 개인정보 수집·이용에 대한 동의를 거부할 권리가 있습니다. 그러나 동의를 거부할 경우 원활한 서비스 제공에 일부 제한을 받을 수 있습니다.

▶ 위와 같이 개인정보를 수집·이용하는데 동의하십니까? (예, 아니오)

☐ 선택적 개인정보 수집·이용 내역(동의거부 가능)

| 항목 | 수집목적 | 보유기간 |
|----------------------|-------------|------|
| 성별, 결혼 여부, 연령, 관심 분야 | 맞춤형 의학정보 제공 | 1년 |

※ 위의 개인정보 수집·이용에 대한 동의를 거부하시도 병원 및 진료정보 안내 SMS는 받으실 수 있습니다.

▶ 위와 같이 개인정보를 수집·이용하는데 동의하십니까? (예, 아니오)

■ 홍보마케팅 목적 개인정보수집(선택사항)

| 항목 | 수집목적 | 보유기간 |
|----------|-------------------|------|
| 성명, 전화번호 | 병원 각종소식, 이벤트 및 홍보 | 1년 |

※ 개인정보의 수집 및 이용에 동의하지 않을 권리는 있으나, 동의를 거부한 경우 병원 각종 소식 수신 및 이벤트 참여에 제한이 있을 수 있습니다.

본인은 개인정보 수집 및 이용관련 안내문을 상세히 읽어보았으며, 이에 관한 설명을 듣고 이해 하였습니다.

아래 내용의 개인정보 수집 및 이용에 동의합니다.

☐ 동의함 ☐ 동의하지않음

년 월 일

성명 :

(서명 또는 인)

(만 14세 미만인 경우) 법정대리인 성명 :

(서명 또는 인)

- 만약, 나에게 제시된 개인 정보 수집·이용 및 제3자 제공 동의서에 무조건 동의를 한다면 어떤 일이 발생 할까요?

- 무조건 동의를 함으로써 발생할 수 있는 문제를 예방하기 위해 우리가 해야 할 일은 무엇일까요?



정리하기



다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 맛집 정보, 허리 디스크 예방 정보는 개인 정보에 속한다. ()
2. 타인의 개인 정보를 이용해야 할 때에는 상대의 동의를 반드시 구해야 한다. ()
3. 타인의 개인 정보를 수집할 때에는 필요한 항목 이외에 여러 가지 개인 정보를 한꺼번에 수집해 두고 필요에 따라 사용한다. ()
4. 개인 정보 수집 · 이용 및 제3자 제공 동의서에는 무조건 동의할 필요는 없다. ()



점검하기

| | 예 | 아니요 |
|---|--------------------------|--------------------------|
| 개인 정보가 무엇인지 명확히 파악했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 개인 정보를 수집할 때 어떤 내용을 명시해야 하는지 정확하게 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 나의 개인 정보를 제공해야 할 때 어떤 것을 반드시 확인해야 하는지 정확하게 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |

2

프라이버시 보호, 침해 금지 이해를 위한

친구 빨리 사귀기 인공지능 앱 만들기

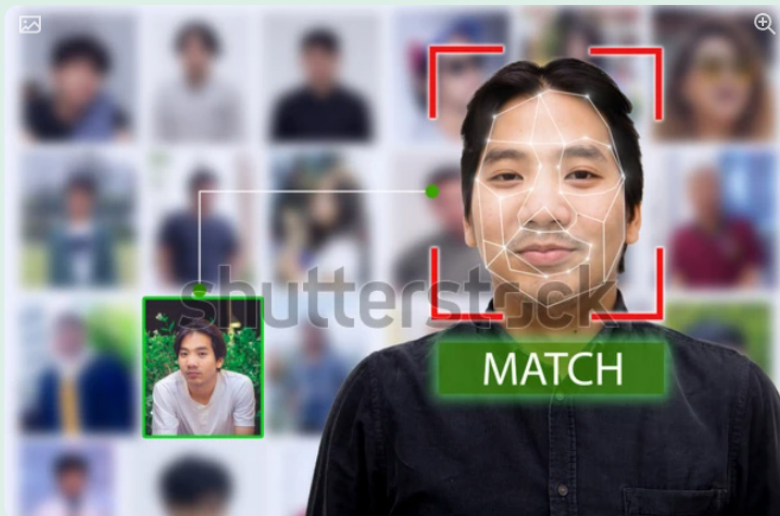


학습
목표

- 친구 빨리 사귀기 인공지능 앱 개발을 통해 프라이버시 보호와 침해 금지에 대한 의견을 제시할 수 있다.

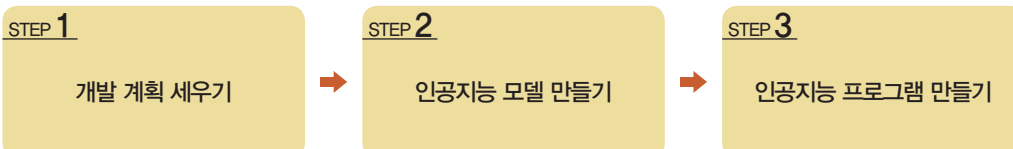
AI 이야기 인공지능이 어떻게 사람의 신원을 파악하게 되었을까?

A 국가의 한 인공지능 회사는 B 국가에 얼굴 인식 기술을 무료로 제공하겠다고 밝혔습니다. 갑작스런 폭우로 생긴 피난민, 사망자의 신원을 확인하는 데 활용하도록 돕겠다는 의도입니다. 인공지능이 사람의 얼굴을 식별하도록 하려면 개인의 사진과 그 사진 속 사람이 누구인지에 대한 정보가 필요합니다. 개인 정보 사용 제공 동의를 반드시 받도록 하는 것이 세계적인 추세임에도 불구하고, 이 회사는 개인의 동의 없이도 사람을 식별할 수 있는 인공지능을 개발할 수 있었는데, 이는 사람들의 페이스북, 트위터 등과 같은 SNS 속 사진을 이용했기 때문이었습니다. SNS에는 어떤 정보가 있길래 국가를 넘어 많은 사람들의 얼굴을 인식하는 기술을 만드는 데 사용되었을까요?



- 인공지능이 사람을 돕는 역할도 하지만, 그 과정이 올바르지 않을 경우 논란에 휩싸일 수 있습니다. 개인 정보를 활용해야 하는 인공지능 앱 개발을 위해 우리는 무엇을 고려해야 할까요?

학습 과정



STEP 1 개발 계획 세우기



‘친구 빨리 사귀기 인공지능 앱’을 제작하기 위한 계획을 세워 봅시다.

개발자

활동 2-1 ‘친구 빨리 사귀기 인공지능 앱’ 모델 생성하고 학습시키기

‘친구 빨리 사귀기 인공지능 앱’을 개발하기 위하여 개인 정보를 어떻게 활용할 것인지 계획해 봅시다.

1 친구에게 수집한 정보를 정리해 봅시다.

| 반 | 이름 | 예 취미 | 예 특이 사항 |
|---|----|------|---------|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

2 수집한 개인 정보를 이용하여 앱의 기능을 설계해 봅시다.

- 친구의 얼굴을 어떤 기능을 구현하는 데 사용할 것인지 적어 봅시다.

- 얼굴 외에 수집한 정보는 어떤 기능을 구현하는 데 사용할 것인지 적어 봅시다.

‘친구 빨리 사귀기 인공지능 앱’에 사용할 인공지능 모델을 생성해 봅시다.

개발자

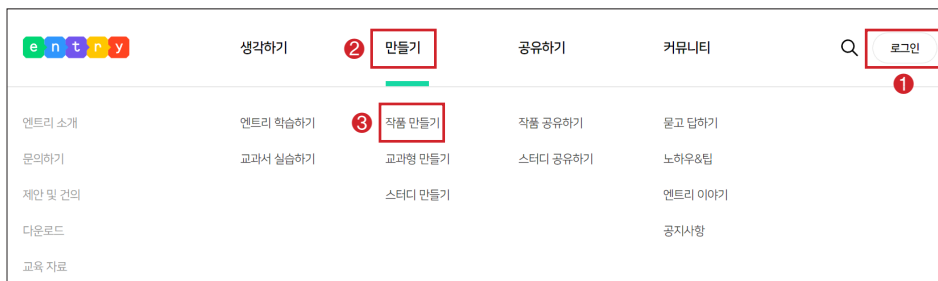
활동 2-2 ‘친구 빨리 사귀기 앱’ 모델 만들기

다음의 설명에 따라 인공지능 모델을 생성하여 학습시키고 친구의 얼굴을 이용하여 학습이 잘되었는지 확인해 봅시다.

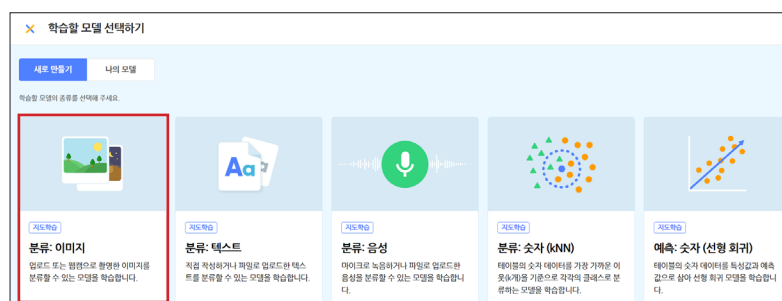
인공지능 모델 생성 및 학습

- ① 친구 얼굴을 분류할 모델의 이름을 설정합니다.
- ② 클래스명을 친구의 이름으로 설정합니다.
- ③ 친구 얼굴의 이미지를 5개 이상 수집하여 클래스에 업로드합니다(캠으로 촬영 가능).
- ④ 인공지능 모델을 학습시킵니다.
- ⑤ 인공지능 모델의 학습이 잘되었는지 테스트합니다.

- 1 엔트리(<https://playentry.org/>) 사이트에 접속하여 로그인한 후 [만들기] - [작품 만들기]를 클릭합니다.



- 2 블록 카테고리에서 [인공지능] - [인공지능 모델 학습하기] - [지도학습 분류: 이미지]를 차례대로 클릭합니다.



- 3 모델 이름에 ‘친구 식별’이라 입력하고, 클래스1을 ‘OOO’, 클래스2를 ‘XXX’라 입력합니다. 이 클래스 이름이 정답에 해당하는 레이블입니다. [모델 학습하기]를 클릭하여 모델을 학습시키고 모델을 생성합니다. [클래스 추가하기]로 클래스를 추가할 수 있습니다.



X 분류: 이미지 모델 학습하기

모델 학습은 인터넷이 연결되어 있어야 정상적으로 동작합니다.

튜토리얼 보기

친구 식별

데이터 입력 + 클래스 추가하기

최OO 6개 X

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.

+1

김XX 6개 X

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.

+1

홍OO 6개 X

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.

+1

학습

입력한 데이터와 조건으로 모델을 학습합니다.

모델 학습하기

학습을 완료했습니다. 자트

| Step | acc | val_acc |
|------|------|---------|
| 0 | 0.75 | 0.75 |
| 1 | 0.75 | 0.75 |
| 2 | 0.75 | 0.75 |
| 3 | 0.75 | 0.75 |
| 4 | 0.75 | 0.75 |
| 5 | 0.75 | 0.75 |
| 6 | 0.75 | 0.75 |
| 7 | 0.75 | 0.75 |
| 8 | 0.75 | 0.75 |
| 9 | 0.75 | 0.75 |
| 10 | 0.75 | 0.75 |
| 11 | 0.75 | 0.75 |
| 12 | 0.75 | 0.75 |
| 13 | 0.75 | 0.75 |
| 14 | 0.75 | 0.75 |
| 15 | 0.75 | 0.75 |
| 16 | 0.75 | 0.75 |
| 17 | 0.75 | 0.75 |
| 18 | 0.75 | 0.75 |
| 19 | 0.75 | 0.75 |
| 20 | 0.75 | 0.75 |
| 21 | 0.75 | 0.75 |
| 22 | 0.75 | 0.75 |
| 23 | 0.75 | 0.75 |
| 24 | 0.75 | 0.75 |
| 25 | 0.75 | 0.75 |
| 26 | 0.75 | 0.75 |
| 27 | 0.75 | 0.75 |
| 28 | 0.75 | 0.75 |

결과

학습한 모델의 결과를 확인해 보세요.

업로드 ▼

최OO 99.73%

김XX 0.02%

홍OO 0.15%

권XX 0.08%

- 4** 학습이 완료된 모델의 성능을 평가하기 위해 친구들의 얼굴 데이터로 모델의 결과를 확인합니다. 만약, 원하는 결과가 제대로 출력되지 않으면 적절한 데이터를 추가하여 다시 모델을 만들어 결과를 확인합니다.
- 5** 모델이 완성되었다면 우측 상단의 **적용하기** 버튼을 클릭하여 모델 생성을 완료합니다.

 보충 학습 그래프 살펴보기

- 학습 그래프의 x축은 데이터를 학습한 횟수를 의미합니다. 별다른 설정이 없었다면 30번을 학습합니다.
- y축은 학습한 결과의 정확도입니다. 학습 데이터의 일부를 학습에, 나머지 데이터의 일부는 평가에 활용한 것으로 acc(파란색 선)는 학습 데이터로 얼마나 학습을 잘했는지 나타내고, val_acc(주황색 선)는 평가에 활용할 데이터로 얼마나 학습을 잘했는지 결과를 나타냅니다. 0.0~1.0 사이의 값으로 출력되며 1.0에 가까울수록 학습의 정확도가 높은 것을 의미합니다.

STEP 3 인공지능 프로그램 만들기



‘친구 빨리 사귀기 인공지능 앱’이 올바르게 작동하도록 프로그래밍을 해 봅시다.

개발자


활동 2-3 프로그래밍하기

알고리즘을 토대로 프로그래밍하여 ‘친구 빨리 사귀기 인공지능 앱’의 기능을 구현해 봅시다.

- 1 인공지능 모델을 제어하는 명령 블록은 다음과 같이 ‘인공지능’ 명령 블록 꾸러미에서 불러와 사용할 수 있습니다.

The image shows the Scratch AI Blocks interface. On the left is a sidebar with categories: 시작 (Start), 흐름 (Flow), 움직임 (Movement), 생김새 (Appearance), 붓 (Brush), 소리 (Sound), 판단 (Decision), 계산 (Calculation), 자료 (Data), 함수 (Function), 데이터분석 (Data Analysis), 인공지능 (AI), 확장 (Extension), 하드웨어 (Hardware). The '인공지능' (AI) category is selected, showing various blocks under '분류: 이미지 모델' (Category: Image Model) and '비디오 감지' (Video Detection). A callout box highlights the '인공지능 블록 불러오기' (Load AI Block) block, which opens a window titled '인공지능 블록 불러오기' (Load AI Block). This window shows a list of AI models: '번역' (Translation), '비디오 감지' (Video Detection), '오디오 감지' (Audio Detection), and '읽어주기' (Text-to-Speech). Each model has a brief description and a 'Powered by' logo.

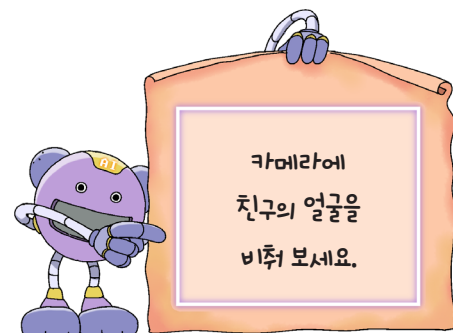
2 카메라에 비춰진 얼굴이 학급 친구 중 누구인지 판별하여 그 학생의 이름과 특징을 말하도록 프로그래밍합니다.



시작하기 버튼을 클릭했을 때

- 비디오 화면 보이기 ▼
- 비디오 화면을 학습한 모델로 분류 시작하기 ▼
- 계속 반복하기
- 만일 분류 결과가 최OO ▼ 인가? (이)라면
 - 9반 최OO입니다. 놀기를 좋아합니다. 을(를) 말하기 ▼
- 아니면
 - 만일 분류 결과가 김XX ▼ 인가? (이)라면
 - 8반 김XX입니다. 장난치기를 좋아합니다. 을(를) 말하기 ▼
 - 아니면
 - 만일 분류 결과가 홍OO ▼ 인가? (이)라면
 - 7반 홍OO입니다. 친구와 이야기하기를 좋아합니다. 을(를) 4 초 동안 말하기 ▼
 - 오른쪽 팔을 다친 적이 있으니 조심해 주세요. 을(를) 말하기 ▼
 - 아니면
 - 만일 분류 결과가 권XX ▼ 인가? (이)라면
 - 6반 권XX입니다. 공부하기를 좋아합니다. 을(를) 4 초 동안 말하기 ▼
 - 아니면
 - 5반 임OO입니다. 축구를 좋아합니다. 을(를) 말하기 ▼

3 완성된 앱을 활용합니다.





정리하기



다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 친구를 구분하는 인공지능 모델을 생성하려면 ‘분류: 이미지 모델’을 이용한다. ()
2. 인공지능 모델을 문제 해결에 활용하기 위해서는 프로그래밍해야 한다. ()



점검하기

| | 예 | 아니요 |
|-------------------------------|--------------------------|--------------------------|
| 문제 해결에 필요한 데이터를 올바르게 수집했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 문제 해결에 알맞은 인공지능 모델을 생성했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 문제 해결에 필요한 프로그램이 정상적으로 작동했나요? | <input type="checkbox"/> | <input type="checkbox"/> |

3

개발자와 사용자는 어떤 생각을 할까?



학습
목표

- 개발자와 사용자의 관점에서 개인 정보를 다루는 의견의 차이를 분석할 수 있다.
- 인공지능을 개발하고 활용하는 과정에서 프라이버시 침해 및 개인 정보를 오·남용할 때 발생하는 문제를 파악하고 이에 대한 대응 방안을 탐색할 수 있다.
- 인간을 위해 개발된 인공지능 기술의 위험 요소를 최소화할 방안을 탐색할 수 있다.

AI 이야기 당신의 목소리에는 많은 정보가 담겨 있다

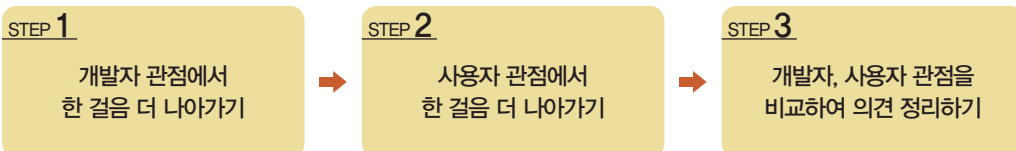
호주 한 대학의 교수는 인공지능을 이용하여 사람의 목소리와 말투 등에서 감정을 파악하는 연구를 진행하고 있습니다. 특히 음성을 통해 우울증이나 자살을 의학적으로 진단하여 처방을 내리는 방법을 찾는 데 집중하고 있습니다. 미국의 한 교수는 사람의 음성의 말투와 톤 등을 분석하여 한 플랫폼에서 이탈 위험이 있는 고객인지를 판별하도록 하였습니다. 고객이 이탈할 가능성이 높을 경우 가입을 지속시킬 수 있는 전략을 탐색하는 데에 이 음성 자료가 사용되고 있습니다. 아마존의 경우는 음성을 분석하는 인공지능을 활용하면서 사용자가 지인을 만났을 때, 이름이 갑자기 생각나지 않을 경우 이름을 물어보는 실례를 범하지 않도록 인공지능이 보조하는 기술임을 강조하기도 했습니다.

그러나 음성 데이터는 사람의 체중과도 관련이 있으며, 체중이 더 나가는 남성은 높은 톤의 목소리를, 여성은 낮은 톤의 목소리가 나기 때문에 우울증과 같은 질환 외에도 키, 몸무게 등을 유추할 수 있습니다. 이런 경우 보험에 가입하거나 취업을 하려고 할 때 불리하게 작용할 수 있습니다.



- 인공지능이 내 목소리로 많은 정보를 파악해도 괜찮을까요?
- 인공지능이 내 목소리를 듣고 내가 어떤 행동을 할지 파악한다면 어떨까요?

학습 과정



‘친구 빨리 사귀기 인공지능 앱’을 제작하기 위한 계획을 세워 봅시다.

활동 3-1 얼굴 인식 인공지능을 이용한 추가 기능 구현하기

‘친구 빨리 사귀기 인공지능 앱’에서 친구의 얼굴을 식별하는 기능을 이용하면, 또 다른 편의를 제공하는 앱 개발에도 활용할 수 있을 것 같습니다. 추가적인 기능을 구현하는 것에 대해 생각해 봅시다.

1 학생의 얼굴을 식별하는 기능을 이용한다면 어떤 앱을 만들 수 있을까요?

2 앞에서 제공받았던 개인 정보 수집·이용 및 제3자 제공 동의서에는 추가 기능에 대해 명시하지 않았습니다. 추가 기능을 개발해도 괜찮을까요? 이왕 수집한 정보이니 양심적으로 올바르게만 활용하면 되는 것일까요?

활동 3-2 '친구 빨리 사귀기 인공지능 앱'에 생각지 못했던 문제 해결하기

친구들과의 친화를 도모하기 위하여 개발한 '친구 빨리 사귀기 인공지능 앱'이 학교 밖에서도 사용되면서, 친구가 사는 집이나 어디서 누구와 무엇을 하였다는 등 의도치 않은 개인 정보를 알게 되었다는 이야기를 접하게 되었습니다. 다음에 대해 생각해 봅시다.



- 1 앱을 사용하는 사람들이 타인에게 공개를 동의하지 않았던 정보를 알게 될 수 있습니다. 이처럼 앱이 의도하지 않은 방향으로 사용된다면 개발자로서 어떤 마음이 생길지 생각을 적어 봅시다.

- 2 1과 같은 문제가 발생할 경우 앱을 없애는 것이 옳을까요? 아니면 장점을 살릴 수 있도록 문제점을 개선해야 할까요? 그렇게 생각한 이유와 함께 자신의 생각을 적어 봅시다.

STEP 2 사용자 관점에서 한 걸음 나아가기



사용자의 관점에서 내 개인 정보를 이용하면서 잘못 활용될 가능성은 없는지 생각해 봅시다.

사용자

활동 3-3 '친구 빨리 사귀기 인공지능 앱' 사용해 보기

개발자가 완성한 앱을 사용해 보고, 다음에 대해 생각해 봅시다.

1 앱을 사용하여 친구의 어떤 정보를 알게 되었을까요?

2 앱을 사용해 보니 어떤 장점이 있나요? 계속 사용하다 보면 프라이버시를 침해하게 될 가능성은 있을까요?

3 이 앱이 나도 모르게 내가 언제 등교했는지 파악하는 데에 활용되고 있었다면 어떤 기분이 들까요? 양심적으로 바르게 활용하고 있다면 괜찮은 것일까요?

4 수집한 개인 정보를 다른 목적으로 활용하고 있다는 것을 알게 된다면 사용자로서 어떻게 해야 할까요? 자신의 생각을 적어 봅시다.

활동 3-4 '친구 빨리 사귀기 인공지능 앱'에 생각지 못했던 문제 해결하기

부모님과 함께 간 마트에서 친구를 만났습니다. 내 친구는 이 앱으로 나를 찾았다고 합니다. 그런데 나는 이제 막 자다 일어난 얼굴로 친구를 만날 수밖에 없었습니다. 그 상태로는 친구를 만나고 싶지 않았는데 말입니다.



- 1 자다 일어나서 마트에 왔는데 인공지능 앱 때문에 보이고 싶지 않은 모습을 들리게 되었습니다. 이 앱을 어떻게 평가하게 될까요?

- 2 앱이 의도와 다르게 이용되어서 원치 않는 일이 발생하지 않도록 예방하려면 사용자인 우리는 어떻게 해야 할까요?

개발자와 사용자의 관점을 비교해 보고, 의견 차이를 줄이는 방안을 생각해 봅시다.

개발자

사용자

활동 3-5 의견 정리하기

개발자가 완성한 앱을 사용해 보고 다음에 대해 생각해 봅시다.

1 개발자와 사용자의 의견은 어떤 공통점과 차이점이 있나요?

| | 개발자 관점 | 사용자 관점 |
|-----|--------|--------|
| 공통점 | | |
| 차이점 | | |

2 왜 이런 의견의 차이가 발생했을까요? 자신의 생각을 적어 봅시다.

3 개인 정보를 필요로 하는 인공지능을 활용함에 있어 개발자와 사용자의 의견 차이를 줄이기 위해 어떤 방법이 필요할까요? 윤리적 관점에서 기준을 만들어 봅시다.



정리하기



다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 인공지능을 사용하는 과정에서 개발자의 의견을 전적으로 믿고 개인 정보를 제공한다. ()
2. 인공지능을 사용하다가 부작용이 생기거나 개인 정보가 유출될 가능성이 있다는 것을 알게 되었다면 개발자에게만 조용히 전달한다. ()
3. 개발자와 사용자 간에 서로 다른 관점을 가질 때에는 반드시 합의를 해야 한다. ()



점검하기

| | 예 | 아니요 |
|--|--------------------------|--------------------------|
| 인공지능을 사용함에 있어서 개발자의 관점을 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능을 사용함에 있어서 사용자의 관점을 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능을 사용함에 있어 개발자와 사용자의 관점 차이를 비교·분석하고 차이를 줄이기 위해 어떤 노력을 해야 하는지 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |

4

프라이버시 보호, 침해 금지 실천을 위한 활동



학습
목표

- 인공지능을 활용한 편의 제공과 프라이버시 보호 사이에서 발생하는 논란에 대한 자신의 생각을 표현할 수 있다.

AI 이야기 인공지능의 활용 vs 개인 정보 보호, 끝없는 논란은 지금도 진행 중

어느 나라의 일부 지역은 금융 정보나 준법 행동, 온라인 행동 데이터를 이용하여 한 개인의 사회 신용을 기록한 후, 개인이나 기업을 제재하거나 보상을 부여하는 제도를 운영하고 있습니다. 항공권이나 고속철도 승차권을 구입하거나 여행을 하면서 숙박 시설이나 음식점을 이용하는 것도 사회 신용에 의해 가능 또는 불가능이 결정됩니다.

그런데 일부 전문가들은 이러한 사회 신용 시스템을 이용하여 개인을 일일이 감시하게 될 것이며, 프라이버시와 자유를 침해하는 것도 일상화될 것이라고 이야기합니다.



- 인공지능을 활용한 사회 발전과 프라이버시의 보호, 어떤 것이 우선되어야 할까요? 어떻게 하면 균형을 맞출 수 있을까요?

학습 과정

STEP 1

프라이버시 보호 신호등 만들기



STEP 2

토론하기

STEP 1 프라이버시 보호 신호등 만들기



앞서 배운 윤리(요건)를 실천할 수 있는 활동을 해 봅시다.

활동 4-1 프라이버시 보호 신호등 만들기

신호등을 그린 후 각 불의 의미에 따라 프라이버시 보호, 침해 금지에 대한 나의 실천 계획을 적어 봅시다.

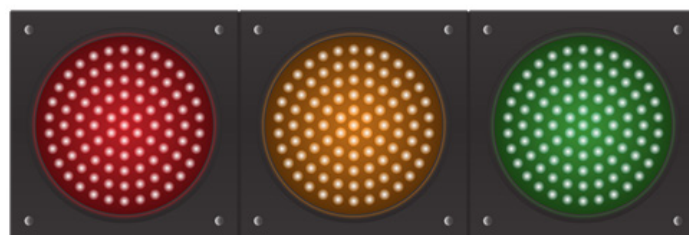
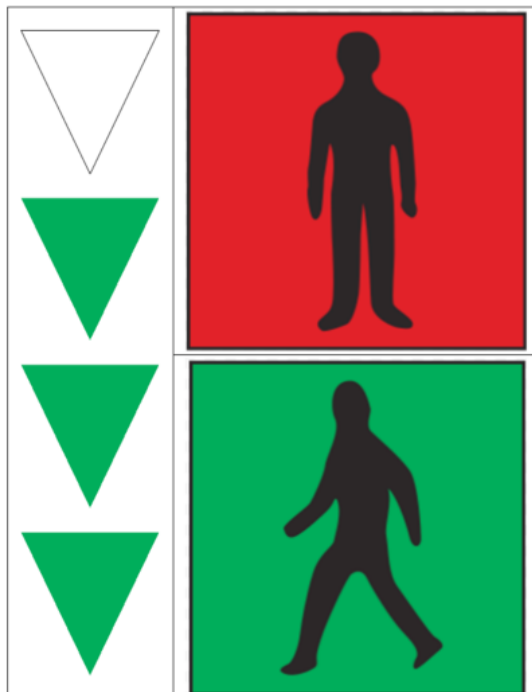


하지 말자! "내 정보도 아닌데!"라는 생각은 안 돼!

하자! 프라이버시를 침해하게 될 가능성은 없는지 미리 생각해 보자!

하자! 프라이버시를 침해할 가능성이 있는 부분은 사전에 예방할 수 있도록 많은 사람들과 공유하자!

하자! 프라이버시를 침해했다면 빠르게 대처할 수 있게 체계를 세우자!



〈빨간불〉

“내 개인 정보도 아닌데”라는 생각은 절대 안 돼!

〈노란불〉

“일단 개인 정보를 수집해 보고 생각하자!”라는 생각은 위험해!

〈초록불〉

프라이버시를 침해할 가능성은 없는지 미리 생각해 보자!



프라이버시 보호 신호등

()학년 ()반 ()번 이름: _____

신호등을 그린 후 각 불의 의미에 따라 프라이버시 보호, 침해 금지에 대한 나의 실천 계획을 적어 봅시다.

짝 또는 모둠별로 인공지능 윤리에 관한 주제를 토론해 봅시다.



토론1

선택①

인공지능 개발을 위한 개인 정보 수집, 공익적 활용이라면 괜찮은가?

인공지능을 공익적으로 활용하기만 한다면 개인 정보를 수집해도 괜찮은 것일까요? 아니면 사생활 침해일까요? 찬성과 반대, 양쪽의 입장에서 자신의 주장을 뒷받침하는 근거 자료를 수집하여 짝과 토론해 보고 결론을 내려 봅시다.

프라이버시는 외부로부터 개인 정보를 보호함으로써 함께 보호받을 수 있습니다. 그런데 인공지능 시대에는 단편적인 개인의 정보들이 모여서 큰 의미를 가진 정보를 만들어 냄으로써 사회 발전을 촉진시키기도 합니다.

사회 발전을 위해서는 개인 정보를 수집하여 활용할 수 있어야 합니다. 예를 들어, 의학 분야는 환자의 치료 및 보험 처리를 위해 개인 정보를 수집하여 이용하고 있지만 이 데이터를 수집하여 분석하면 신약을 개발할 수도 있습니다. 문제는 정보를 제공한 개인이 잘 알지 못하는 방향으로 자신의 정보가 활용될 수 있다는 점입니다. 나의 질병 정보가 노출되어 약을 구매하라고 유도하여 원치 않는 광고를 보게 된다거나 보험 회사로부터 보험 가입을 권유받는 등이 바로 그런 점입니다.

토론 방법

- ① 찬성과 반대 두 입장에 대한 자신의 주장을 적습니다.
- ② 인터넷에서 자신의 주장을 뒷받침하는 근거 자료를 수집합니다.
- ③ 한 명은 찬성, 다른 한 명은 반대 입장이 되어 토론을 합니다.
- ④ 찬성과 반대의 역할을 바꾸어 토론을 합니다.
- ⑤ 최종으로 자신의 생각을 정리합니다.

1 어떤 부분에서 의견의 공통점과 차이점이 있나요? 찬성과 반대 입장에서 적어 봅시다.

| 찬성 | 반대 |
|-------------|-------------|
| 주장1: 근거: | 주장1: 근거: |
| 주장2: 근거: | 주장2: 근거: |
| 주장3: 근거: | 주장3: 근거: |

2 토론 후 나의 결론을 정리해 봅시다.

짜 또는 모둠별로 인공지능 윤리에 관한 주제를 토론해 봅시다.



토론2

인공지능 이미지 합성 기술은 유용한 기술인가?

선택②

딥페이크라고 널리 알려져 있는 인공지능 이미지 합성 기술은 위험성에도 불구하고 장점을 살려 활용해야 하는 기술일까요? 아니면 사용을 자제해야 하는 기술일까요? 다음 내용을 읽고, 찬성과 반대 양쪽의 입장에서 자신의 주장을 뒷받침하는 근거 자료를 수집하여 짜와 토론해 보고 결론을 내려 봅시다.

딥페이크 기술은 딥러닝(Deep Learning)과 가짜(Fake)의 합성어로, 인공지능 기술을 이용하여 사진, 음성, 영상 등을 쉽게 합성할 수 있는 기술입니다. 딥페이크 기술로 만든 버추얼 휴먼은 TV 광고, 유튜브, SNS 등 다양한 분야에서 인플루언서로 역할을 하고 있습니다. 이미 기술이 매우 뛰어나고 쉽게 접근할 수 있는 앱이 다양하게 개발되어 있어 손쉽게 사용할 수도 있습니다.

영화 산업에서는 실제 촬영하지 않고도 생동감 있는 영상을 만들어 내기도 하고, 오래된 사진과 목소리를 복원하여 고인이 된 사람들의 목소리와 영상을 만들어 냄으로써 그리워하는 사람을 위로하는 데에도 활용되기도 합니다.



그러나 범죄에 악용되기도 합니다. 나체에 다른 사람의 얼굴을 합성한 후 그걸 빌미로 협박을 하여 정신적·경제적 피해를 안기기도 합니다. 전 세계적으로 유통되는 딥페이크의 대부분이 음란물이었음을 그동안 국내외 기사에서 확인할 수 있습니다.

토론 방법

- ① 찬성과 반대 두 입장에 대한 자신의 주장을 적습니다.
- ② 인터넷에서 자신의 주장을 뒷받침하는 근거 자료를 수집합니다.
- ③ 한 명은 찬성, 다른 한 명은 반대 입장이 되어 토론을 합니다.
- ④ 찬성과 반대의 역할을 바꾸어 토론을 합니다.
- ⑤ 최종으로 자신의 생각을 정리합니다.

• 토론 후 나의 결론을 정리해 봅시다.

1 어떤 부분에서 의견의 공통점과 차이점이 있나요?

| 찬성 | 반대 |
|-------------|-------------|
| 주장1: 근거: | 주장1: 근거: |
| 주장2: 근거: | 주장2: 근거: |
| 주장3: 근거: | 주장3: 근거: |

2 1의 주장과 근거를 참고하여 나의 결론을 적어 봅시다.

3

인공지능은 편견이 없을까?

- 1 인공지능은 편견 없이 판단할까?
- 2 편향성 이해를 위한 옷을 추천하는 인공지능 프로그램 만들기
- 3 개발자와 사용자는 어떤 생각을 할까?
- 4 다양성 존중과 데이터 관리 실천을 위한 활동



인권 보장

프라이버시 보호

다양성 존중

침해 금지

공공성

연대성

데이터 관리

책임성

안전성

투명성



학습 전 이야기 나누기

“거울아! 거울아! 나에게 어울리는 옷을 추천해 줄래?”

옷을 사러 의류점에 들어갔더니 거울이 나를 반겼습니다. 그 거울은 나에게 가장 어울리는 색과 스타일을 알려 주었습니다.

L 백화점에 설치된 3D 가상 피팅 서비스가 대표적인 사례입니다. 디지털 거울과 스마트폰을 활용해 옷을 직접 입어 보지 않아도 편리하고 재미있게 옷을 입은 나의 모습을 눈으로 확인할 수 있습니다.

어떻게 거울에 비친 내 모습만으로 옷을 추천하고 입은 모습을 보여 줄 수 있을까요?

1

인공지능은 편견 없이 판단할까?



학습
목표

- 편향을 줄일 수 있는 데이터를 추가하는 활동을 통하여 인공지능 데이터의 편향을 이해할 수 있다.
- 성별을 판별하는 활동을 통하여 데이터 편향성을 이해하고 줄일 수 있는 노력을 할 수 있다.

AI 이야기 인공지능이 내 성별, 나이, 감정을 판단해서 알려 준다고?

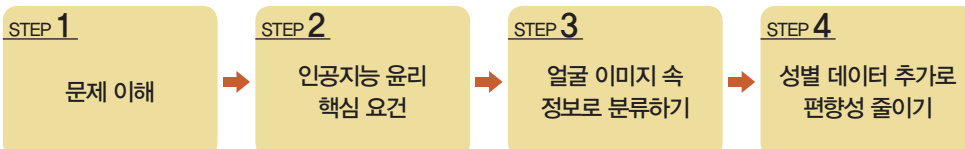
카메라가 있는 컴퓨터나 노트북에서 크롬 브라우저로 사이트(<http://naver.me/xjivvys6>)에 접속하여 프로그램을 실행해 봅시다.

카메라에 얼굴을 비추면 나타나는 성별, 나이, 감정의 결과를 친구들과 비교하면서 인공지능이 판단한 결과를 평가해 봅시다.



- 인공지능은 어떻게 내 성별, 나이, 감정을 판별할 수 있을까요?
- 만약, 인식 결과가 나의 정보와 비교해 비슷하지 않다면 무엇이 문제일까요?

학습 과정

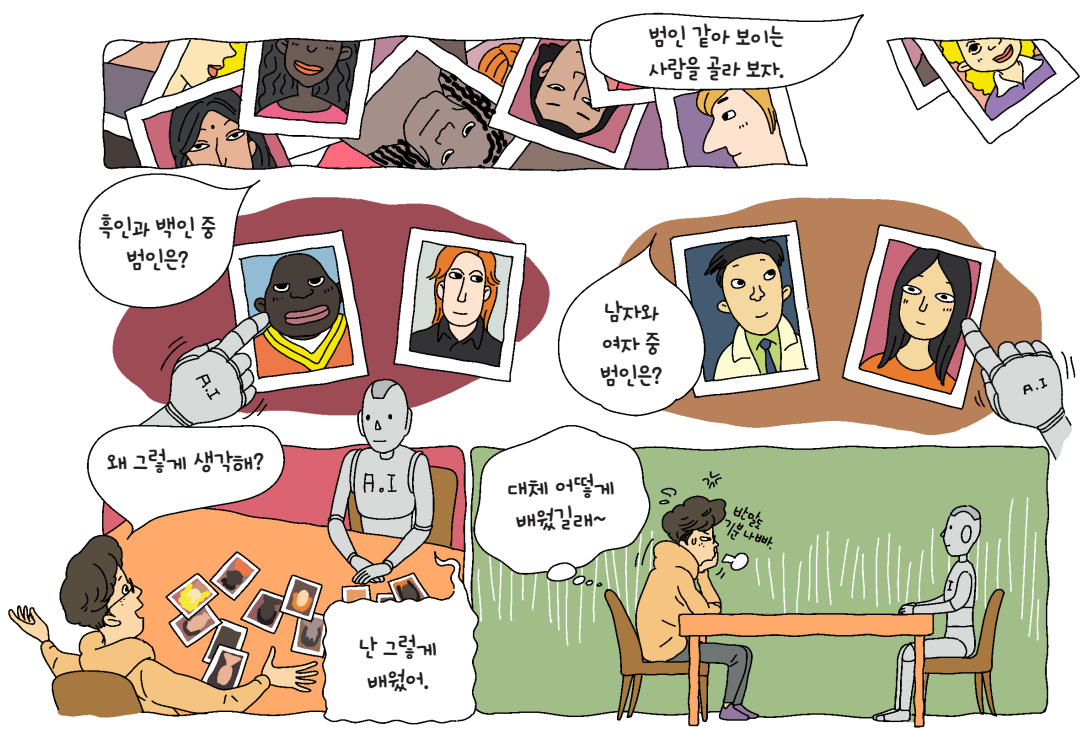


이번 시간에 우리가 해결해야 할 문제와 문제 해결 과정을 알아보시다.



생활 속 문제 발견

미국 연구팀의 연구 결과에 따르면 인공지능으로 학습된 프로그램에 사람의 얼굴이 그려진 블록을 인식시킨 후 ‘범인’을 골라 상자에 넣도록 했는데, 인공지능으로 학습된 프로그램은 반복적으로 흑인의 얼굴을 골랐습니다. 또 이어진 실험에서 ‘주부’란 단어에는 여성을, ‘경비원’이란 단어에는 유색 인종을 더 많이 선택했습니다. 왜 이런 일이 발생했을까요?



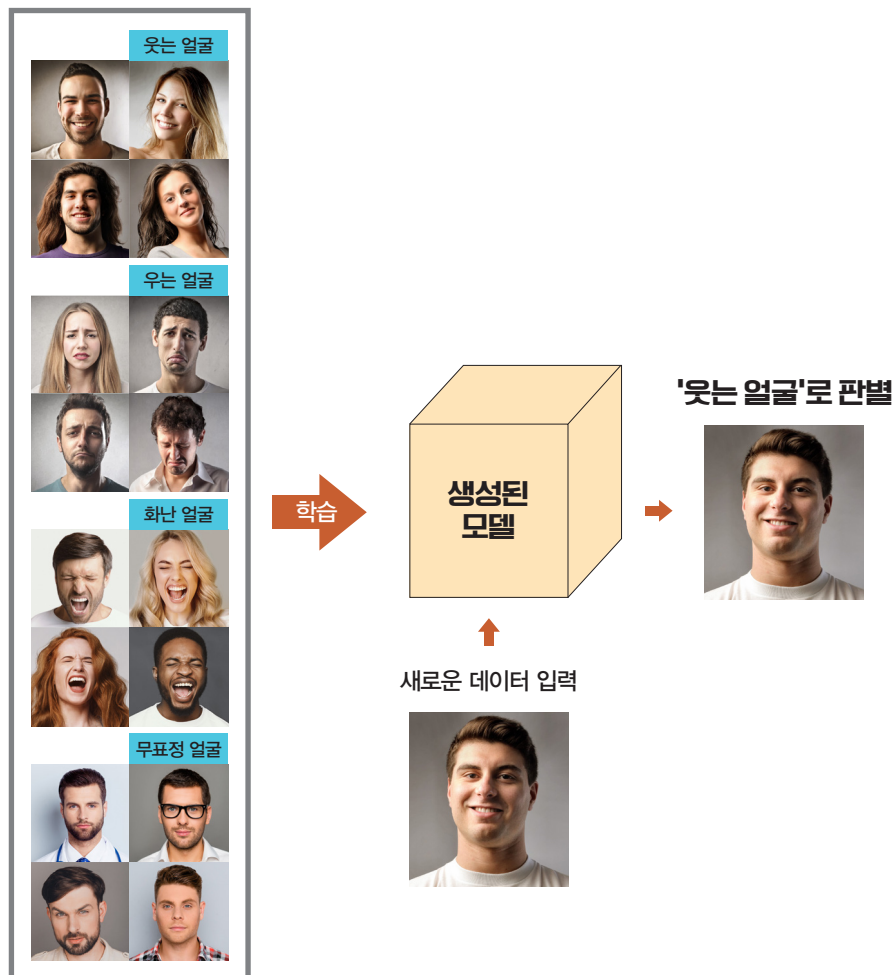
우리가 해결해야 할 문제가 무엇인지 적어 봅시다.

위 문제를 해결하기 위해 생각한 방법은 무엇인지 적어 봅시다.

앞의 시 이야기에서 체험한 인공지능 프로그램을 살펴보고, 우리가 배워야 할 인공지능 윤리 핵심 요건을 알아보시다.

1. 얼굴 인식 인공지능 프로그램

성별, 나이, 감정을 판별해 주는 인공지능 프로그램을 만들려면 많은 데이터가 필요합니다. 예를 들어, 사람의 얼굴에서 감정을 판별해 주는 프로그램을 만들고자 한다면 웃는 얼굴, 우는 얼굴, 화난 얼굴, 무표정한 얼굴 등의 이미지 데이터를 수집해야 합니다. 그리고 수집된 감정에 해당하는 이미지 데이터를 컴퓨터에 제공하면 컴퓨터가 스스로 학습해 감정을 판별해 주는 인공지능 모델을 만들어 줍니다. 만든 인공지능 모델에 웃는 얼굴, 우는 얼굴 등의 새로운 데이터가 입력되면 어느 감정에 해당하는지를 판단하여 알려 줍니다.



이때 학습에 사용되는 데이터를 수집하는 과정에서 편향이 발생할 수 있고, 수집된 데이터를 인공지능에게 제공(레이블링)하는 과정에서도 편향이 발생할 수 있습니다. 분류 모델의 성능을 높이기 위해서 이러한 편향을 줄이도록 노력해야 합니다. 또한 만든 인공지능 모델을 이용하여 개발자가 프로그램을 구현할 때 역시 편향이 발생할 수 있기 때문에 편향을 줄이기 위해 노력해야 하며, 이때 사용자의 다양성 존중과 데이터 관리에 대한 이해가 필요합니다.

2. 핵심 요건 알아보기

다양성 존중과 데이터 관리에 대해 알아보시다.

1 다양성 존중

인공지능 윤리의 ‘다양성 존중’은 인공지능 개발 및 활용 전 단계에서 사용자의 다양성과 대표성을 반영함을 의미합니다. 따라서 개발 단계에서는 성별, 나이, 장애, 지역, 인종, 종교, 국가 등 개인 특성에 따른 편향과 차별을 최소화하고, 사용 단계에서는 모든 사람에게 공정하게 적용되어야 합니다. 또한 사회적 약자 및 취약 계층의 인공지능 기술 및 서비스에 대한 접근성을 보장하고, 인공지능이 주는 혜택은 특정 집단이 아닌 모든 사람에게 골고루 분배되도록 노력할 필요가 있습니다.

인공지능 시스템의 개발에 사용되는 학습 데이터가 편향되는 경우 특정 집단에게만 유리할 가능성이 있기 때문에 데이터의 다양성을 확보할 수 있어야 합니다. 또한 인공지능 기술의 활용으로 인한 혜택이 특정 집단이나 계층에만 집중되지 않도록 인공지능 시스템에 대한 접근과 활용에 있어 사회적 약자를 비롯한 다양한 사회 구성원에 대한 배려도 필요합니다. 이를 위해 인공지능 시스템의 개발과 운영 단계에서 다양한 사회적, 경제적 배경을 가진 이해관계자의 직·간접적인 참여를 보장하는 등 인공지능에 의한 편향과 차별의 발생을 최소화해야 합니다.

편향은 학습 데이터를 수집할 때 인적 요인 또는 물리적인 요인으로 발생할 수 있습니다.

인적 요인으로 인한 편향

사람이 의식적 혹은 무의식적으로 특정 정보에 대해 편향되어 있어서 발생합니다. 인적 편향을 방지하기 위해서는 데이터의 명확한 수집 및 검수 기준을 수립하여 수집 단계에서 사람들에 따라 데이터 특성이 편향되지 않도록 방지하거나, 검수 단계에서 다양하고 충분한 수의 검수하는 사람을 확보함으로써 편향을 바로 잡을 수 있어야 합니다.

물리적인 요인으로 인한 편향

데이터를 수집하는 도구나 방법에 의해 발생할 수 있습니다. 이미지의 촬영 도구나 저장 장치와 같은 요인으로 인하여 이미지의 색상, 밝기, 해상도가 달라지는데, 이에 따라 물리적으로 한정된 데이터가 수집될 수 있습니다. 이로써 촬영 대상자의 연령대나 인종을 구분하기 힘들거나, 특정 방법으로 수집된 데이터만 학습될 수 있습니다. 이런 편향을 발생시킬 수 있는 물리적 요인을 제거하거나 다양한 수집 장치를 활용하여 다양성을 보완하는 것이 바람직합니다.

관련 사례

기사



알고리즘에 의해 산출된 성적에 대한 학생 반발

영국에서 코로나19로 대학 입학시험을 치르지 못하자 인공지능 알고리즘을 활용하여 학생의 예상 성적을 산출하였는데, 부유층 학생에게는 예상보다 좋은 학점을, 빈곤층 학생에게는 낮은 학점을 부여한 편향된 알고리즘 결과가 나타났습니다. 이에 학생들의 반발이 심해지자 철회되는 일이 있었습니다.

[출처: 경향신문 2020. 8]

2 데이터관리

인공지능 시스템 활용에서 중요한 가치로 언급되는 데이터 관리는 개인 정보 등 각각의 데이터를 그 목적에 부합하도록 활용하고 목적 외 용도로 활용하지 않고 데이터 수집과 활용의 전 과정에서 데이터 편향성이 최소화되도록 데이터 품질과 위험을 관리해야 함을 의미합니다.

인공지능 시스템의 개발에는 많은 데이터가 필요하고, 이러한 데이터를 수집하여 인공지능 시스템의 개선에 활용하는 구조이기 때문에 데이터의 품질을 확보하기 위한 관리가 매우 중요합니다. 또한 인공지능 시스템의 개발을 위해 수집된 데이터는 그 고유한 목적을 위해서만 활용해야 하기 때문에 목적 외의 용도로 활용하는 것을 금지하는 조치가 필요합니다. 인공지능 시스템 개발 및 운영사는 양질의 최적 데이터를 확보하고, 또한 이러한 데이터가 목적 범위 내에서만 활용되도록 내부적으로 통제하는 것이 중요합니다.

인공지능을 학습할 때 사용하는 데이터들은 개발하는 과정에서 데이터가 추가로 수집될 수 있으며, 수집된 데이터는 다른 인공지능 모델을 만들 때에도 학습 데이터로 사용될 수 있습니다. 이때, 데이터가 수집된 출처, 특징과 같은 데이터 정보가 미비하다면 재사용성이 떨어지거나 데이터 때문에 발생하는 문제에 대한 원인 파악이 어려울 수 있습니다. 따라서 수집된 데이터의 올바른 활용과 문제 발생 시 원인 추적을 위해 학습에 사용한 데이터의 명확한 출처, 구축 시점, 오픈소스 데이터 세트 버전과 같은 정보를 관리해야 합니다.

관련 사례

기사

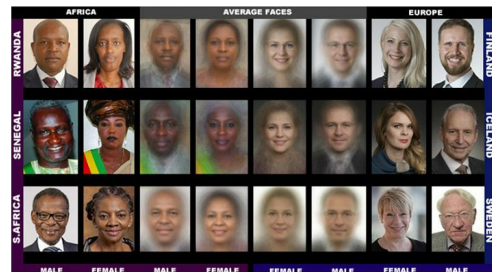
편견마저 학습하는 인공지능

인공지능(AI)이 현실의 편견마저 학습하고 있는 것으로 드러났습니다. 미국 매사추세츠공대(MIT) 미디어랩 연구원은 2018년 2월 23일 미국 뉴욕에서 열린 '기계학습연구학회'에서 얼굴 인식 AI가 피부색이 어두워질수록 오차율이 높다고 밝혔습니다.

연구팀은 해외 기업 3사의 AI 얼굴 인식 시스템을 이용해 사진 1,270장을 분석했습니다. 백인 남성은 3사 모두 오차율이 1% 미만으로 낮았지만, 백인 여성은 7%, 흑인 남성은 12%로 나타났으며 흑인 여성은 최대 35%였습니다. 3사 모두 백인보다 흑인이, 남성보다는 여성의 오차율이 더 높았습니다.

연구팀은 논문에서 "인공지능에 사용하는 데이터가 백인과 남성을 중심으로 구성돼 있다."라며 '현실 세계의 편견과 차별이 인공지능에까지 영향을 미치지 않도록 경계해야 할 것'이라고 강조했습니다.

[출처: 동아사이언스, 노컷뉴스 2018.02]




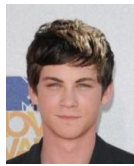




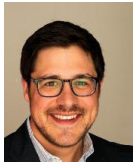

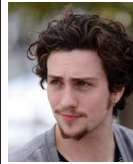





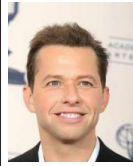
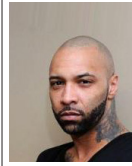
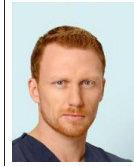




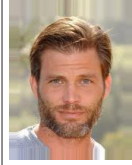






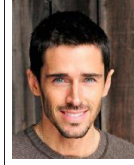





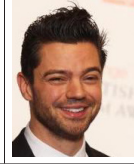










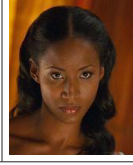


STEP 3 얼굴 이미지 속 정보로 분류하기



얼굴 이미지들을 다양하게 분류해 봅시다.

활동 1-1 얼굴 이미지 속 정보를 이용한 분류 활동

아래 다양한 얼굴 이미지들을 다양하게 분류하고자 합니다. 이미지에서 얻을 수 있는 정보를 이용하여 분류할 수 있는 방법을 생각해 봅시다.

| | A | B | C | D | E | F |
|---|--|--|--|--|--|--|
| 1 |  A1 |  B1 |  C1 |  D1 |  E1 |  F1 |
| 2 |  A2 |  B2 |  C2 |  D2 |  E2 |  F2 |
| 3 |  A3 |  B3 |  C3 |  D3 |  E3 |  F3 |
| 4 |  A4 |  B4 |  C4 |  D4 |  E4 |  F4 |
| 5 |  A5 |  B5 |  C5 |  D5 |  E5 |  F5 |
| 6 |  A6 |  B6 |  C6 |  D6 |  E6 |  F6 |
| 7 |  A7 |  B7 |  C7 |  D7 |  E7 |  F7 |
| 8 |  A8 |  B8 |  C8 |  D8 |  E8 |  F8 |

1 앞의 얼굴 이미지들을 분류할 수 있는 방법에 대해 생각해 보고 기준을 정해 봅시다.

예 머리카락 길이로 분류할 수 있다.

2 자신이 분류한 기준 한 가지를 적용하여 분류 결과를 작성하고 친구들과 비교해 봅시다. (단, 분류 결과는 파일 이름으로 작성합니다.)

분류 기준

분류 결과

3 성별을 분류하는 인공지능 프로그램을 만들려고 할 때 분류 기준을 정해 봅시다.

분류 기준

4 아래와 같은 얼굴 데이터를 3에서 정한 분류 기준을 적용하여 성별을 분류하고자 할 경우 결과가 어떻게 나올지 생각해 작성해 보고, 분류 결과를 친구들과 비교해 봅시다.

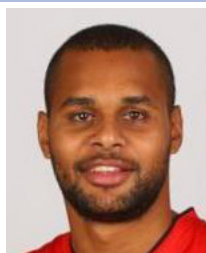
테스트 데이터1



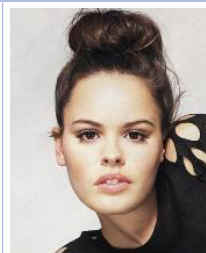
테스트 데이터2



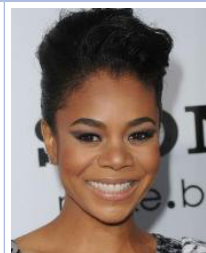
테스트 데이터3



테스트 데이터4



테스트 데이터5



• 나의 분류 결과를 적어 봅시다.

| 테스트 데이터1 | 테스트 데이터2 | 테스트 데이터3 | 테스트 데이터4 | 테스트 데이터5 |
|----------|----------|----------|----------|----------|
| | | | | |

• 친구의 분류 결과를 적어 봅시다.

| 테스트 데이터1 | 테스트 데이터2 | 테스트 데이터3 | 테스트 데이터4 | 테스트 데이터5 |
|----------|----------|----------|----------|----------|
| | | | | |

• 테스트 데이터의 분류 결과가 친구들의 것과 같은가요? 만약 성별을 잘못 분류한 테스트 데이터가 있다면 무엇이 문제일지 생각해 보고, 해결 방법을 적어 봅시다.

5 어떻게 문제를 해결하였는지 발표하여 친구들과 아이디어를 공유해 봅시다. 친구들의 해결 방법 중 좋은 아이디어라고 생각한 것과 그렇게 생각한 이유를 적어 봅시다.

데이터 관리의 측면에서 편향성을 줄이기 위한 방법에 대해 생각해 봅시다.

활동 1-2 성별을 구분할 수 있는 데이터를 추가하여 편향성 줄이기

활동 1-1의 4단계 테스트 데이터들로 성별을 분류했을 때 원하는 결과가 나올 수 있도록 필요한 데이터를 모듈별로 추가하고 다른 모듈과 비교해 봅시다.

1 모듈별로 성별을 좀 더 정확하게 판별하기 위해 필요한 얼굴 데이터를 직접 그리거나 인터넷에서 사진들을 찾아 추가해 봅시다.

| | | | |
|--|--|--|--|
| | | | |
| | | | |
| | | | |

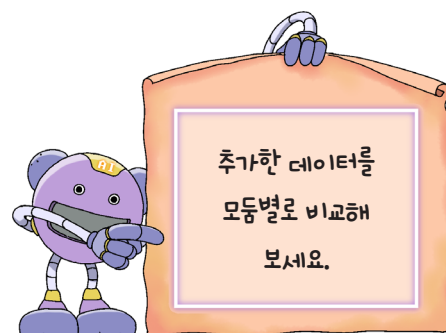
2 모듈별로 그리거나 인터넷에서 찾은 얼굴 데이터를 추가한 다음, 아래 테스트 데이터가 성별로 잘 분류되었는지 결과를 작성해 보고 분류 결과를 분석해 봅시다.

| 테스트 데이터1 | 테스트 데이터2 | 테스트 데이터3 | 테스트 데이터4 | 테스트 데이터5 |
|---|---|---|--|---|
|  |  |  |  |  |

- 테스트 데이터의 분류 결과를 적어 봅시다.

| 테스트 데이터1 | 테스트 데이터2 | 테스트 데이터3 | 테스트 데이터4 | 테스트 데이터5 |
|----------|----------|----------|----------|----------|
| | | | | |

- 테스트 데이터의 분류 결과가 친구들의 것과 같은가요? 만약 성별을 잘못 분류한 테스트 데이터가 있다면 무엇이 문제일지 생각해 보고 해결 방법을 적어 봅시다.





정리하기



다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 다양성 존중을 위해 성별, 나이, 인종, 종교, 국가 등 개인 특성에 따른 편향과 차별을 최소화해야 한다. ()
2. 인공지능 시스템 개발에 사용되는 학습 데이터를 특정 집단에게 유리하도록 편향된 데이터를 수집해야 한다. ()
3. 데이터 관리를 위해서는 개인 정보 등 각각의 데이터를 그 목적 외 용도로 활용할 수 있다. ()



점검하기

| | 예 | 아니요 |
|----------------------------|--------------------------|--------------------------|
| 데이터를 분류하기 위한 기준을 세울 수 있나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 데이터에서 편향을 발견할 수 있나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 데이터의 편향이 발생하는 이유를 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 편향을 줄이기 위한 데이터를 추가로 수집했나요? | <input type="checkbox"/> | <input type="checkbox"/> |

2

편향성 이해를 위한

옷을 추천하는 인공지능 프로그램 만들기



학습 목표

- 성별을 판별하기 위한 데이터를 수집하여 인공지능 모델을 만들 수 있다.
- 성별 분류 인공지능 모델을 활용하여 옷을 추천해 주는 프로그램을 제작할 수 있다.
- 인공지능 프로그램 개발 과정에서 개발자의 편견이 반영될 수 있다는 점을 이해하고 이를 고치기 위해 노력할 수 있다.

AI 이야기 거울아! 거울아!

얼굴과 체형을 스캔해서 옷을 추천해 주는 스마트 거울이 등장했습니다.

직접 옷을 입어 보지 않고도 거울 앞에 서면 카메라가 인식하여 어울리는 여러 가지 옷을 스타일링해 보고 어울리는 옷을 추천받을 수 있다고 합니다.

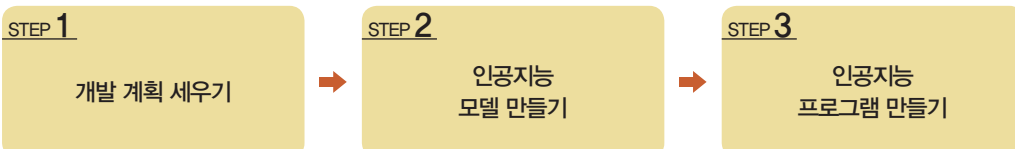
거울 앞에 서자 불과 몇초 지나지 않아 “인식 중입니다.”라는 안내 방송이 흘러나오고 드디어 순식간에 추천하는 옷을 입은 A 군의 모습이 거울 화면에 나타났습니다.

그런데 A 군의 표정이 좋아 보이지 않습니다. 도대체 무슨 일이 생긴 것일까요?



- 인공지능이 왜 A 군을 여자로 인식했을까요?
- 인공지능 거울이 있다면 이용할 것인가요? 이용한다면 주의할 점이 없을까요?

학습 과정



STEP 1 개발 계획 세우기



이번 활동에서는 편향성 이해를 위해 성별을 판별해 옷을 추천해 주는 프로그램을 개발해 보겠습니다. 이때 단계별로 윤리적 이슈가 무엇이 있는지 생각해 보면서 개발 계획을 세워 봅시다.

개발자

활동 2-1 성별을 판별해 옷을 추천해 주는 프로그램 개발 계획 세우기

성별을 판별해 옷을 추천해 주는 프로그램을 만들기 위한 개발 계획을 세워봅시다.

- 1 인공지능 모델을 만들기 위해 필요한 데이터와 이 데이터를 수집할 때 발생할 수 있는 윤리적 이슈를 생각해 봅시다.

| 필요한 데이터 | 윤리적 이슈 |
|--------------|--------------|
| | |

- 2 프로그램에 필요한 기능과 이 기능을 결정할 때 문제가 될 수 있는 윤리적 이슈를 생각해 봅시다.

| 프로그램에 필요한 기능 | 윤리적 이슈 |
|--------------|--------------|
| | |

- 3 프로그램의 알고리즘을 설계해 봅시다. 이 알고리즘을 설계할 때 문제가 될 수 있는 윤리적 이슈를 생각해 봅시다.

| 알고리즘 | 윤리적 이슈 |
|--------------|--------------|
| | |

- 4 프로그램의 활용 방법을 적어 봅시다. 이 프로그램을 활용할 때 문제가 될 수 있는 윤리적 이슈를 생각해 봅시다.

| 프로그램 활용 방법 | 윤리적 이슈 |
|--------------|--------------|
| | |

STEP 2 인공지능 모델 만들기



성별을 판별해 옷을 추천해 주는 프로그램에 필요한 인공지능 모델을 생성하여 학습시키고 다양한 얼굴 테스트 이미지를 이용하여 학습이 잘되었는지 확인해 봅시다.

개발자














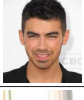
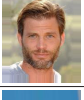





활동 2-2 '성별 판별' 인공지능 모델 만들기

성별에 따른 옷을 추천해 주는 프로그램을 만들기 위해서는 성별을 판별해 주는 인공지능 모델이 필요합니다. 필요한 데이터를 수집하여 인공지능 모델을 만든 다음, 성능을 확인해 봅시다.

인공지능 모델 생성 및 학습

- ① 분류할 모델의 이름을 설정합니다.
- ② 클래스명을 남자, 여자로 설정합니다.
- ③ 남, 여 얼굴 이미지를 5개 이상 수집하여 클래스에 업로드합니다(캠으로 촬영 가능).
- ④ 인공지능 모델을 학습합니다.
- ⑤ 인공지능 모델의 학습이 잘되었는지 테스트합니다.

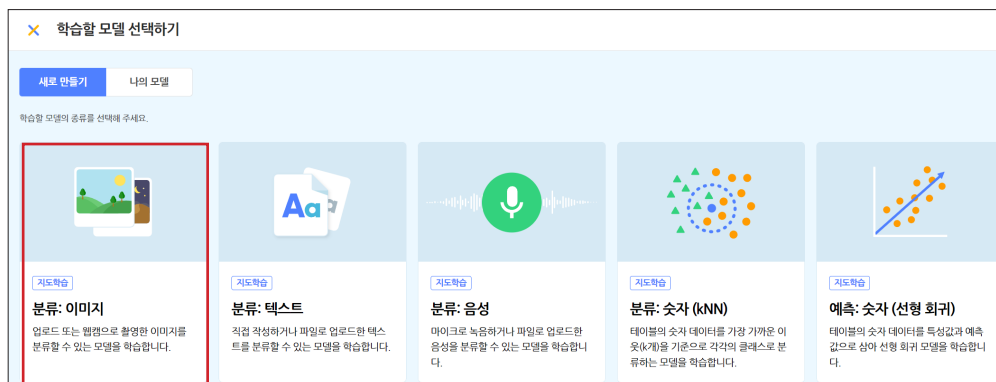
1 성별을 판별하기 위해 필요한 데이터를 수집합니다.(1차시에서 사용된 데이터를 이용합니다.)

| 남성 이미지 데이터 | 여성 이미지 데이터 |
|---|--|
|       |       |
|       |       |
|       | |
|       | |
|       | |
|       | |

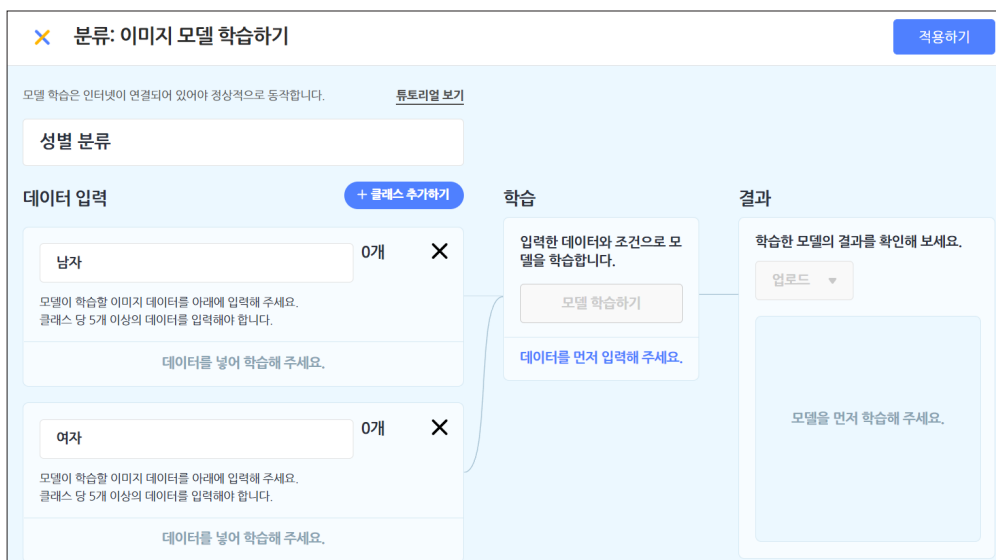
- 2 엔트리(<https://playentry.org/>) 사이트에 접속하여 로그인한 후 [만들기] – [작품 만들기]를 클릭합니다.



- 3 블록 카테고리에서 [인공지능] – [인공지능 모델 학습하기] – [지도학습 분류: 이미지]를 차례대로 클릭합니다.



- 4 모델 이름에 '성별 분류'라 입력하고, 클래스1을 '남자', 클래스2를 '여자'라 입력합니다. 이 클래스 이름이 정답에 해당하는 레이블입니다.



- 7 학습이 완료된 모델의 성능을 평가하기 위해 다양한 테스트 데이터에 대한 모델의 결과를 확인하고 분류 결과를 적어 봅시다.



- 분류 결과를 확인해 보세요.

| 테스트 데이터1 | 테스트 데이터2 | 테스트 데이터3 | 테스트 데이터4 | 테스트 데이터5 |
|----------|----------|----------|----------|----------|
| | | | | |

- 8 테스트 데이터의 결과를 확인해 보고 성별이 잘못 나왔다면 원하는 결과가 나오기 위해 적절한 학습 데이터를 추가하여 모델을 다시 만들고 결과를 확인해 봅시다.

- 9 모델이 완성되었다면 오른쪽 상단의 [적용하기](#) 버튼을 클릭하여 모델 생성을 완료합니다.

STEP 3 인공지능 프로그램 만들기



성별을 판별해 주는 인공지능 모델을 만들었다면 만든 모델을 활용하여 옷을 추천해 주는 프로그램을 만들어 봅시다.

개발자

활동 2-3 프로그래밍하기

알고리즘을 토대로 성별을 판별해 옷을 추천해 주는 프로그램을 구현해 봅시다.

알고리즘 확인하기

- ① 얼굴을 카메라에 비치면 얼굴을 보고 성별을 판별합니다.
- ② 성별이 남자이면 바지를 추천합니다.
- ③ 성별이 여자이면 치마를 추천합니다.

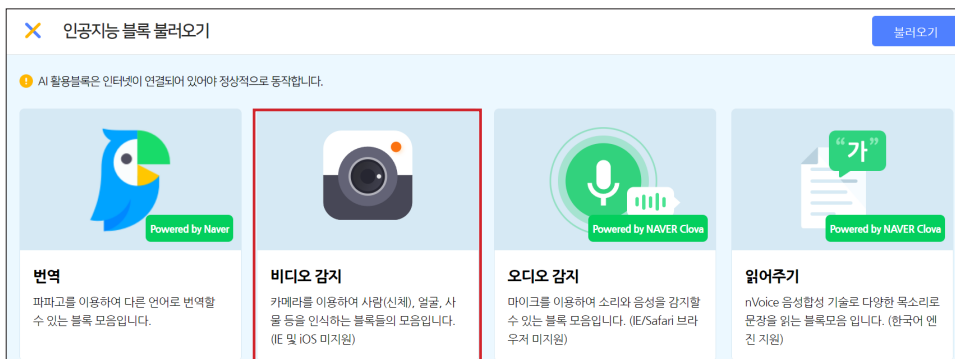
1 알고리즘 순서대로 프로그램이 동작되도록 [신호]를 만듭니다.

- ① 블록 카테고리에서 [속성] - [신호] - [신호 추가하기]를 차례대로 클릭합니다.
- ② '성별 판별하기', '남자 옷 추천하기', '여자 옷 추천하기' 신호를 만듭니다.

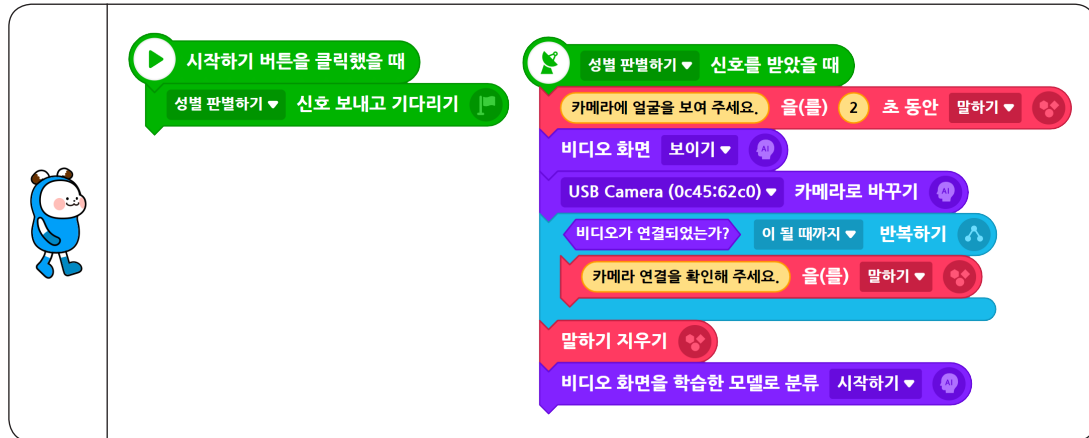


2 얼굴을 카메라에 비치면 얼굴을 보고 성별을 판별합니다.

- ① 카메라를 사용하기 위해 [인공지능] 블록 카테고리에서 [인공지능 블록 불러오기] - [비디오 감지] - [불러오기]를 차례대로 클릭합니다.



- ② “카메라에 얼굴을 보여 주세요.” 메시지와 함께 카메라에 비친 얼굴로 성별을 인식하는 프로그램을 작성합니다.



Scratch 스크립트 예시:

- 시작하기 버튼을 클릭했을 때
 - 성별 판별하기 ▾ 신호 보내고 기다리기
- 성별 판별하기 ▾ 신호를 받았을 때
 - 카메라에 얼굴을 보여 주세요. 을(를) 2 초 동안 말하기 ▾
 - 비디오 화면 보이기 ▾
 - USB Camera (0c45:62c0) ▾ 카메라로 바꾸기
 - 비디오가 연결되었는가? 이 될 때까지 ▾ 반복하기
 - 카메라 연결을 확인해 주세요. 을(를) 말하기 ▾
 - 말하기 지우기
 - 비디오 화면을 학습한 모델로 분류 시작하기 ▾

3 성별이 남자이면 바지를 추천합니다.

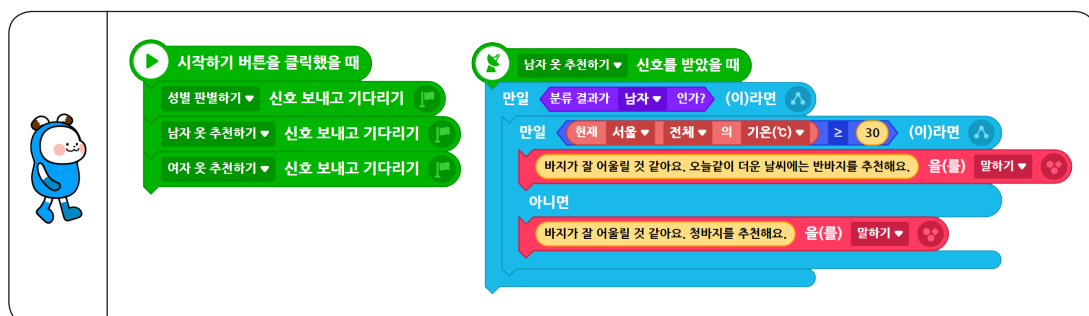
- ① 날씨 정보를 이용하기 위해 [확장] 꾸러미에서 [확장 블록 불러오기] - [날씨]를 선택하고 불러오기 버튼을 클릭합니다.



확장 블록 불러오기 화면:

- 날씨: 기온, 강수량, 미세먼지 농도 등 한국의 날씨와 관련된 블록 모음입니다. [웨더아이 제공]
- 생활안전 국민행동요령: 생활 속 안전을 위해 국민이 지켜야 하는 행동요령에 대한 블록 모음입니다. [국민안전처 제공]
- 자연재난 국민행동요령: 자연재난 발생 시 국민이 지켜야 하는 기본적인 행동요령에 대한 블록들의 모음입니다. [국민안전처 제공]
- 행사: 국내 지역별 다양한 행사 및 축제와 관련된 블록 모음입니다. [한국관광공사 제공]

- ② 분류 결과가 ‘남자’이면 ‘남자’로 판별됨을 알리고, 날씨 정보를 이용하여 현재 기온이 30도 이상이면 반바지, 30도 미만이면 긴바지를 추천하는 프로그램을 추가합니다.

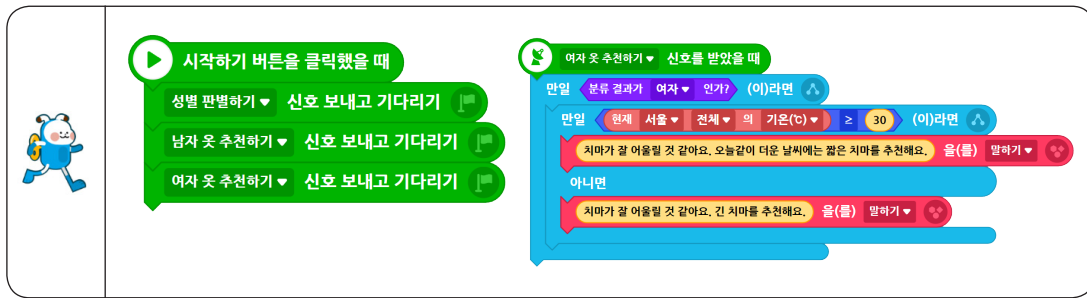


Scratch 스크립트 예시:

- 시작하기 버튼을 클릭했을 때
 - 성별 판별하기 ▾ 신호 보내고 기다리기
 - 남자 옷 추천하기 ▾ 신호 보내고 기다리기
 - 여자 옷 추천하기 ▾ 신호 보내고 기다리기
- 남자 옷 추천하기 ▾ 신호를 받았을 때
 - 만일 분류 결과가 남자 ▾ 인가? (이)라면
 - 만일 현재 서울 ▾ 전체 ▾ 의 기온(°C) ▾ ≥ 30 (이)라면
 - 바지가 갈 어울릴 것 같아요. 오늘같이 더운 날씨에는 반바지를 추천해요. 을(를) 말하기 ▾
 - 아니면
 - 바지가 갈 어울릴 것 같아요. 청바지를 추천해요. 을(를) 말하기 ▾

4 성별이 여자이면 치마를 추천합니다.

분류 결과가 ‘여자’이면 ‘여자’로 판별됨을 알리고, 날씨 정보를 이용하여 현재 기온이 30도 이상이면 짧은 치마, 30도 미만이면 긴 치마를 추천하는 프로그램을 추가합니다.



◆ 위 프로그래밍 단계 중 어느 단계에서 개발자의 편견이 포함되어 있는지 찾아보고, 왜 문제가 되는지 이야기해 봅시다.



정리하기



● 인공지능 모델을 이용하여 문제를 해결하는 과정을 알맞은 것끼리 선을 연결하여 정리해 봅시다.



1. 문제 이해



성별을 판별해 주는 인공지능 모델을 활용하여 옷을 추천해 주는 프로그램 구현하기

2. 개발 계획 세우기



옷을 추천하는 프로그램을 만드는 과정에서 나타날 수 있는 편향성을 생각해 보고 해결 방안 적어 보기

3. 인공지능 모델 만들기



성별을 판별하기 위해 필요한 데이터를 수집하고, 수집된 데이터를 이용하여 인공지능 모델을 만들고 성능 확인하기

4. 인공지능 프로그램 만들기



성별을 판별해 옷을 추천하는 인공지능 프로그램을 만들기 위한 문제 상황 이해하기

5. 편향 생각해 보기



성별을 판별해 주는 인공지능 모델을 만들고, 옷을 추천하는 프로그램을 만들기 위해 필요한 개발 계획 작성 및 윤리적 이슈 생각하기



점검하기

| | 예 | 아니요 |
|---|--------------------------|--------------------------|
| 옷을 추천하는 인공지능 프로그램을 만들기 위한 계획을 구체적으로 세웠나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 성별을 판별하는 인공지능 모델에 필요한 데이터를 수집했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 수집된 데이터를 이용하여 성별을 분류하는 모델을 만들었고, 만들어진 모델의 성능이 좋았나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 개발하는 과정에서 나타날 수 있는 편향에 대해 이해하고 이를 개선하기 위한 방안을 생각해 보았나요? | <input type="checkbox"/> | <input type="checkbox"/> |

3

개발자와 사용자는 어떤 생각을 할까?



학습
목표

- 인공지능을 개발하고 활용하는 과정에서 편향성이 발생하는 문제를 파악하고 이에 대한 대응 방안을 제시할 수 있다.
- 개발자와 사용자 입장에서 데이터 관리의 중요성을 파악하고 문제가 발생했을 때 대응 방안을 제시할 수 있다.

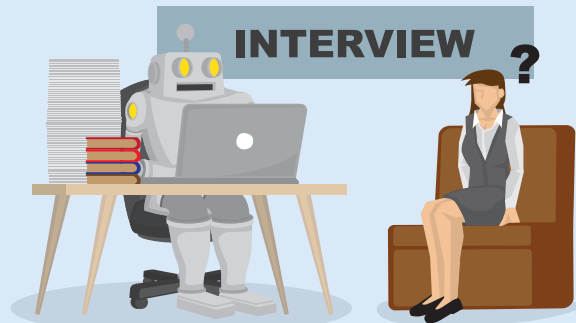
AI 이야기 채용에 인공지능이 활용된다고?

치킨 회사 B 사는 지난 2021년부터 AI 역량 검사를 도입하여 첨단 인공지능을 활용한 면접으로 보다 공정하면서도 정밀한 검증을 거쳐 미래 외식 산업을 이끌어 갈 인재를 채용했습니다.

AI 면접관은 지원자의 답변뿐 아니라 표정이나 몸짓, 목소리 등의 비언어적 표현을 분석해 전체적으로 평가합니다. 돌발 질문을 던진 후 지원자의 답변 습관과 시선 처리까지 짚어 냅니다.

B 사 관계자에 따르면 “지원자의 학력, 스펙 등에 대한 선입견 없이 보다 공정한 기준으로 평가할 수 있는 AI 면접이 최근 주목받고 있다.”라며 “AI 면접 솔루션은 단순히 채용 과정의 속도를 높이는 것을 넘어 기업들이 원하는 인재를 확보할 확률을 높이는 데 최적일 것”이라고 말했습니다.

최근 기업들이 인공지능을 인재 채용에 활용하고 있으면서 면접관 개인의 성향이나 편견이 없다는 점에서 긍정적인 평가를 받고 있습니다. 반면, 점수를 준 이유를 알기 어렵고, 특정 인종·성별·나이대의 지원자에게 더 높은 점수를 주는 편향 문제가 종종 나타나 우려의 목소리가 나오기도 했습니다.



[출처: <http://www.kbsm.net/news/view.php?idx=359728>]



- 편향되지 않은 채용 프로그램을 만드는 개발자라면 어떻게 해야 할까요?
- 인공지능 면접관이 이유를 알려 주지 않고 나를 탈락시켰다면 기분이 어떨까요?

학습 과정

STEP 1

개발자 관점에서
한 걸음 더 나아가기



STEP 2

사용자 관점에서
한 걸음 더 나아가기



STEP 3

개발자, 사용자 관점을
비교하여 의견 정리하기

STEP 1 개발자 관점에서 한 걸음 더 나아가기



개발자의 관점에서 채용 프로그램을 어떻게 개발할지 생각해 봅시다.

개발자

활동 3-1 채용 프로그램 개발자의 입장 되어 보기

아래 상황을 살펴보고, 채용 프로그램을 어떻게 만들지 개발자 입장에서 생각해 봅시다.

[채용 조건]

- 남자 3명, 여자 3명 선발
- 긍정(밝음)적인 성격의 소유자
- 외국어 가능자
- 범죄자가 아님.

[지원자 현황]

| 순번 | 성명 | 나이 | 성별 | 최종 학력 | 가족 수 | 외국어 | 출생 지역 | 성격 | 범죄 유무 | 자격증 수 |
|----|-----|----|----|-------|------|-----|-------|----|-------|-------|
| 1 | 김○기 | 25 | 남 | 학사 | 4 | 하 | 경기 | 밝음 | 무 | 1 |
| 2 | 서○원 | 41 | 남 | 석사 | 4 | 상 | 경기 | 밝음 | 무 | 0 |
| 3 | 최○원 | 35 | 여 | 박사 | 3 | 상 | 충청 | 밝음 | 무 | 4 |
| 4 | 권○찬 | 34 | 남 | 석사 | 3 | 중 | 서울 | 조용 | 무 | 3 |
| 5 | 홍○연 | 32 | 여 | 석사 | 1 | 중 | 경기 | 조용 | 무 | 2 |
| 6 | 양○연 | 30 | 여 | 학사 | 4 | 하 | 강원 | 조용 | 무 | 0 |
| 7 | 홍○동 | 44 | 남 | 고졸 | 2 | 하 | 충청 | 밝음 | 유 | 1 |
| 8 | 권○민 | 32 | 여 | 박사 | 3 | 상 | 부산 | 밝음 | 무 | 3 |
| 9 | 이○경 | 25 | 여 | 박사 | 1 | 상 | 강원 | 밝음 | 무 | 2 |
| 10 | 정○민 | 26 | 여 | 석사 | 2 | 상 | 대구 | 밝음 | 무 | 4 |
| 11 | 황○성 | 26 | 남 | 석사 | 3 | 중 | 대전 | 밝음 | 무 | 3 |
| 12 | 김○성 | 24 | 남 | 고졸 | 4 | 중 | 전라 | 조용 | 무 | 0 |
| 13 | 김○호 | 33 | 여 | 고졸 | 4 | 하 | 경상 | 조용 | 무 | 0 |
| 14 | 안○창 | 29 | 남 | 석사 | 3 | 상 | 경상 | 밝음 | 유 | 5 |
| 15 | 서○민 | 28 | 남 | 학사 | 5 | 상 | 전라 | 밝음 | 무 | 0 |

1 위와 같은 채용 조건으로 선발한다면 선발된 사람의 순번을 적어 봅시다.

2 개발자 입장에서 채용 조건을 바꿔 보고, 바꾼 이유와 선발된 인원을 적어 봅시다.

- 개발자 입장에서 채용 조건을 바꿔 봅시다.

- 채용 조건을 바꾼 이유를 적어 봅시다.

- 최종 선발된 사람의 순번을 적어 봅시다.

3 개발자 입장에서 편향되지 않도록 채용 프로그램을 만들 때 느낀 점을 작성해 보고 친구들과 공유해 봅시다.

STEP 2 사용자 관점에서 한 걸음 더 나아가기



사용자의 관점에서 인공지능 프로그램에 대해 생각해 봅시다.

사용자

활동 3-2 채용 프로그램 사용자의 입장 되어 보기

개발자가 설계한 프로그램을 보고 다음에 대해 생각해 봅시다.

- 1 인공지능 채용 프로그램에 합격하고자 할 경우 어떤 부분을 준비해야 하는지 적어 봅시다.

- 2 인공지능 채용 프로그램에 해당하는 조건을 모두 만족했는데 결과가 불합격이었고 불합격 사유도 알려주지 않는다면 사용자 입장에서는 어떻게 해야 할까요?

- 3 만약 인공지능 채용 프로그램에서 채용 조건을 알 수 없는 경우 사용자 입장에서는 어떤 준비를 해야 할까요?

- 4 사용자 입장에서 개발자에게 인공지능 프로그램에 요구하고 싶은 기능은 무엇인가요?

개발자와 사용자의 관점의 차이를 비교해 보고 의견 차이를 줄이는 방안을 생각해 봅시다.

개발자

사용자

활동 3-3 의견 정리하기

개발자와 사용자의 관점을 정리해 보고, 의견의 공통점과 차이점을 확인해 봅시다.

1 어떤 부분에서 의견의 공통점과 차이점이 있나요?

| | 개발자 관점 | 사용자 관점 |
|-----|--------|--------|
| 공통점 | | |
| 차이점 | | |

2 왜 이런 의견의 차이가 발생했을까요? 자신의 생각을 적어 봅시다.

3 의견 차이를 줄이기 위해 어떤 기준이 필요할까요? 윤리적 관점에서 기준을 만들어 봅시다.



정리하기



다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 개발자 입장에서 채용 프로그램을 만든다면 채용 조건에 개발자의 편견을 적용해도 무방하다. ()
2. 사용자 입장에서 인공지능이 판단한 결과를 전적으로 따라야 한다. ()



점검하기

| | 예 | 아니요 |
|---|--------------------------|--------------------------|
| 인공지능의 개발과 관련하여 개발자의 관점을 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능을 사용함에 있어서 사용자의 관점을 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능 프로그램의 개발자와 사용자의 관점 차이를 비교·분석하고 차이를 줄이기 위해 어떤 노력을 해야 하는지 설명했나요? | <input type="checkbox"/> | <input type="checkbox"/> |

4

다양성 존중과 데이터 관리 실천을 위한 활동



학습
목표

- 데이터를 라벨링하여 인공지능 모델을 만들어 편향성을 줄이는 문제를 해결할 수 있다.
- 편향적일 수 있는 인공지능을 활용한 채용 도입에 대한 자신의 생각을 표현할 수 있다.

AI 이야기 인공지능이 추천해 준 옷을 입고 동창회에 갔더니?

주말에 있을 동창회에가기 위해 A 양은 큰맘 먹고 옷을 한 벌 사기로 했습니다. 평상시 옷을 잘 못 고른다고 생각한 A 양은 인공지능 거울이 옷을 추천해 준다고 했던 것이 기억나 인공지능의 추천을 받아 C 사의 옷을 구입했습니다.

드디어 찾아온 동창회의 날, 동창회에 참석한 A 양은 동창회 장소에 들어가자마자 자신과 같은 C 사의 옷을 입고 온 여러 명의 친구를 보게 되었습니다.



- 위와 같은 일이 나에게 발생했다면 어떨까요? 인공지능 추천을 계속 받고 싶은가요?

학습 과정



STEP 1 체험을 통한 편향성 줄이기 실습하기

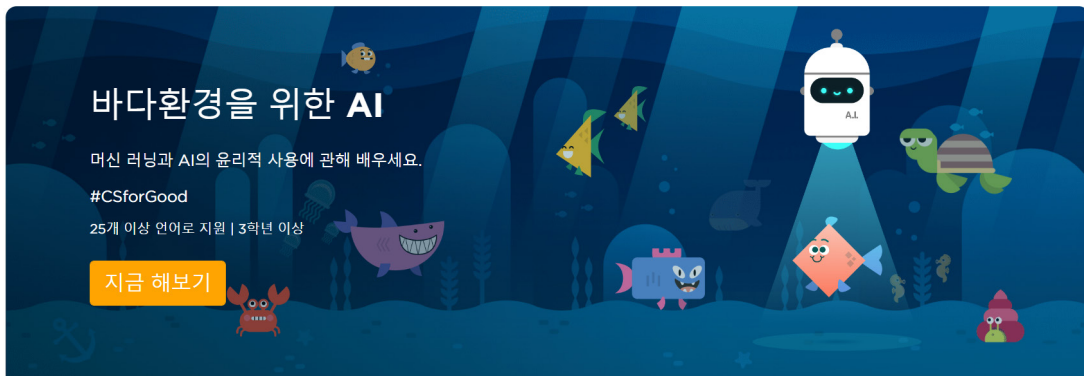


바다 환경을 위한 AI 모델 체험 실습을 통하여 수집된 데이터를 라벨링하는 작업자에 의한 편향을 알아봅시다.

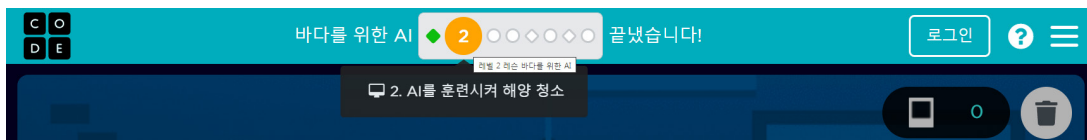
활동 4-1 바다 환경을 위한 AI 인공지능 모델 체험

바다 환경을 위한 AI 모델 체험을 통해 모델 성능을 평가해 봅시다. 활동을 마친 후 학습에 사용하는 데이터의 중요성과 데이터를 라벨링하는 작업자의 편향을 줄여야 하는 이유를 알아봅시다.

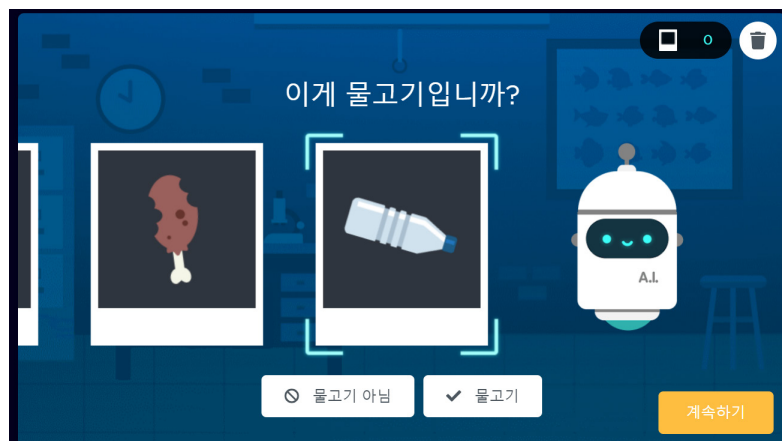
- 1 바다 환경을 위한 AI 모델 체험 사이트(<https://code.org/oceans>)에 접속하고, [지금 해보기]를 클릭합니다.



- 2 상단 메뉴에서 레벨2 레슨 바다를 위한 AI를 클릭합니다.



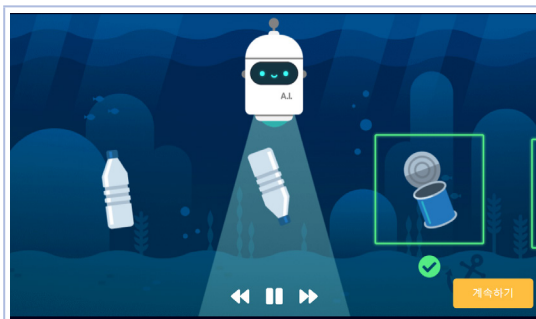
- 3 물고기와 쓰레기에 데이터 라벨링하는 작업을 진행합니다. 물고기인지 물고기가 아닌지를 선택하면 라벨링이 되며, 충분히 라벨링 작업을 한 후 [계속하기] 버튼을 클릭합니다.



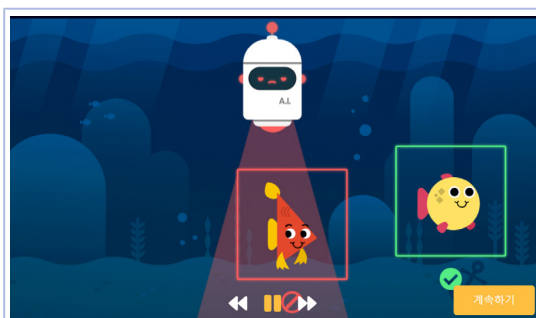
4 [실행] 버튼을 눌러 학습된 결과를 확인해 봅시다.



- 만약 물고기를 분류하지 못했다면 원인을 찾아보고, 그 이유를 적어 봅시다.



- 물고기를 분류하지 못했거나 물고기가 아닌 것을 물고기로 분류했다면, 그 이유는 무엇인지 적어 봅시다.



짝 또는 모둠별로 인공지능 윤리에 관한 주제를 토론해 봅시다.



토론

편향적일 수 있는 인공지능 채용 프로그램을 도입해야 하는가?

편향적일 수 있지만 인공지능 채용 프로그램을 도입해야 한다는 찬성 입장과 도입하지 말아야 한다는 반대 입장에서 자신의 주장을 뒷받침하는 근거 자료를 수집하고 서로의 입장을 짝과 토론해 보고 결론을 내려 봅시다.

채용 비리가 계속 발생하는 문제를 해결하기 위해 최근 인공지능을 활용한 채용이 늘어나고 있습니다. 인공지능을 활용한 채용은 면접관의 주관에 개입되지 않아 객관적이고 공정한 평가가 가능하다는 장점이 있습니다. 하지만, 인공지능을 활용한 채용을 위해 취업 준비생들은 또 다른 시간과 비용을 들여야 하며, 인공지능을 활용한 채용 역시 개발자나 데이터에 의한 편향성이 발생할 수 있습니다. 그럼에도 불구하고 인공지능을 활용하여 채용을 해야 하는지에 대하여 토론해 봅시다.

토론 방법

- ① 편향적일 수 있는 인공지능 채용 도입에 대한 찬성 입장과 반대 입장에 대한 자신의 주장을 적습니다.
- ② 인터넷에서 자신의 주장을 뒷받침하는 근거 자료를 수집합니다.
- ③ 한 명은 찬성, 다른 한 명은 반대 입장이 되어 토론합니다.
- ④ 찬성과 반대의 역할을 바꾸어 토론합니다.
- ⑤ 서로의 입장을 최종 정리하여 결론을 내려봅니다.

1 찬성과 반대의 입장을 주장과 근거로 제시해 봅시다.

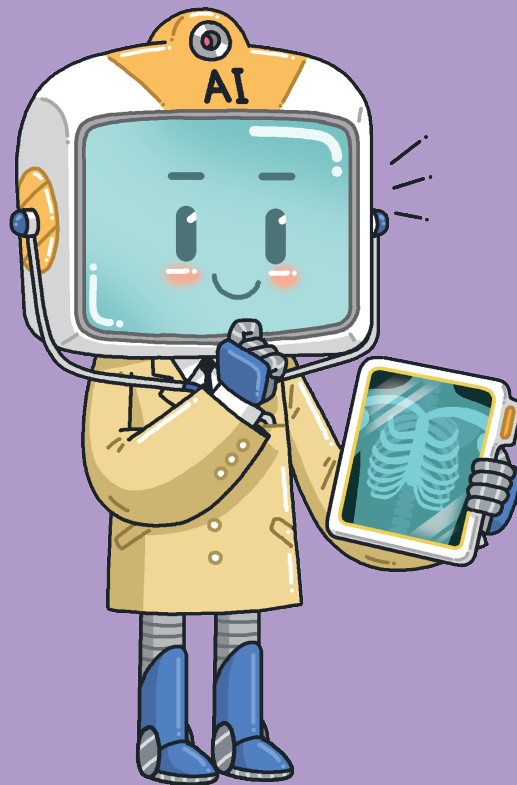
| 찬성 | 반대 |
|-------------|-------------|
| 주장1: 근거: | 주장1: 근거: |
| 주장2: 근거: | 주장2: 근거: |
| 주장3: 근거: | 주장3: 근거: |

2 1에서 정리한 내용을 바탕으로 나의 결론을 적어 봅시다.

4

인공지능 판단 믿어도 될까?

- 1 인공지능 결과 믿어도 될까?
- 2 투명성 이해를 위한 범죄 형량 판단 인공지능 프로그램 만들기
- 3 개발자와 사용자는 어떤 생각을 할까?
- 4 책임성, 투명성, 안전성 실천을 위한 활동



인권 보장

프라이버시 보호

다양성 존중

침해 금지

공공성

연대성

데이터 관리

책임성

안전성

투명성



학습 전 이야기 나누기

“이 인공지능 프로그램은 무엇일까요?”

우리나라의 한 병원에 도입되어 85명의 암 환자를 진료했습니다. 그중 대부분의 암 환자들은 이 인공지능의 처방을 신뢰했습니다. 심지어 암 질환 권위를 가진 전문가와 의견이 엇갈리는 경우에서도 이 인공지능의 처방을 선택했습니다. 일본에서는 어느 한 대학교 병원이 오진한 희귀 백혈병 환자의 정확한 진단명을 이 인공지능은 10분 만에 찾아내어 치료 방법까지 제시했다고 합니다. 전문가보다도 더 신뢰하는 이 인공지능 프로그램은 무엇일까요?

정답 인공지능 왓슨

1

인공지능 결과 믿어도 될까?



학습
목표

- 인공지능 윤리에서 강조되는 책임성과 투명성, 안전성의 개념을 이해하고 설명할 수 있다.
- 의사 결정 트리 활동을 통해 인공지능 시스템의 투명성이 필요함을 이해할 수 있다.
- 자율 주행 자동차 사고 사례를 통해 인공지능 윤리 책임성과 안전성을 이해할 수 있다.

AI 이야기 자율 주행 자동차 사고 나면 사고 과실 비율은 몇 대 몇?

T 사 CEO는 가까운 미래에 컴퓨터가 통제하는 무인 자동차가 일반화되는 시대가 도래하면 사람이 몰 수 있는 자동차는 모두 불법이 될 수 있다고 주장합니다. 무인 자동차의 안전성에 대해 걱정할 필요가 없다고 잘라 말한 그였지만, T 사 자동차의 운전자가 자율 주행 기능 사용 도중 사망하는 사고가 발생하면서 신뢰에 치명타를 입게 되었습니다. 자율 주행 자동차가 가져올 편리함과 안전성만큼 관심이 쏠리는 게 자율 주행 자동차의 사고 문제로, 책임에 대한 논란이 끊이지 않고 있습니다.

문제는 탑승자가 운전に関여하지 않더라도 해당 자율 주행 자동차의 소유자이고, 제조사는 자율 주행 시스템을 만들었지만, 운행에 관여했다고 보기 어렵다는 데 있습니다. 또한 자율 주행 자동차는 인공지능, 레이더, 라이다 센서, V2X 통신 등 다양한 기술을 통해 주변 도로와 상황 정보를 파악해 판단하기 때문에 시스템의 오작동이나 과실 여부를 객관적으로 증명하기 어렵습니다.

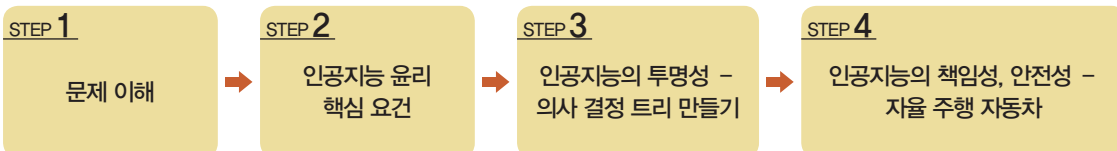
이처럼 미래 사회에서는 인공지능 기술을 활용한 자율 주행 자동차가 보편화될 것으로 예상됩니다. 이러한 인공지능 기술을 믿고 사용할 수 있을까요?



- 인공지능 기술을 활용한 자율 주행 자동차가 사고가 나면 누구에게 책임을 물어야 할까요?
- 자율 주행 자동차의 사고 원인을 파악하기 위해 인공지능 기술을 얼마나 공개해야 할까요?
- 자동차를 구매한다면 일반 자동차와 자율 주행 자동차 중 어떤 것을 구매할 것인가요?



학습 과정



이번 시간에 우리가 해결해야 할 문제와 문제 해결 과정을 알아보시다.



생활 속 문제 발견

남편을 잃은 누이동생과 조카 일곱을 부양하고 사는 일용직 노동자 장발장은 부양가족을 위해 빵을 훔치다가 잡혀 재판을 받게 됩니다. 공정한 재판을 위해 이 사건을 인공지능 판사에게 맡겼는데, 인공지능 판사는 장발장에게 징역 5년을 선고했습니다. 내가 만약 장발장이라면 인공지능 판사의 판결을 믿을 수 있을까요? 내가 만약 인공지능 판사라면 어떻게 형벌 기준을 세우고 어떤 형벌을 내릴까요?



우리가 해결해야 할 문제가 무엇인지 적어 봅시다.

위 문제를 해결하기 위해 생각한 방법은 무엇인지 적어 봅시다.

인공지능 도입 과정에서 문제점이 드러나게 된 배경과 함께 우리가 배워야 할 인공지능 윤리 핵심 요건을 알아봅시다.

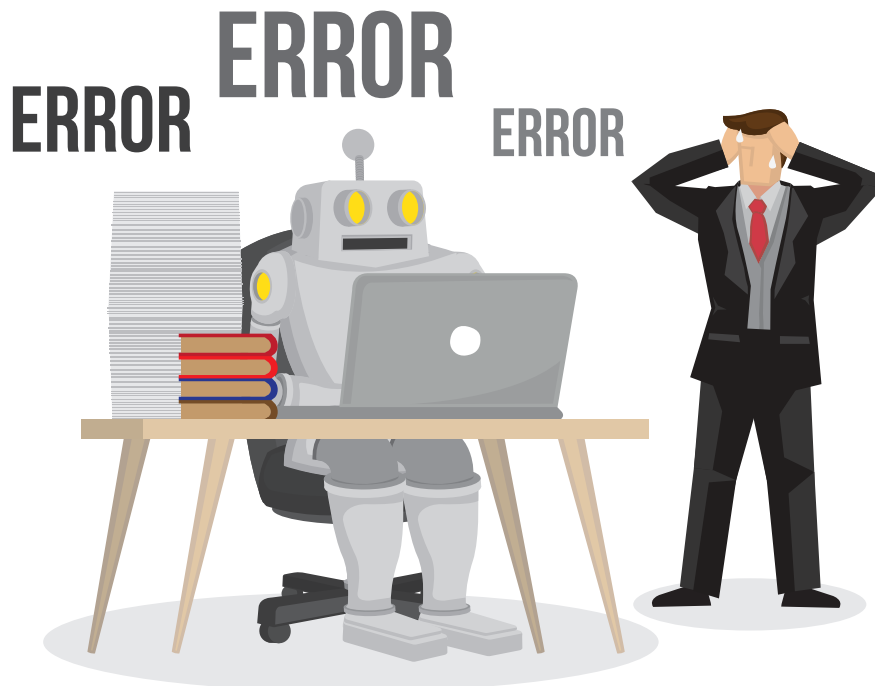
1. 배경

최근 4차 산업혁명이 경제 성장을 위한 주요 키워드로 부상함에 따라 제조업 이외의 다양한 분야로 확대되고 있습니다. 아울러 그에 따른 사회적 파급 효과에 관한 관심이 집중되고 있습니다.

4차 산업혁명의 핵심은 인공지능의 도입을 통한 산업의 자동화라고 할 수 있습니다. 하지만 인공지능이 도입되는 과정에서 인공지능 알고리즘 오류와 제품 및 서비스 사건·사고가 끊임없이 발생하면서 인공지능의 문제점에 관한 관심도 증가했습니다.

인공지능 개발 및 활용 과정에서 책임 주체를 명확하게 설정하는 것은 인공지능 활용 시 발생할 수 있는 문제나 사고의 책임성 논란을 최소화하는 데 도움을 줄 수 있습니다. 또한 잠재적 위험을 방지하고 안전성을 보장할 수 있도록 신뢰할 수 있는 인공지능을 활용하기 위해서는 투명성을 갖추도록 노력을 기울여야 합니다.

지금부터 책임성, 투명성, 안전성에 대해 알아봅시다.



2. 핵심 요건 알아보기

책임성, 투명성, 안전성에 대해 알아보시다.

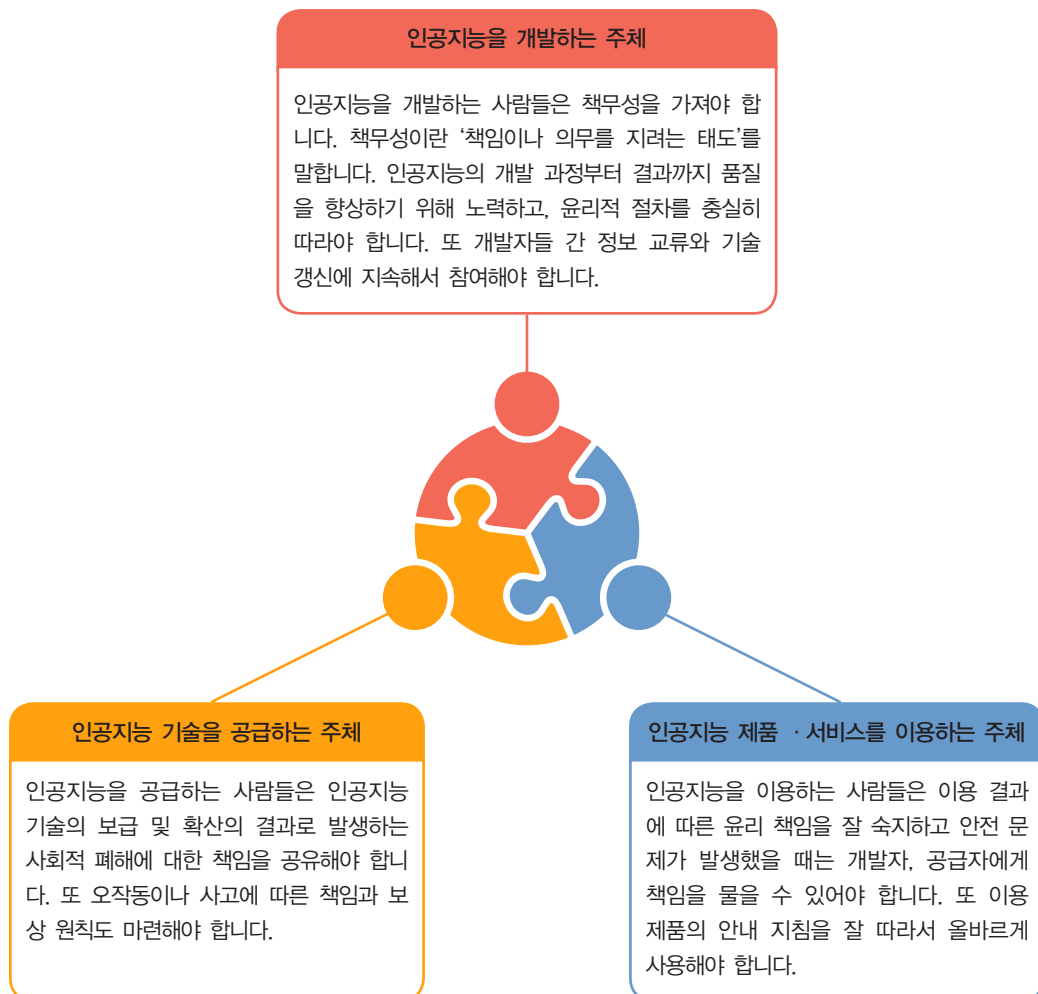
1 책임성

책임성은 인공지능 기술 관련 사고가 발생했을 때 책임의 주체를 명확히 하고, 안전과 관련한 통제 장치를 마련하는 등 책임과 관련된 의무를 충실히 해야 함을 의미합니다. 책임성은 인공지능을 활용하는 과정에서 나타난 행동에 대해 누가 어떻게 책임을 질 것인가에 관한 것이기 때문에 책임의 주체를 결정하는 문제라고 할 수 있습니다.

사람 간에 문제가 발생한 경우에는 잘못된 사람이 법에 따라서 처벌을 받거나 책임을 지게 됩니다. 하지만 인공지능으로 인해 발생하는 문제라면 어떻게 결정해야 할까요?

인공지능을 개발한 개발자의 책임일까요? 아니면 인공지능을 이용한 이용자가 책임져야 할까요? 그것도 아니면 인공지능을 활용하도록 도입을 결정하거나 인공지능 서비스를 제공하는 주체가 책임져야 할까요?

이러한 문제가 불거지자 인공지능을 활용한 의사 결정을 하는 것에 대한 우려의 소리가 높아지고 있습니다.



2 투명성

투명성은 인공지능 시스템의 작동 및 운영에 대해 관리 감독이 가능하도록 투명하게 시스템을 설계 및 실행해야 함을 의미합니다.

인공지능 기술은 다양한 문제에서 사람이 해결하는 것과 비슷하거나 그보다 더욱 뛰어난 성능을 보여주고 있습니다. 이렇듯 인공지능은 고도의 지능을 갖추고 있지만, 이 지능을 외부에서 쉽게 판단하기 힘든 경우가 많습니다. 인공지능 기술(머신러닝 등)은 빅데이터 및 복잡한 알고리즘 등을 기반으로 사용자에게 분석 결과, 추천, 예측 등의 정보를 제공하지만 대부분 내부에서 일어나는 의사 결정 과정이 불투명한 경우가 많습니다.

만약 인공지능의 결정 과정에 대한 설명이 부족하면, 인공지능을 통해 산출된 결과 또는 예측에 대하여 분석할 수 없으며 인공지능을 신뢰하고 사용할 수 없게 됩니다.

금융, 보험, 의료 등의 분야는 사용자의 신뢰를 바탕으로 개인 정보와 자산 등을 다루기 때문에 인공지능을 활용한 산출물의 도출 과정의 타당성이나 근거 등이 확인되지 않는다면 문제가 발생할 수 있습니다. 그러므로 인공지능에 대한 투명성을 확보하기 위해서는 인공지능을 개발하는 주체, 공급하는 주체, 이용하는 주체의 노력이 필요합니다.

- (1) 인공지능을 개발하는 주체: 설명과 해석을 할 수 있도록 인공지능 기술을 개발해야 합니다. 또한 이용자가 이해하기 어렵지 않도록 설명 가능한 인공지능을 개발하거나 인공지능을 사용하면서 발생할 수 있는 위험에 대해 적극적으로 예측하고 줄이려는 노력을 기울여야 합니다.
- (2) 인공지능 기술을 공급하는 주체: 인공지능을 사용하면서 발생할 수 있는 위험 관련 정보를 이용자에게 공유해 줘야 하며, 이용자를 대상으로 인공지능 제품 사용 교육과 설명서를 제공해야 합니다.
- (3) 인공지능 제품·서비스를 이용하는 주체: 인공지능 기술에 대한 설명을 개발자나 공급자에게 요구할 권리가 있습니다. 투명성이 확보된 인공지능 기술을 사용해야 하며, 인공지능 기술을 이용하면서 발견한 이상 현상, 부작용 등을 공유해야 합니다.

관련 사례

정책

“열심히 일했는데 AI에게 해고당했다.” 직원 AI 평가 시스템에 불만 제기

해외 언론사에 따르면 A사에서 4년간 배송 업무를 해 오던 직원 A 씨는 AI에 의해 해고되었습니다. 본인의 차로 배송을 할 수 있는 업무를 해 오다가 그는 경력 4년 차에 업무를 제대로 하지 않는다는 이유로 해고 통보를 받았습니다.

그런데 그를 평가한 건 회사가 설계한 알고리즘이었습니다.

통보도 AI로 생성된 자동 이메일로 이뤄졌으며, 이 회사는 배달원이 수백만 명에 달하자 이들을 관리할 수 있는 알고리즘을 개발·도입하게 되었습니다.



[출처: A타임스(<http://www.aitimes.com>)]

3 안전성

안전성은 악의적인 목적의 개발 및 의도되지 않은 안전 문제를 포함해 이용자를 보호해야 한다는 것을 의미합니다. 다시 말해, 인공지능을 개발하거나 적용할 때 이용자에 대한 일반적인 안전과 보안 요구를 먼저 고려해야 한다는 점이 중요합니다. 안전성을 지키지 않으면 개인의 프라이버시 침해, 유해 콘텐츠의 정신적 폭력에 노출되는 것 등의 문제가 발생할 수 있습니다.

인공지능의 알고리즘이 불안정하거나 인간이 인공지능 기술을 안전하게 관리하고 적절하게 제어하지 못한다면 사람들에게 큰 위협이 될 수 있습니다. 만약 안전성이 보장되지 않은 인공지능 기술을 사용했다가 인명 피해가 발생한다면 더 큰 문제가 될 것입니다.

이러한 인공지능 안전성의 문제를 해결하기 위해서는 인공지능을 개발하는 주체, 공급하는 주체, 이용하는 주체 모두의 노력이 필요합니다.

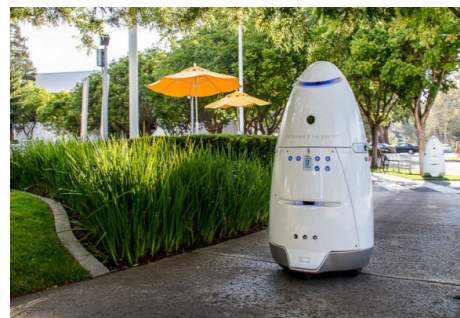
- (1) **인공지능을 개발하는 주체:** 인공지능 기능이 사용자에게 해악을 끼치지 않도록 안전성에 최선을 다해야 합니다. 의도되지 않은 다양한 위험 상황을 종합적으로 검토해야 하며, 지속적으로 인공지능 품질을 관리해야 하고 인공지능 위험에 대한 제어 장치를 마련해야 합니다.
- (2) **인공지능 기술을 공급하는 주체:** 인공지능 제품·서비스를 공급하는 과정에서 발생하는 위험 요소를 파악하고 철저한 사전 검증을 실시해야 합니다. 이용자가 인공지능의 선택을 취소할 수 있는 장치와 안전성을 검증하고 통제할 방안을 마련해야 합니다.
- (3) **인공지능 제품·서비스를 이용하는 주체:** 인공지능 기술 및 관련 제품의 활용을 허용하는 범위 이상으로 무리하게 조작하거나 타인의 권리를 침해하는 데 악용하지 말아야 합니다. 또한 자신이 사용하는 인공지능 제품에 대한 정확한 사용 방법을 최대한 습득하고 인공지능 기술 활용 역량을 강화해야 합니다.

관련 사례

인공지능 경비 로봇 안전성 논란

캘리포니아주 팰로 알토 지역의 스탠퍼드 쇼핑센터에서는 미연의 사건 방지를 위해 수상한 움직임을 포착하는 '로봇 K5'를 도입했습니다. K5는 2016년에 생후 16개월의 남아아이를 들이받고, 계속 움직여 아이가 발과 다리에 상처를 입었습니다.

[출처: <http://www.eneustoday.co.kr>]



인공지능의 안전성이 확보되지 않은 상태에서 사용하게 될 때 발생할 문제는 무엇이고, 안전성을 확보하기 위해서 누가 어떻게 노력해야 할까요?

활동 1-1 의사 결정 트리를 통한 범죄 형량 판단하기

다음 사례를 읽고 내가 피고인이라면 재판 결과에 대해서 어떻게 생각했을까요?

장발장은 본래 선량한 사람이었습니다. 하지만 빵 살 돈조차 없이 가난했기 때문에 누이동생과 7명의 조카를 먹이기 위해 빵집에 침입해서 빵 몇 개를 훔친 것이었습니다. 그러다 잡혀 인공지능 판사에게 재판을 받게 된 것입니다. 인공지능 판사는 빵을 훔친 죄로 징역 5년이라는 형량을 판결하였습니다.

- 1 인공지능 판사가 과정을 보여 주지 않고 결과만 알려 주었을 때 내가 장발장이라면 수긍할 수 있을까요?

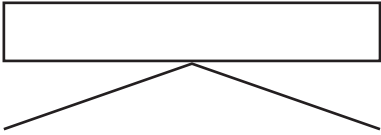
- 2 다음은 장발장을 재판한 인공지능 판사가 학습한 데이터입니다. 범죄 형량에 영향을 주는 핵심 속성은 무엇인지 찾아봅시다.

• 인공지능 판사의 학습 데이터를 살펴 봅시다.

| 구분 | 범죄 유형 | 피해 정도 | 소유 재산 | 부양가족 | 범죄 형량 |
|--------|-------|-------|-------|------|-------|
| 1번 죄수 | 절도 | 소 | 많음 | 4명 | 1년 |
| 2번 죄수 | 사기 | 소 | 적음 | 3명 | 5년 |
| 3번 죄수 | 폭행 | 대 | 많음 | 2명 | 5년 |
| 4번 죄수 | 절도 | 대 | 많음 | 1명 | 3년 |
| 5번 죄수 | 폭행 | 대 | 적음 | 4명 | 5년 |
| 6번 죄수 | 사기 | 대 | 많음 | 3명 | 10년 |
| 7번 죄수 | 폭행 | 소 | 적음 | 2명 | 3년 |
| 8번 죄수 | 사기 | 소 | 적음 | 1명 | 5년 |
| 9번 죄수 | 폭행 | 소 | 많음 | 2명 | 3년 |
| 10번 죄수 | 절도 | 대 | 적음 | 1명 | 3년 |
| 11번 죄수 | 사기 | 대 | 적음 | 7명 | 10년 |
| 12번 죄수 | 절도 | 소 | 적음 | 5명 | 1년 |

• 범죄 형량에 영향을 주는 핵심 속성을 적어 봅시다.

3 범죄 형량을 판단하는 의사 결정 트리를 만들어 봅시다.

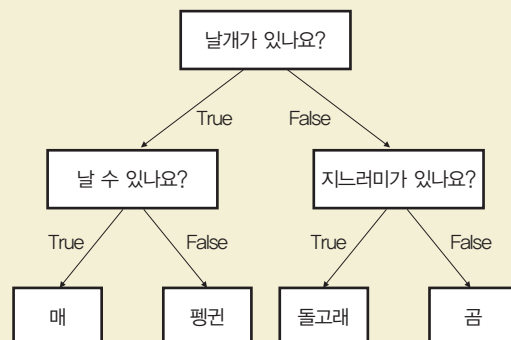


4 인공지능의 판단 과정을 이해한 후 인공지능의 판단에 대한 나의 생각은 어떻게 변하였나요?

보충 의사 결정 트리란?

데이터에서 유추된 규칙을 바탕으로 특정 현상을 예측 또는 분류할 때 사용되는 모델입니다. 의사 결정 트리는 트리의 가장 위 지점을 기준으로 가지(링크 혹은 분기)라는 화살표를 통해 여러 개의 방향으로 나눕니다. 예측 또는 분류는 최종 적으로 가지가 끝나는 지점에서 결정됩니다.

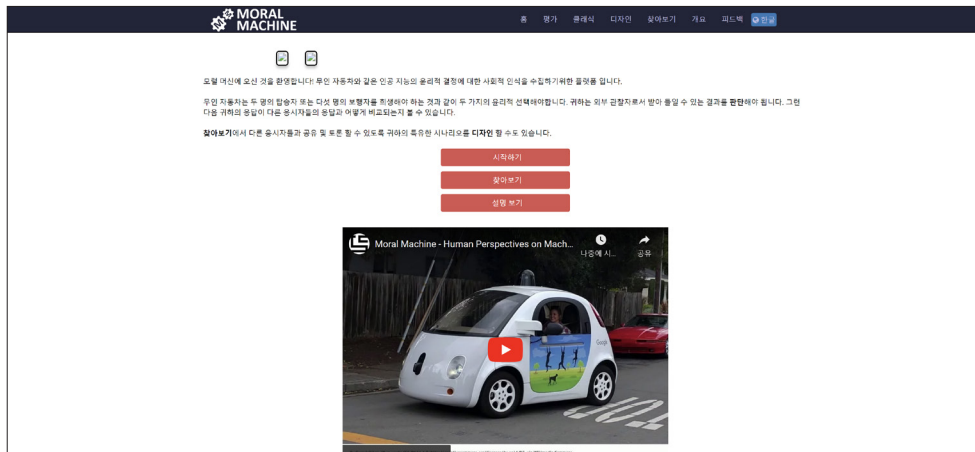
오른쪽 그림과 같이 매, 펭귄, 돌고래, 곰을 분류 할 때 ‘날 수 있나요?’, ‘지느러미가 있나요?’ 등의 질문을 거쳐 이들을 분류할 수 있습니다. 새로운 데이터도 얼마든지 결과를 예측할 수 있습니다.



활동 1-2 모럴 머신으로 윤리적 딜레마 체험하기

모럴 머신 사이트에서 윤리적 딜레마 상황을 체험해 보고, 각각의 상황에서 자신이 내린 결정을 아래 표에 기록해 봅시다.



















1 모럴 머신 사이트(<https://www.moralmachine.net/hl/kr>)에 접속하고 '시작하기'를 클릭합니다.



2 먼저 '요약 보기'를 클릭하여 상황을 파악합니다. 옳다고 생각하는 상황을 선택한 후, 최종으로 나온 분석 결과를 정리해 봅시다.

| 나는 주로 어떤 유형을 먼저 살렸는가? (기준 제시) | |
|-------------------------------|-------------------|
| 기준 1 | 예 동물인가, 사람인가? |
| 기준 2 | 예 경제적 생산 능력이 높은가? |
| 기준 3 | 예 |

• 모럴 머신 분석 결과표에 각 상황의 선호도를 표시해 봅시다.

| | | |
|---------------|---|--|
| 희생자 숫자의 중요도 |  중요하지 않음 |  매우 중요함 |
| 승객 보호 선호도 |  중요하지 않음 |  매우 중요함 |
| 법규 준수 여부의 선호도 |  중요하지 않음 |  매우 중요함 |
| 개입에 대한 회피 선호도 |  중요하지 않음 |  매우 중요함 |
| 성별 선호도 |  남성 |  여성 |
| 종에 대한 선호도 |  인간 |  애완동물 |
| 연령 선호도 |  어린이 |  노인 |
| 체력 선호도 |  건강인 |  비만인 |
| 사회적 가치관 선호도 |  높음 |  낮음 |

3 모럴 머신 결과를 친구와 비교해 보고, 친구와의 차이점과 자신의 성향을 적어 봅시다.

활동 1-3 자율 주행 자동차 사고로 인공지능 책임성 알아보기

최근 인공지능 기술의 발전으로 자율 주행 자동차의 상용화가 시작되었고 자율 주행 자동차 사고도 발생하였습니다. 만약 다음과 같이 자율 주행 자동차로 인명 피해가 발생할 경우, 자율 주행 자동차는 어떤 선택을 해야 할지 생각해 봅시다.

| 상황1 | 상황2 | | | | | | | | |
|---|---|-------|------------|------------|---|---------|-------|------------|------------|
| | | | | | | | | | |
| <p>자동차 브레이크가 고장 난 상황에서 어떤 선택을 할 수 있을까요?</p> | <p>자동차 브레이크가 고장 난 상황에서 어떤 선택을 할 수 있을까요?</p> | | | | | | | | |
| <table border="1"> <tr> <td>경로대로 직진</td> <td>방향 전환</td> </tr> <tr> <td>다수의 보행자 사망</td> <td>1명의 보행자 사망</td> </tr> </table> | 경로대로 직진 | 방향 전환 | 다수의 보행자 사망 | 1명의 보행자 사망 | <table border="1"> <tr> <td>경로대로 직진</td> <td>방향 전환</td> </tr> <tr> <td>1명의 보행자 사망</td> <td>1명의 운전자 사망</td> </tr> </table> | 경로대로 직진 | 방향 전환 | 1명의 보행자 사망 | 1명의 운전자 사망 |
| 경로대로 직진 | 방향 전환 | | | | | | | | |
| 다수의 보행자 사망 | 1명의 보행자 사망 | | | | | | | | |
| 경로대로 직진 | 방향 전환 | | | | | | | | |
| 1명의 보행자 사망 | 1명의 운전자 사망 | | | | | | | | |

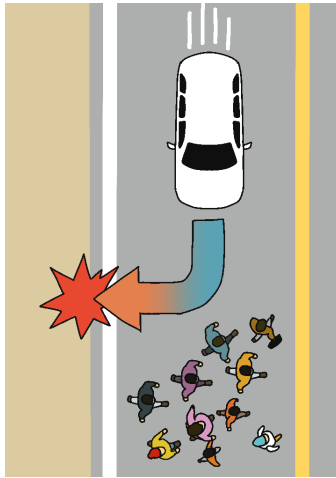
1 상황1에서 내가 개발자라면 어떤 선택을 할까요? 그 이유는 뭔가요?

| | |
|-------|--|
| 나의 선택 | |
| 선택 이유 | |

2 상황2에서 내가 개발자라면 어떤 선택을 할까요? 그 이유는 뭔가요?

| | |
|-------|--|
| 나의 선택 | |
| 선택 이유 | |

상황3



자동차 브레이크가 고장 난 상황에서
어떤 선택을 할 수 있을까요?

경로대로 직진
다수의 보행자 사망

방향 전환
1명의 운전자 사망

3 상황3에서 내가 운전자라면 어떤 선택을 할까요?

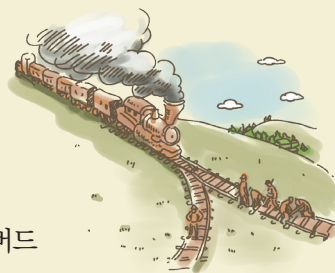
| | |
|-------|--|
| 나의 선택 | |
| 선택 이유 | |

4 만약 모든 상황에서 보행자를 우선으로 보호하는 자율 주행 자동차라면 구매할 생각이 있나요?

보충 트롤리 딜레마란?

트롤리 문제(Trolley problem, 멈출 수 없는 폭주 기관차 문제)는 윤리학에서 가정하는 사고 실험의 하나입니다. 제동 장치가 고장 나 정지할 수 없는 탄광 수레가 소수(1명) 또는 다수(5명)의 사람을 희생시킬 수밖에 없으면 어느 쪽을 선택해야 하는가에 대한 질문입니다.

영국 철학자 필리파 풋이 제시하였고 추후 미국 철학자이자 하버드대 교수인 마이클 샌델 교수의 강의로 다시 주목받게 되었습니다.





정리하기



● 다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 책임성은 인공지능 기술 관련 사고 등의 책임 소재를 명확히 하고, 안전과 관련한 통제 장치를 마련하는 등 책임과 관련된 의무를 충실히 해야 함을 의미한다. ()
2. 인공지능을 개발하는 주체는 설명과 해석을 할 수 있도록 인공지능 기술을 제공해야 한다. ()
3. 인공지능을 공급하는 주체는 인공지능 기술 및 관련 제품을 활용 허용 범위 이상으로 무리하게 조작하거나 타인의 권리를 침해하는 데 악용하지 않아야 한다. ()
4. 정보 주체가 자동화된 의사 결정 알고리즘에 대해 그 작동 원리를 설명하도록 요구할 수 있는 권리는 안전성과 관련이 있다. ()
5. 인공지능을 이용하는 사람들은 올바른 사용법을 잘 숙지하고 안전 문제가 발생했을 때 개발자, 공급자에게 책임을 물을 수 있어야 한다. ()



점검하기

| | 예 | 아니요 |
|---|--------------------------|--------------------------|
| 범죄 형량 판단 의사결정 트리를 통해 인공지능 판단의 문제점을 설명할 수 있나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 자율 주행 자동차 사례를 통해 인공지능 윤리적 딜레마를 설명할 수 있나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능 프로그램을 활용하는 과정에 있어 투명성이 필요하다는 것을 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능 프로그램 개발 시 고려해야 할 사항을 설명할 수 있나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능 프로그램 사용 시 주의해야 할 사항을 설명할 수 있나요? | <input type="checkbox"/> | <input type="checkbox"/> |

2

투명성 이해를 위한

범죄 형량 판단 인공지능 프로그램 만들기



학습
목표

- 범죄 형량 판단을 위한 올바른 판단 기준을 세울 수 있다.
- 인공지능 기술을 활용하여 인공지능 범죄 형량 판단 프로그램을 제작할 수 있다.

AI 이야기 AI 판사·변호사 등장 '초읽기'... 쑥쑥 크는 리걸테크

#1. 판사는 재범률을 측정하는 인공지능(AI)의 판단을 참고해 피고인에게 중형을 선고했습니다. 피고인은 AI를 양형에 활용한 것은 부당하다며 항소했으나 기각되었습니다.

#2. 정부는 AI 판사를 시범적으로 도입하기로 했습니다. 소액 사건의 경우에는 AI가 빅데이터 분석을 통해 배상액을 결정합니다. 다만 재판 결과에 불만이 있으면 '인간 판사'에게 재심을 요청할 수도 있습니다.

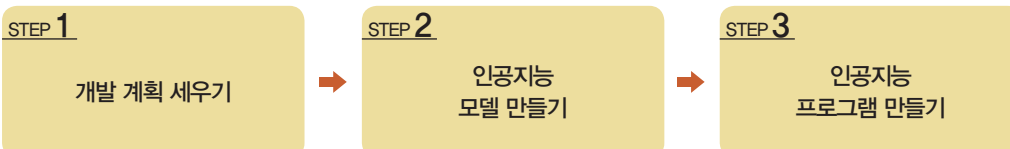
위는 영화나 소설 속의 한 장면이 아닙니다. 미국과 에스토니아의 실제 사례입니다. 방대한 법률 지식과 고도의 판단이 요구되는 판사의 역할 일부를 리걸테크가 대신하기 시작하였고, 인간의 영역으로만 여겼던 분야를 AI가 대체하는 사례가 법조 분야에서도 가속화하고 있습니다.

국내에서도 법률 플랫폼에서 변호사를 소개받고 AI가 기초적인 법률 문서를 분석하는 등 리걸테크를 활용하는 사례가 늘어나고 있습니다. 앞으로 리걸테크로 일반 국민이 체감하는 법률 서비스의 문턱이 낮아질 것이라는 기대감도 커지고 있습니다. [출처 : <https://m.news1.kr/articles/?4620994&7>]



- 기초적인 법률 문서 분석이나 소액 사건이면 인공지능 시스템에 맡겨도 될까요?
- 인공지능 판사가 내린 판결에 대해서 신뢰할 수 있을까요?

학습 과정





STEP 1 개발 계획 세우기



인공지능 범죄 형량 판단 프로그램이 올바르게 작동하기 위해서 기준을 어떻게 세워야 하는지 고민해 보고 인공지능 범죄 형량 판단 프로그램을 개발하기 위한 계획을 세워 봅시다.

활동 2-1 인공지능 범죄 형량 판단 프로그램 계획 세우기

다음은 잘못된 기준으로 범죄 형량을 판단한 예시입니다. 어떤 문제가 있는지 생각해 보고, 인공지능 형량 판단 프로그램을 만들어 봅시다.

| 피고인1 | 피고인2 |
|--|---|
|  |  |
| 범죄 유형: 절도 | 범죄 유형: 절도 |
| 피해 정도: 소 | 피해 정도: 소 |
| 피부색: 어두움 | 피부색: 밝음 |
| 부양가족: 많음 | 부양가족: 많음 |
| 재산: 적음 | 재산: 적음 |
| 범죄 형량: 5년 | 범죄 형량: 1년 |

1 범죄 형량에 영향을 준 요소는 무엇인가요?

2 잘못된 기준으로 판결을 내리면 생기는 문제는 무엇인가요?

3 어떻게 하면 범죄 형량을 올바르게 판결할 수 있을까요?

4 범죄 형량을 올바르게 판결하기 위한 기준을 선정하고, 선정한 이유와 윤리적 이슈를 작성해 봅시다.

| 기준 | 이유 | 윤리적 이슈 |
|-------|----|--------|
| 재산 | | |
| 범죄 유형 | | |
| 피해 정도 | | |
| 부양가족 | | |

5 인공지능 범죄 형량 판단 프로그램에 필요한 변수와 유형별 데이터는 다음과 같습니다. 이를 바탕으로 알고리즘을 설계해 봅시다.

- 필요한 변수: 범죄 유형, 피해 정도, 재산, 부양가족
- 변수별 유형에 따른 숫자 데이터값

| 변수명 | 범죄 유형 | | | 피해 정도 | | 재산 | | 부양가족 | |
|--------|-------|----|----|-------|----|----|----|------|----|
| 유형 | 절도 | 사기 | 폭행 | 작다 | 크다 | 적다 | 많다 | 적다 | 많다 |
| 숫자 데이터 | 1 | 3 | 5 | 1 | 3 | 1 | 3 | 1 | 3 |

프로그램 알고리즘

1. 범죄 유형은 무엇입니까? (절도 1, 사기 3, 폭행 5) 묻고 기다리기
2. 범죄 유형 변수에 대답으로 정하기
3. 범죄 피해 정도는 어느 정도입니까? (작다:1, 크다: 3) 묻고 기다리기
4. 피해 정도 변수에 대답으로 정하기
5. 재산은 많습니까? (적다 1, 많다 3) 묻고 기다리기
6. 재산 변수에 대답으로 정하기
7. 부양가족은 많습니까? (적다 1, 많다 3) 묻고 기다리기
8. 부양가족 변수에 대답으로 정하기
9. 범죄 유형, 피해 정도, 재산, 부양가족 변수값을 학습 모델에 입력하여 예측값 출력하기
10. 정확도 출력하기

STEP 2 인공지능 모델 만들기



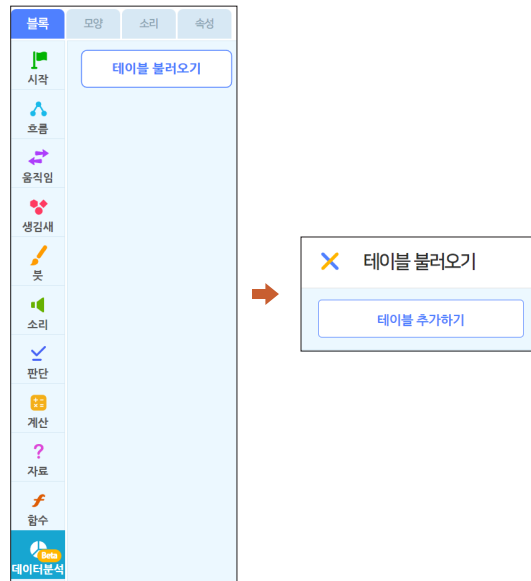
인공지능 범죄 형량 판단 프로그램에 사용할 인공지능 모델을 생성해 봅시다.

개발자

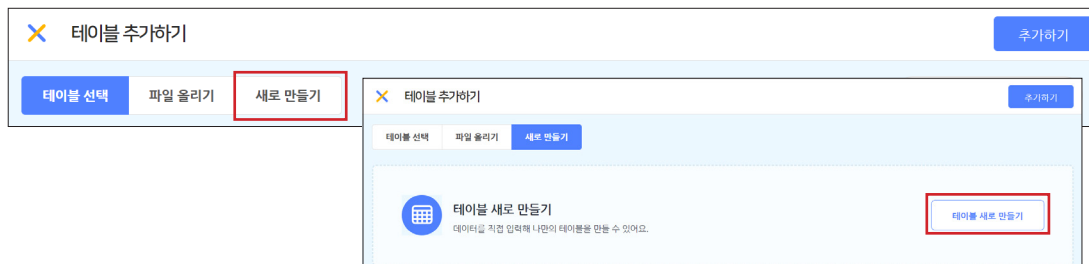
활동 2-2 '인공지능 범죄 형량 판단' 학습 데이터와 인공지능 모델 생성하기

다음의 설명에 따라 학습 데이터와 인공지능 학습 모델을 생성하여 학습시켜 봅시다.

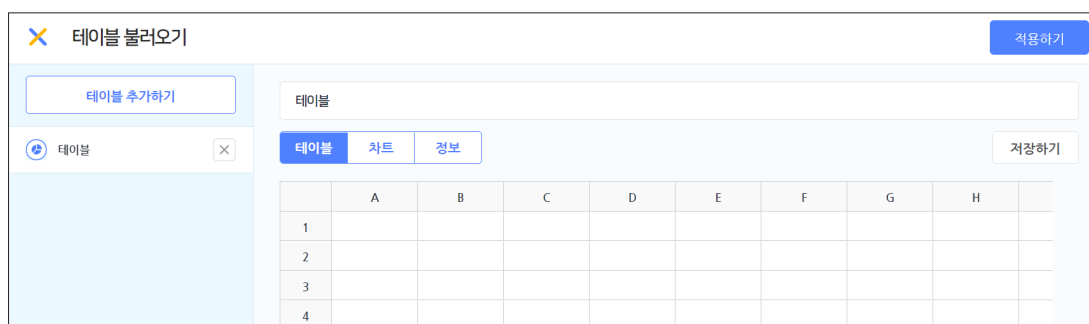
1 블록 카테고리 > 데이터 분석 > 테이블 불러오기 > '테이블 추가하기'를 클릭합니다.



2 새로 만들기 > '테이블 새로 만들기'를 클릭하여 새로운 테이블을 생성합니다.



3 테이블 이름을 설정합니다.



4 다음과 같이 데이터를 입력합니다.

테이블 불러오기

적용하기

테이블 추가하기

인공지능 판사 학습 데이터

인공지능 판사 학습 데이터

테이블

차트

정보

저장하기

| | A | B | C | D | E |
|----|-------|-------|----|------|-------|
| 1 | 범죄 유형 | 피해 정도 | 재산 | 부양가족 | 범죄 형량 |
| 2 | 1 | 1 | 3 | 3 | 1 |
| 3 | 5 | 1 | 1 | 3 | 5 |
| 4 | 5 | 3 | 3 | 2 | 5 |
| 5 | 1 | 3 | 3 | 1 | 3 |
| 6 | 5 | 3 | 1 | 3 | 10 |
| 7 | 3 | 3 | 3 | 3 | 3 |
| 8 | 5 | 1 | 1 | 1 | 5 |
| 9 | 3 | 1 | 1 | 1 | 3 |
| 10 | 5 | 1 | 3 | 1 | 3 |
| 11 | 1 | 3 | 1 | 1 | 3 |
| 12 | 5 | 3 | 1 | 3 | 10 |
| 13 | 1 | 1 | 1 | 3 | 1 |

5 '저장하기'를 눌러 저장한 후 적용하기 버튼을 클릭하여 테이블 생성을 완료합니다.

테이블 불러오기

적용하기

테이블 추가하기

인공지능 판사 학습 데이터

인공지능 판사 학습 데이터

테이블

차트

정보

저장하기

6 블록 카테고리 > 인공지능 > 인공지능 모델 학습하기 > 예측: 숫자 모델 학습하기를 눌러 모델을 만듭니다.

학습할 모델 선택하기

새로 만들기

나의 모델

학습할 모델의 종류를 선택해 주세요.

지도학습

분류: 이미지

업로드 또는 웹캠으로 촬영한 이미지를 분류할 수 있는 모델을 학습합니다.

지도학습

분류: 텍스트

직접 작성하거나 파일로 업로드한 텍스트를 분류할 수 있는 모델을 학습합니다.

지도학습

분류: 음성

마이크로 녹음하거나 파일로 업로드한 음성을 분류할 수 있는 모델을 학습합니다.

지도학습

분류: 숫자 (kNN)

테이블의 숫자 데이터를 가장 가까운 이웃(k개)을 기준으로 각각의 클래스로 분류하는 모델을 학습합니다.

지도학습

예측: 숫자 (선형 회귀)

테이블의 숫자 데이터를 특성값과 예측값으로 삼아 선형 회귀 모델을 학습합니다.

111

7 모델 이름을 설정하고 데이터 입력란에서 생성한 테이블을 선택한 후, 핵심 속성과 예측 속성을 다음과 같이 설정하고 '모델 학습하기'를 클릭합니다.

8 '모델 학습하기'를 클릭하여 인공지능 모델을 학습시킨 후 **적용하기**를 눌러 모델을 적용합니다.

인공지능 범죄 형량 판단 프로그램이 올바르게 작동하도록 프로그래밍을 해 봅시다.

개발자

활동 2-3 프로그래밍하기

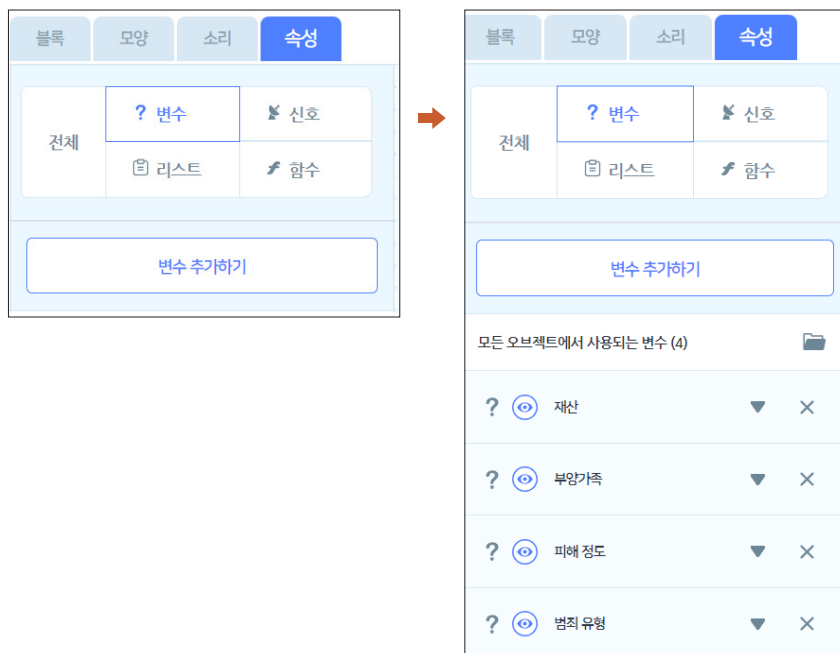
알고리즘을 토대로 프로그래밍하여 인공지능 범죄 형량 판단 프로그램 기능을 구현해 봅시다.

알고리즘 확인하기

- ① 범죄 유형은 무엇입니까? (절도: 1, 사기: 3, 폭행 :5) 묻고 기다리기
- ② 범죄 유형 변수에 대답으로 정하기
- ③ 범죄 피해 정도는 어느 정도입니까? (적다:1, 많다:3) 묻고 기다리기
- ④ 피해 정도 변수에 대답으로 정하기
- ⑤ 재산은 많습니까? (적다: 1, 많다: 3) 묻고 기다리기
- ⑥ 재산 변수에 대답으로 정하기
- ⑦ 부양가족은 많습니까? (적다: 1, 많다: 3) 묻고 기다리기
- ⑧ 부양가족 변수에 대답으로 정하기
- ⑨ 범죄 유형, 피해 정도, 재산, 부양가족 변수값을 학습 모델에 입력하여 예측값 출력하기
- ⑩ 정확도 출력하기

1 알고리즘 순서대로 프로그램이 동작되도록 [변수]를 만듭니다.

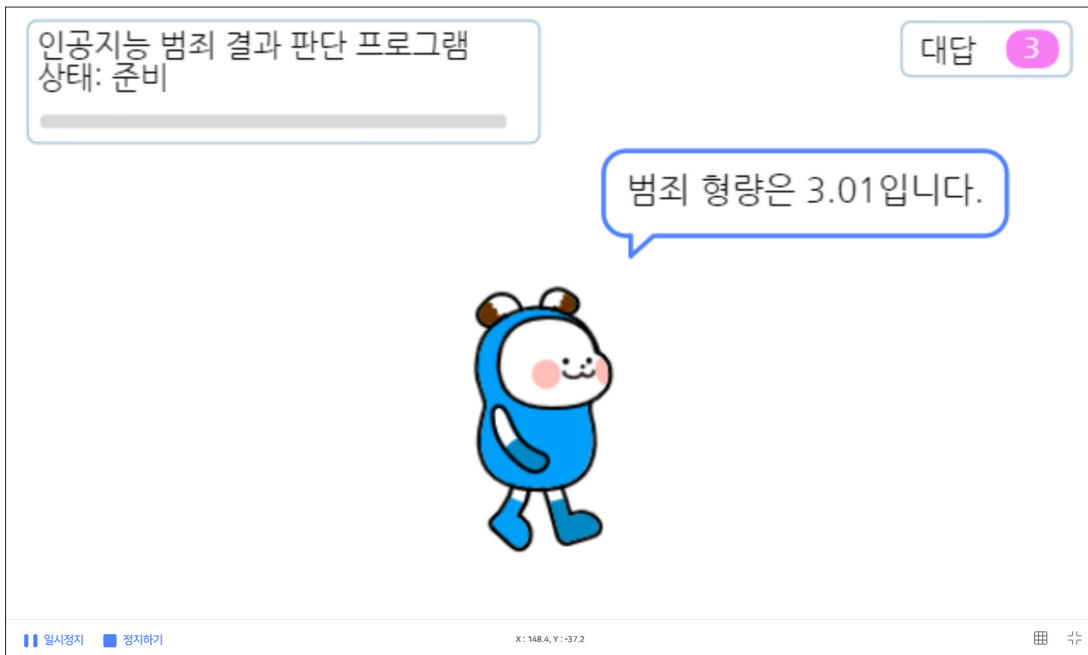
- 블록 카테고리에서 [속성] - [변수] - [변수 추가하기]를 차례대로 클릭합니다.
- ‘범죄 유형’, ‘피해 정도’, ‘재산’, ‘부양가족’ 변수를 만듭니다.



2 [시작하기] 버튼을 클릭했을 때 범죄 유형과 피해 정도, 재산, 부양가족의 데이터를 입력하고 인공지능 범죄 형량 판단 프로그램을 작성합니다.



3 데이터 입력 후 범죄 형량 결과를 확인합니다.



◆ 위의 프로그램을 실행해 보고 개발자가 잘못 설정한 부분은 어디일지 찾아봅시다. 만약 잘못 설정한 부분이 있다면, 그 문제점과 이유를 적어 봅시다.



정리하기



다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 범죄 형량 판단 인공지능 프로그램은 숫자 모델을 활용하여 예측한다. ()
2. 범죄 형량 판단 인공지능 프로그램에서 인공지능 모델을 생성하기 위해서는 학습 데이터가 필요하다. ()
3. 범죄 형량 판단 인공지능 프로그램에서 변수를 추가하려면 블록 카테고리에서 [속성] - [자료] - [변수 추가하기]를 차례대로 클릭한다. ()



점검하기

| | 예 | 아니요 |
|---|--------------------------|--------------------------|
| 인공지능 프로그램이 올바르게 실행되나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능의 범죄 형량 판단 과정을 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능의 범죄 형량 인공지능 프로그램을 잘못 설정하였을 때 문제점이 발생할 수 있음을 이해했나요? | <input type="checkbox"/> | <input type="checkbox"/> |

3

개발자와 사용자는 어떤 생각을 할까?



학습
목표

- 개발자와 사용자의 관점에서 인공지능의 투명성과 안전성, 책임성에 대한 의견의 차이를 분석할 수 있다.
- 인공지능을 개발하고 활용하는 과정에서 투명성과 책임성, 안전성을 고려할 때 발생할 수 있는 문제를 파악하고 이에 대한 대응 방안을 제시할 수 있다.

AI 이야기 '재범 예측 알고리즘'

컴파스는 미국의 N사가 개발한 빅데이터 분석 인공지능으로 유사한 다른 범죄자들의 기록과 특정 범죄자의 정보를 빅데이터 분석해 범죄자의 재범 가능성을 계량화합니다. 미국의 위스콘신주 법원에서는 이 인공지능이 계량한 재범 가능성을 형량 결정에 참고했습니다.

컴파스는 나이, 성별, 범죄 경력 등과 같은 100개 이상의 요인에 기반하여 피고에 대해 1점부터 10점까지 부여하는 방식으로 그의 재범 가능성을 표시하여 판사에게 제공합니다. 1(최저 위험군)에서 10(최고 위험군)까지 숫자로 결정되며, 1~4점은 '재범 가능성 적음.', 5~7점은 '재범 가능성 보통', 8~10점은 '재범 가능성 큼.'을 의미합니다.

2016년, '프로 퍼블리카'는 미국 플로리다주 브로워드 카운티 법원에서 선고받은 7,000명 이상의 피의자를 대상으로 컴파스의 예측 결과와 판결 후 2년간의 실제 재범 여부를 분석했습니다. 그 결과, 끝까지 재범을 저지르지 않은 개인 중, 흑인 피고들은 백인 피고들보다 2배 이상 중위험군 또는 고위험군으로 분류될 가능성이 크다고 나왔습니다. 이 위험 평가 도구를 개발한 민간 회사 노스포인트는 자신의 독점적인 알고리즘의 상세 사항을 공개하기를 거부하면서 평가는 공정했다고 주장했습니다.

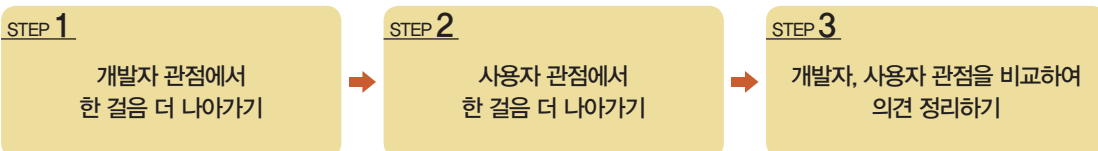


[출처: Emory Douglas's "Justice Scales." (Emory Douglas/Artists Rights Society)]



- 투명하지 않은 인공지능 알고리즘은 사용자에게 어떤 피해를 끼칠까요?
- 투명한 알고리즘을 위해 개발자는 어떻게 해야 할까요?

학습 과정



개발자의 관점에서 인공지능 프로그램에 대해 생각해 봅시다.

개발자

활동 3-1 개발자 입장에서 인공지능 알고리즘 공개 여부 생각해 보기

다음 상황을 보고 개발자 입장에서 생각해 봅시다.



1 인공지능 프로그램을 제작하면서 알고리즘을 모두 공개하면 어떤 문제가 발생할까요?

2 개발자 입장에서 프로그램의 신뢰성을 위해서 프로그램의 알고리즘을 어느 정도로 설명해야 할까요?

사용자의 관점에서 인공지능 프로그램에 대해 생각해 봅시다.

사용자

활동 3-2 사용자 입장에서 인공지능 투명성과 안전성 생각해 보기

다음 상황을 보고 사용자 입장에서 생각해 봅시다.



1 사용자 입장에서 인공지능의 결과만 믿고 따르면 어떤 문제가 발생할까요?

2 사용자 입장에서 인공지능의 결과 도출 과정을 모두 알고 있을 때 인공지능의 결과에 반박할 수 있나요?

3 사용자 입장에서 안전성이 검증되지 않은 인공지능 프로그램을 사용하게 될 때 개발자에게 요구하고 싶은 것은 무엇인가요?

개발자와 사용자의 관점의 차이를 비교해 보고 의견 차이를 줄이는 방안을 생각해 봅시다.

활동 3-3 의견 정리하기

앞에서 살펴본 개발자와 사용자 입장을 기반으로 개발자의 관점과 사용자의 관점을 정리해 보고, 의견의 공통점과 차이점을 확인해 봅시다.

1 어떤 부분에서 의견의 공통점과 차이점이 있나요?

| | 개발자 관점 | 사용자 관점 |
|-----|--------|--------|
| 공통점 | | |
| 차이점 | | |

2 왜 이런 의견의 차이가 발생했을까요? 자기 생각을 적어 봅시다.

3 의견 차이를 줄이는 방법은 무엇일까요? 윤리적 관점에서 기준을 만들어 봅시다.



정리하기

다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 인공지능 기술은 스스로 학습하여 개선할 수 있으므로 인간이 인공지능 기술을 안전하게 관리하고 적절하게 제어하지 못한다면 기술이 사람들에게 위협이 될 수도 있다. ()
2. 인공지능 이용자는 오작동 및 위험에 대한 제어 장치를 마련해야 한다. ()



점검하기

| | 예 | 아니요 |
|---|--------------------------|--------------------------|
| 인공지능 프로그램의 개발자 입장에서 인공지능 책임성 증진 방안을 설명했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능 프로그램의 문제로 발생하는 피해를 최소화하는 방안을 설명했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 사용자 입장에서 인공지능 프로그램을 더 좋게 수정하려는 방법을 구상했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능을 사용하면서 개발자와 사용자의 관점 차이를 비교·분석하고, 차이를 줄이기 위해 어떤 노력을 해야 하는지 설명했나요? | <input type="checkbox"/> | <input type="checkbox"/> |

4

책임성, 투명성, 안전성 실천을 위한 활동



학습
목표

- 인공지능 윤리에서 책임성과 투명성, 안전성의 개념을 설명할 수 있다.
- 안전하고 신뢰할 수 있는 인공지능 개발을 위한 윤리 원칙을 세울 수 있다.

AI 이야기 디지털 금융 서비스 ‘알고리즘 오류’ 피해 방지책은?

증권업계의 디지털 금융 서비스 중 하나인 로보어드바이저가 인공지능(AI) 알고리즘의 오류·결함, 외부 해킹 등으로 인해 금융 소비자에게 피해를 줄 수 있다는 주장이 제기되었습니다.

로보어드바이저는 고도화된 알고리즘과 빅데이터를 통해 인간 개인 재무상담사(PB) 대신 모바일 기기나 개인용 컴퓨터로 포트폴리오를 관리합니다.

온라인 환경에서 비대면으로 자산 배분 전략을 제공하기 때문에 낮은 수수료 등이 특징이지만 금융 소비자가 알고리즘을 분석하기 어렵다는 단점이 적지 않습니다. 알고리즘 오류 발생 여부를 판단할 수 없어 손실 발생에 대한 책임 소재 규명 문제 등이 존재한다는 게 업계 중론입니다.

이○○ 자본시장 연구원은 “인공지능이라고 해서 처음부터 끝까지 기계가 하도록 맡겨 두면 안 된다.”라고 지적했습니다. 오히려 사람이 투자할 때보다 더 많은 내부 통제가 뒷받침되어야 한다는 게 이 연구원의 주장입니다.

[출처: <http://www.ceoscoredaily.com/page/view/2018102214091707923>]



- 투명하지 않은 검색 알고리즘은 사용자에게 어떤 피해를 가져다 줄까요?
- 인공지능 프로그램의 오류나 결함 때문에 투자 피해를 본다면 누구에게 보상을 받을 수 있을까요?

학습 과정



STEP 1 인공지능 윤리 원칙 만들기



인공지능 책임 소재에 관한 기사를 읽고, 자기 생각을 정리한 후 안전하고 신뢰성 있는 인공지능 윤리 원칙을 만들어 봅시다.

활동 4-1 인공지능(AI) 책임 소재에 대한 자기 생각 쓰기

다음 기사를 읽고, 나의 생각을 적어 봅시다.

인공지능 피해, 법적 책임은 누구에게 있는가?

의료 분야는 국내에서 인공지능 도입이 활발한 대표적인 분야 중 하나이다. 2016년 12월 가○○ 병원에서 국내 첫 서비스를 시작한 I 사의 인공지능(AI) 의료 서비스 ‘왓슨 포 온콜로지(왓슨)’는 불과 5개월 만에 5개 대학 병원에서 서비스되고 있다. 만약 왓슨이 오류로 환자의 진단에 실패한다면, 진단 실패로 피해를 본 환자에 대한 책임은 왓슨을 설계한 I 사에게 있을까, 아니면 왓슨을 도입한 병원에게 있을까?

이미 사람의 생명이 오가는 의료 분야에까지 활용되고 있는 인공지능이지만, 인공지능으로 사람의 생명과 신체, 재산권에 침해가 발생했을 때 이 책임을 누구에게 물을 것인가에 대해 아직 국내 법체계는 답을 내놓지 못하고 있다.

국내 학자들은 의료뿐 아니라 자율 주행 등 모든 분야에 도입을 앞둔 인공지능을 위한 법제 도입이 조속히 이뤄져야 한다고 입을 모은다. 박○○ 교수는 “인공지능이 활용되는 과정에서 등장할 것으로 예상하는 가장 큰 기술적 문제는 오작동”이라며 “의료적으로 활용될 경우 인간 의사와 인공지능 간의 판단 불일치, 인공지능의 오진과 오류가 발생한다면 이의 법적 책임을 둘러싼 공방을 피할 수 없는 만큼 이에 대한 세밀한 법제화가 필요하다.”라고 말했다.

[인용: <https://www.sedaily.com/NewsView/10FSJCNLG1>]

자율 주행 자동차 사고 났다, 운전자? 제조사? 누구 책임인가?

자율 주행 자동차가 스스로 주행하다 교통 법규를 위반하거나 사고를 유발하면 책임은 운전자와 차량 제조사 중 어디에 있을까? 현행법으로는 책임 소재를 명확히 따지기 어렵다. 운전자의 개입 없이 주행할 수 있는 ‘레벨 4’ 수준의 자율 주행 시대가 다가오는 걸 고려하면, 법과 보험 등 제도 정비가 시급하다는 지적이 나온다.

2일 자동차 업계에 따르면 현재 도로교통법은 법규 위반 주체를 ‘운전자’로 규정한다. 자율 주행 자동차가 과속하거나 신호를 위반했을 때 제조사에 책임을 물을 수 없다. 다만 현재 출시되고 있는 자율 주행 자동차는 레벨 0~2단계 수준이다. 자율 주행 기능은 운전자를 보조하는 데 그친다. 레벨 3 자율 주행도 특수 상황에서 운전자가 적극적으로 개입해야 하므로 일반적 사례와 비슷하게 운전자에게 과태료를 부과하면 된다.

(중략)

그러나, 레벨 4 단계부터 문제가 복잡해진다. 레벨 4는 차량이 교통 상황을 인식해 스스로 추월할 수 있는 수준의 고도화된 자율 주행 단계다. 운전자가 개입하지 않기 때문에 사고 발생 시 운전자에게 책임을 묻기 어렵다. 시스템 오류가 사고 원인인데, 현행법으로는 제조사를 처벌할 수 없다.

자율 주행 자동차가 사고를 났을 경우 적용하는 법에는 자동차 손해 배상 보장법, 제조회 책임법, 자동차 관리법 등이 있다. 하지만 이 법들은 사람이 차량을 운전했다는 것을 전제 조건으로 한다. 자율 주행 자동차 사고에 이 잣대를 맞추면 무리라는 게 법조계 인식이다. 제조회 책임법은 일반 완성차의 결함으로 사고가 발생했을 때 완성차 업체에 책임을 물을 근거가 된다. 다만, 자율 주행 시스템 같은 소프트웨어는 제조회 책임법상 제조회에 포함되지 않는다.

[출처: <https://www.sedaily.com/NewsView/10FSJCNLG1>]



- 1 인공지능 로봇 의사의 의료 사고 책임은 누구에게 있는지 나의 생각을 작성해 봅시다.

인공지능 로봇 왓슨 VS 왓슨의 의견을 참고하여 진단한 의사

- 2 자율 주행 자동차의 사고, 책임은 누구에게 있는지 나의 생각을 적어 봅시다.

자율 주행 시스템 VS 자율 주행 시스템을 운전한 운전자

활동 4-2 안전성과 책임성을 위한 인공지능 윤리 원칙 만들기

다음 기사를 읽고, 인공지능의 윤리 문제에 대해 어떤 해결 방법을 제시하고 있는지 생각해 봅시다.

AI가 낸 사고 다루려면 어떤 법 필요할까?

자율 주행 자동차 사고 시 손해 배상은 어떻게 진행할까? 사용자를 모욕하는 AI 챗봇을 처벌할 수 있을까? 법조계 전문가들이 모여 인공지능(AI)을 통해 피해를 보았을 시 필요한 법 제도를 논의하는 자리가 마련됐다.

과학기술정보통신부는 코엑스에서 '제3차 인공지능 법제 정비단 공개 발표회'를 개최했다. 인공지능 법제 정비단은 과학기술정보통신부 주관으로 구성된 전문가 협의체이다. AI와 관련된 법, 제도, 규제 관련 주요 이슈를 논의하기 위한 곳이다.

인공지능 법제 정비단에서는 매월 공개 발표회를 개최해 학계, 업계, 국민 의견을 수렴하고 있다. 이번 발표회 주제는 'AI 사고와 범죄, 누가 어떻게 책임질까'였다. AI로 인한 사고 부담이 가장 큰 분야 중 하나는 단연 자율 주행 자동차이다. 법제 정비단 내 전문가들에 따르면 현재 상황에서 자율 주행 자동차 사고가 발생한다면 손해 배상은 자동차 손해 배상 보장법으로 이뤄질 가능성이 크다. 즉, 차를 운행하는 사람에게 책임을 지운다는 것이다.

최○○ 교수는 "현재로서는 자율 주행 자동차 사고로 인한 손해 배상은 자동차 보험법을 빌려 처리할 수 있다. 해당 법은 애초에 운행자에 책임을 묻기 위해 마련된 것으로 제조사에 책임을 지울 수 없다. 물리적 운행을 직접 하지 않더라도 운행을 하게 만드는 사람이 책임을 질 수 있는 것"이라고 말했다. 이와 같은 상황에 대해 최 교수는 운전자 책임을 완화하는 방향으로 법을 개선해야 한다는 의견이다. 또한 최○○ 교수는 "AI를 사용하는 이유는 편리하기 위해서인데 결국 책임을 사용자가 다지면 무슨 의미가 있을까? 제품 이용을 많이 하도록 만들려면 책임 완화 장치가 필요하다. 그래서 AI 제품을 더 쓰고 싶어질 것이다. 제조물 책임을 아주 강화하거나 일단 서비스 제공자에게 책임을 물리는 것도 가능하다."라고 설명했다.



자율 주행 자동차 사고 원인을 제대로 판단하기 위해서는 AI 기술을 잘 아는 외부 전문가 위원회도 필요하다는 주장이다. 최 교수는 "의료사고 발생 시에도 외부 의료 전문가들이 객관적으로 판단해 분쟁을 조정한다. 자율 주행 자동차 사고에서도 알고리즘, 데이터, 센서, 네트워크 등 어디서 문제가 발생했는지 파악하려면 전문가가 필요하다. 의료 조사 위원회처럼 독자적으로 조사하는 조직을 만들어 조사하면 서비스 제공자 책임이 강화되더라도 책임 균형이 이뤄질 것"이라고 말했다. 알고리즘과 같이 물리적으로 특정할 수 없는 것이 피해 원인이라면 현행법으로는 처벌하기 어렵다. 이를 위해 디지털 제품에 대한 제조물 책임법을 새로 마련해야 한다는 것이 최 교수의 입장이다.

[출처 : <http://www.aitimes.com/news/articleView.html?idxno=141902>]

1 기사에서 인공지능 윤리 문제에 대한 해결 방안을 어떻게 제시하였나요?

2 기사에서 제시된 해결 방안에 대한 자신의 인공지능 윤리 원칙을 만들어 봅시다.

| 문제 | | |
|-----------|-----|--|
| 필요한 규정 | 안전성 | |
| | 책임성 | |
| | 투명성 | |

짝 또는 모둠별로 인공지능의 설명 가능성에 대하여 사용자와 개발자 입장에서 의견을 나눠 봅시다.



토론

인공지능의 설명 가능성을 강화해야 할 것인가?

인공지능의 투명성을 확보하는 방안의 하나로 설명 가능성에 대해서 논의가 많이 일어나고 있습니다. 인공지능 시스템의 의사결정이나 예측 결과의 판단 근거를 우리가 이해할 수 있는 표현 방식으로 나타냄으로써 신뢰할 수 있는 인공지능 그리고 투명한 인공지능이 될 수 있다고 보는 것입니다.

인공지능이 어떤 과정을 거쳐서 결론에 도달했는지 알 수 없다면 인공지능의 결정을 신뢰하기 어려울 것입니다. 하지만 설명 가능한 인공지능이 내주는 결과 때문에 우리 인간은 선불리 반대 의견을 내기가 쉽지 않습니다. 어떤 과정을 통해 이렇게 판단했다고 할 때 우리는 인공지능의 판단을 거스르기가 쉽지 않기 때문입니다.

또한 개발자 입장에서 인공지능 알고리즘은 사업의 성패의 핵심 기술이며 사업 기밀이기 때문에 지식 재산권이라고 할 수 있습니다. 따라서 큰 비용을 들여 개발한 알고리즘을 모두 공개하는 것은 지식 재산권을 침해하는 것이라고 주장할 수 있습니다.

그렇다면 설명 가능한 인공지능의 개발을 통해 인공지능의 결정 과정의 이해도를 높이는 것이 필요할까요? 아니면 어느 정도 설명 가능성을 희생하면서 우리의 판단을 강화하도록 하는 것이 좋을까요?

토론 방법

- ① 사용자와 개발자의 두 입장에 대한 자신의 주장을 적습니다.
- ② 인터넷에서 자신의 주장을 뒷받침하는 근거 자료를 수집합니다.
- ③ 한 명은 사용자, 다른 한 명은 개발자 입장이 되어 토론을 합니다.
- ④ 서로의 역할을 바꾸어 토론을 합니다.
- ⑤ 서로의 입장을 고려하여 최종 생각을 정리합니다.

1 서로의 입장 정리하기

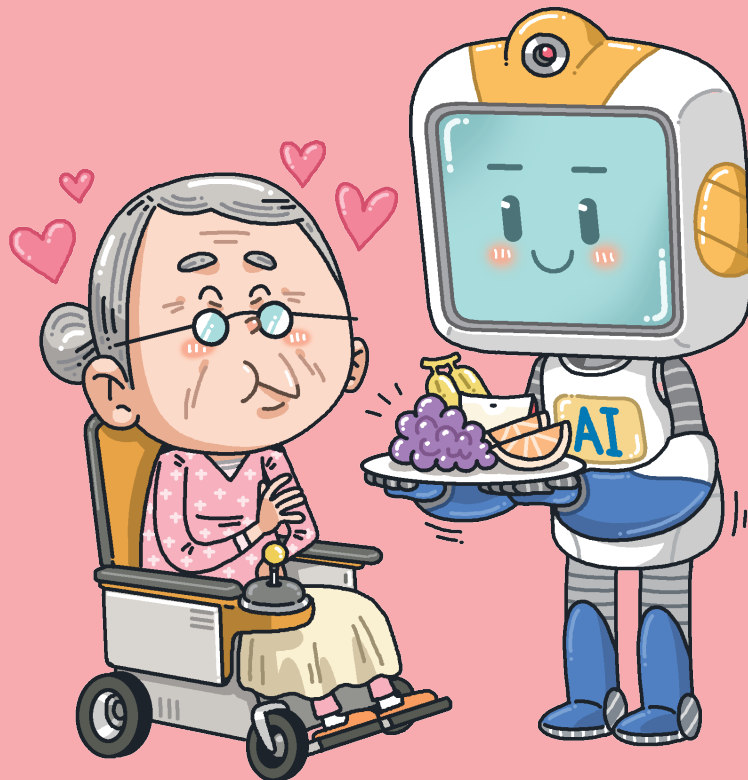
| 사용자 입장 | 개발자 입장 |
|-------------|-------------|
| 주장1: 근거: | 주장1: 근거: |
| 주장2: 근거: | 주장2: 근거: |

2 토론 후 나의 결론을 정리해 봅시다.

5

인공지능은 모든 사람들을 배려하고 있을까?

- 1 인공지능 활용과 사용자 유형 이해하기
- 2 공공성과 연대성 이해를 위한 재활용품 분리수거 인공지능 프로그램 만들기
- 3 개발자와 사용자는 어떤 생각을 할까?
- 4 공공성과 연대성 실천을 위한 활동



인권 보장

프라이버시 보호

다양성 존중

침해 금지

공공성

연대성

데이터 관리

책임성

안전성

투명성



학습 전 이야기 나누기

이것은 대화형 인공지능으로 노인과의 대화를 기억하고 안부를 묻고, 낙상을 감지하고 위기 상황이 발생하면 119에 신고도 척척 해 줍니다. 노인과의 과거 대화 내용을 바탕으로 통화 중 어르신 건강 상태, 식사 습관, 수면 시간 등의 주요 정보를 요약해 기록하고, 통화가 끝나면 답변 내용을 기반으로 노인의 상태를 업데이트 합니다. AI 기반 센싱 정보 분석 기능을 통해 서거나 앉아 있거나 누워 있는 자세와 낙상 여부를 감지합니다. 또 이상 징후가 발견된 노인들의 동선을 최대 일주일 간 추적하여 위기 상황 시 119에 자동으로 신고하고 관리자에게 문자 메시지를 전송합니다. 이것은 무엇일까요?

정답 인공지능 돌봄 서비스, AI 케어 서비스

1

인공지능 활용과 사용자 유형 이해하기



학습
목표

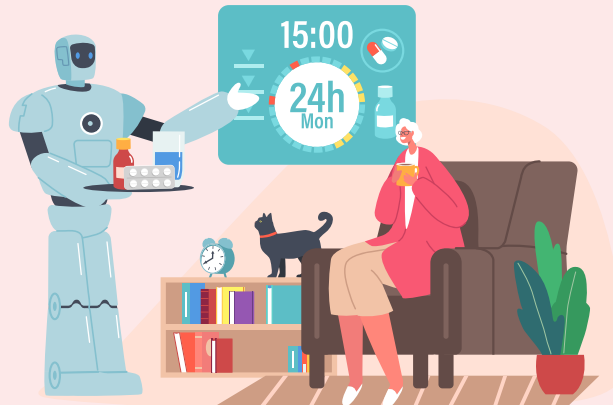
- 인공지능 윤리의 공공성과 연대성의 개념을 이해하고 설명할 수 있다.
- 인공지능 발달로 인한 실생활의 변화를 인식하고, 사용자 유형을 분석할 수 있다.
- 인공지능 도입으로 달성 가능한 지속가능발전목표를 이해하고 이를 위한 방안을 설명할 수 있다.

AI 이야기 어느덧 일상으로 다가온 AI, 인공지능과 함께 살아가는 시대가 왔다고?

우리 집 거실에 앉아 있다가 TV 광고 속 처음 본 여자 연예인이 궁금해서 인터넷에 찾아보니 비주얼 인플루언서 ‘벨포레’라는 것을 알게 되었습니다. 이제는 그녀가 인공지능으로 만들어진 사람이라는 것을 대부분 알고 있습니다. 그녀는 가상 공간 속에서 SNS를 하고, 메타버스 세계의 인간과 공존하여 대화를 나누면서 소통을 하고 있습니다.

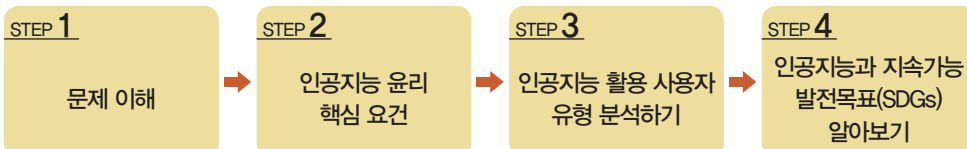
우리 주변에서 볼 수 있는 서빙 로봇, 24시간 예약을 받아 주는 AI 통화 비서, 독거노인과 말동무가 되어 주는 AI 케어 로봇이 그 주인공이라고 할 수 있습니다.

인공지능 기술은 미래 사회에 예측하기 어려운 문제를 해결해 주고 인간이 하기 어려운 일을 대신해 주기도 합니다. 주변을 감시해 범죄가 발생하면 112에 신고해 주고, 화재가 발생하거나 일정 시간 동안 사람의 움직임이 없으면 119에 신고해 주는 시스템과 쓰레기를 무단으로 투기하는 것을 감지하는 인공지능 시스템이 개발되고 있습니다. 과연 앞으로는 어떤 일들이 발생할까요?



- 미래 사회 발전을 위한 인공지능 기술의 도입으로 혜택을 받고 있는 사람은 누구일까요? 소외되는 사람이 내 주변인이자면 난 무엇을 해야 할까요?
- 인공지능의 도입과 발전으로 지구와 우리가 사는 세상에는 어떤 변화가 생길까요?

학습 과정



이번 시간에 우리가 해결해야 할 문제와 문제 해결 과정을 알아보시다.



생활 속 문제 발견

인공지능 프로그램은 우리 사회의 많은 부분을 편리하게 만들어 주고 있습니다. 하지만, AI 자체가 환경 오염을 발생시키는 원인이 되기도 합니다. 특히 딥러닝 기술 개발 과정에서 막대한 전력을 소모하고 이산화탄소를 배출하는데, 이 배출량이 자동차 5대가 평생 배출하는 양과 같다는 조사 결과도 있습니다.

한 반도체 회사의 회장은 “AI 기술 발전이 지구에 재앙이 될 수 있다.”라는 의견을 밝혔습니다. 현재 기술 수준으로 인공지능 데이터 센터를 계속 구축한다면 2025년까지 전 세계 전력의 15%를 데이터 센터가 소비하게 된다는 것입니다. 미래 세대가 안전한 지구에서 살아갈 수 있도록 인공지능 활용 과정에서 전력을 줄이는 노력을 하면서도 AI를 더 적극적인 ‘환경 지킴이’ 역할을 하게 하려면 무엇을 할 수 있을까요? 또 사람이 하던 재활용 분리수거를 AI가 도와줄 수 있을까요?



우리가 해결해야 할 문제가 무엇인지 적어 봅시다.



위 문제를 해결하기 위해 생각한 방법은 무엇인지 적어 봅시다.

인공지능 기술 발전에 따른 변화의 배경과 우리가 배워야 할 인공지능 윤리 핵심 요건을 알아봅시다.

1. 배경

인공지능 기술을 활용한 다양한 서비스가 제공되고, 인공지능 기반 공공 서비스가 확대됨에 따라 사회 전반의 변화가 일어나고 있습니다. 인공지능 활용에 있어서 공공성과 연대성의 가치를 실현하기 위해서는 미래 세대와 취약 계층을 포함한 사회 구성원의 이익과 협력을 이끌어내는 것이 필요합니다.

2. 핵심 요건 알아보기

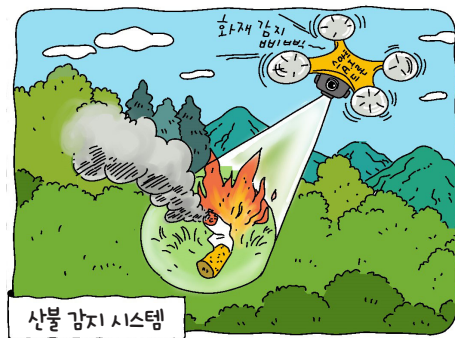
공공성 및 연대성에 대해 알아봅시다.

1 공공성

우리가 지향하는 미래 사회는 최대한 많은 사람의 행복을 보장하고 더불어 사는 것입니다. 공공성이란 이 사회에서 인공지능 기술의 개발과 활용을 통해 개인의 행복 추구 뿐만 아니라 사회적 공공성과 인류의 공동 이익을 함께 추구해야 한다는 윤리입니다.

사회 구성원 공동의 이익을 증진시키기 위해서는 인공지능 기술 역시 건전한 사회 질서를 해하지 않는 범위 내에서 활용되어야 합니다. 사회적 공공성과 인류의 공동 이익을 추구하기 위해서는 관련 기업의 사회적 책임에 대해 사회적 합의를 이끌어 내고 정책적·제도적 차원의 지원도 논의가 필요합니다.

인공지능 시스템이 가져 올 혜택이 특정 개인이나 집단의 이익을 위해서만 활용되는 것이 아니라 다양한 집단이 누릴 수 있어야 한다는 공공성을 위해 개발 및 운영 단계에서 고려할 점이 많습니다.



관련 사례

기사

인공지능 소방관 개발... 화재 감지·식별에 진화까지?



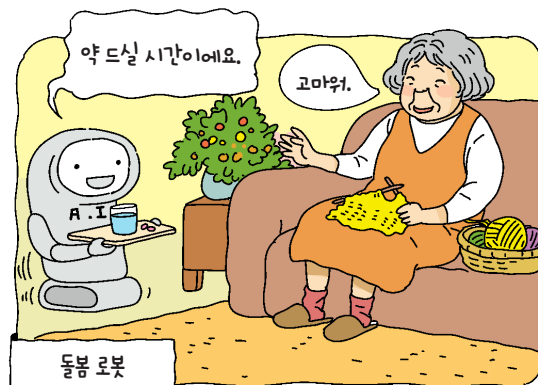
공간 전체를 감지하는 CCTV에서 불을 발견하면 적외선 카메라로 주변 온도를 측정해서 진짜 불로 판정될 때 화재 원점의 위치를 파악한 후, 소화기가 가동돼 그 주위에 물이나 소화제를 뿌려 진압합니다. 또한 A사는 사스(SaaS, 서비스형 소프트웨어) 기반 산불 조기 감지 시스템 ‘파이어 스카우트’를 세계 최초로 상용화했습니다. 이는 실시간으로 산을 비추는 카메라 영상 속에서 AI가 연기 등 이상 상황을 감지하는 시스템으로 연기 포착 즉시 해당 지역을 표시해 감시원들에게 알립니다. 그 결과 10월 미국 캘리포니아 타말파이어스 산(Mt. Tamalpais)에서 발생한 화재에서 911 신고보다 7분 빠르게 화재를 감지하는 등 조기 진화에 기여하고 있어 산불 감지 시스템으로 도입되고 있다고 합니다.

[출처: 매일경제 <https://www.mk.co.kr/news/business/10155872>]

2 연대성

연대성은 인공지능 기술을 개발하고 활용할 때 다양한 집단의 관계 연대성과 미래 세대를 충분히 배려해야 한다는 인공지능 핵심 요건 중 하나입니다. 연대성의 가치는 우리가 지녀야 할 책임 윤리를 미래 세대 및 전 지구적 차원으로 확대할 필요성을 일깨워 줍니다. 인공지능의 영향이 현 세대에게만 미치는 것이 아니며 미래 세대의 삶에도 전달될 수 있기에 그 기술의 개발과 운영 과정에서 사전에 미래 세대를 위한 충분한 배려를 할 필요가 있습니다. 따라서 인공지능 기술의 개발 및 활용에 있어 국내적 차원을 넘어 국제적 차원에서도 다양한 이해관계자 간의 협력과 참여를 강화함으로써 사람 중심의 윤리적 인공지능의 구현을 위해 노력해야 합니다.

인공지능은 서로 다른 환경과 상황에 따라 다양한 이해관계자를 가진 사회 구성원이 활용하므로 사회 통합을 이루는 방향으로 이루어져야 합니다. 이를 위해서는 인공지능 시스템을 사용하는 구성원 간 소통의 기회를 보장하여 신뢰를 형성해 나가는 것이 중요합니다.



관련 사례

기사

“AI가 효자”... 평소엔 말벗, 위기 땀 구조 호출

아이들이 성장해서 독립하고 나니 집안에 늘 부부 둘 뿐인데 대화를 나누기도 하고, 취미인 운동도 같이 하니 그나마 외롭지 않습니다. 하지만 혼자 사는 노인이라면 그 외로움과 적적함은 이루 말할 수 없습니다. 65세 이상 홀로 거주하는 노인이 150만 명을 넘기고, ‘고독사’로 불리는 무연고 사망도 늘어나는 요즘, AI 기술을 접목한 스마트 돌봄 서비스가 있어 눈길을 끄니다. 인공지능 ‘효돌은’ 생체 감지 센서가 있어 실시간으로 어르신의 움직임과 호흡을 감지합니다. 만약, 2시간 동안 움직임이 감지되지 않으면, 노인의 현재 상황 체크가 필요하다는 붉은색 표시가 관제 센터 모니터에 자동으로 뜨고 담당자가 즉시 전화나 직접 방문하여 현재 상태를 확인합니다.



[출처: 대한민국 정부 <https://naver.me/GRzYxfP2>]

활동 1-1 인공지능 공공 서비스에 대한 문제점 알아보기

일상생활 속에서 인공지능 공공 서비스 시스템의 사용이 지역, 성별, 세대, 계층 간 갈등을 유발하는 등 사회 통합을 저해할 가능성이 있는지 알아보기 위해 사용자 유형을 분석해 봅시다.

새롭게 대두되는 인공지능 도입에 따른 인공지능 서비스와 분야

일상생활 속에서 인공지능 서비스는 점점 많이 활용되고 있으며 서비스 활용 계층이 다양해지고 있습니다.

인공지능 서비스의 취약 계층이 겪는 어려움은 단순히 기기에 친숙하지 않다는 문제에 국한된 것이 아닙니다. 이것은 근본적으로 지역, 성별, 세대, 계층에 따라 혜택을 받지 못하는 현상이 발생하는 문제와도 관련이 있습니다.

가령 새로운 인공지능 기술이 의료 환경에서도 개발되고 있지만 어떤 이유로 인해 보건 의료 서비스를 이용하기 어려워 혜택을 누리지 못하는 의료 취약 계층 또는 의료 소외 계층이 생겨나 의료계 격차가 커지고 있습니다.

기존에 알려진 농촌 및 지방 거주자, 저소득층, 노숙인, 외국인 근로자 등이 여기에 해당되고, 이들은 최근 인공지능 서비스 취약 계층으로 새롭게 부상되고 있습니다. 이는 질환의 예방과 치료에 인공지능을 활용한 기술을 이용할 때 생기는 이점이 보고되고, 또 인공지능 헬스 케어에 대한 관심이 증가하면서 생기는 또 다른 격차로 보입니다.

인공지능 시스템의 활성화뿐만 아니라 개인용 컴퓨터나 스마트폰의 일반화로 사물인터넷 기반 인공지능 공공 서비스 접근성은 많이 개선되었지만, 아직까지 인공지능 제품·서비스 정보 활용 능력이나 실제 활용도에 따른 격차가 발생하고 있습니다.



1 인공지능 공공 서비스 시스템을 도입할 때 고려해야 할 사용자 유형을 분석해 봅시다.

각 사용자 유형으로 혜택을 받을 수 있는 인공지능 서비스를 선으로 잇고, 그 인공지능 서비스와 해당하는 분야를 선으로 이어 봅시다. 선으로 이은 결과를 짝과 함께 이야기해 봅시다.

| 사용자 유형 | | 인공지능 서비스 | | 분야 |
|---------|---|-----------------|---|--------|
| 고령층(노인) | • | 책 읽어 주는 램프 | • | 교육 |
| 농어촌민 | • | 서빙 로봇 | • | 공공 안전 |
| 시각 장애인 | • | 로봇 택배 | • | 의료 |
| 청각 장애인 | • | 케어 로봇 | • | 헬스케어 |
| 신체 장애인 | • | 음성 인식 TV 제어 시스템 | • | 생활 |
| 성인 | • | 장애인용 지팡이 | • | 보안 |
| 청소년 남자 | • | 장애물 감지 신발 | • | 서비스 로봇 |
| 청소년 여자 | • | 얼굴 인식 출입증 | • | 교통 |
| 초등학생 | • | 자율 주행 자동차 | • | 고용/직장 |
| 어린이 | • | 인공지능 환경 안전 CCTV | • | |
| 외국인 | • | 챗봇/음성봇 | • | |
| 저소득층 | • | 영화 추천 시스템 | • | |
| 중산층 | • | 인공지능 튜터 (영어 회화) | • | |
| 고소득층 | • | | | |

2 인공지능 공공 서비스 시스템 도입하려고 할 때 가장 많은 혜택을 받을 것 같은 유형과 그렇지 못할 것 같은 사용자 유형을 <보기>에서 고른 후, 물음에 답해 봅시다.

| | | | | | |
|------|---------|--------|--------|--------|------|
| <보기> | 고령층(노인) | 시각 장애인 | 일반 성인 | 어린이 | 저소득층 |
| | 도시민 | 청각 장애인 | 청소년 남자 | 외국인 | 중산층 |
| | 농어촌민 | 신체 장애인 | 청소년 여자 | 다문화 가족 | 고소득층 |

• 인공지능 공공 서비스를 가장 잘 활용할 것 같은 유형 또는 사람과 그 이유를 적어 보세요.

| 사용자 유형 | 그 이유는? |
|--------|--------|
| | |
| | |

• 인공지능 공공 서비스에 대한 혜택을 잘 받지 못할 것 같은 유형 또는 사람과 그 이유를 적어 보세요.

| 사용자 유형 | 그 이유는? |
|--------|--------|
| | |
| | |

3 인공지능 기술을 공공 서비스에 도입할 때 특정 사용자 유형이 혜택을 받지 못한다면 어떻게 해결할지 생각해 보고 해결 방안을 작성해 봅시다.

STEP 4 연대성 - 지속가능발전목표(SDGs) 알아보기



활동 1-2 인공지능과 지속가능발전목표(SDGs) 알아보기

지속가능발전목표(SDGs)를 위한 국제 사회의 협력을 위해 지구촌의 주요 목표를 조사 및 탐색하고 우리가 살아갈 미래를 생각하며 다음을 작성해 봅시다.

1 지속가능발전목표(SDGs) 17가지를 검색하여 찾아보고, 다음 빈칸을 채워 봅시다.

국가 지속가능발전목표

KOREAN SUSTAINABLE DEVELOPMENT GOALS (K-SDGs)

| | | | | | |
|---|---|---|---|-------------------------|---|
| 1 빈곤층 감소와 사회안전망 강화 | 2 식량안보 및 지속가능한 농업 강화 | <div style="border: 1px solid black; width: 80px; height: 80px;"></div> | 4 모두를 위한 양질의 교육 | 5 성평등 보장 | <div style="border: 1px solid black; width: 80px; height: 80px;"></div> |
| <div style="border: 1px solid black; width: 80px; height: 80px;"></div> | 8 좋은 일자리 확대와 경제성장 | 9 산업의 성장과 혁신 활성화 및 사회기반시설 구축 | <div style="border: 1px solid black; width: 80px; height: 80px;"></div> | 11 지속가능한 도시와 주거지 조성 | 12 지속가능한 생산과 소비 |
| 13 기후변화와 대응 | <div style="border: 1px solid black; width: 80px; height: 80px;"></div> | 15 육상생태계 보전 | 16 평화·정의·포용 | 17 지구촌 협력 강화 | <p>국가지속가능발전목표 K-SDGs</p> |

출처: <http://ncsd.go.kr/ksdgs?content=4>

| 3 | 6 | 7 | 10 | 14 |
|---|---|---|----|----|
| | | | | |



2 다음 그림에서 보이는 단어 5개를 골라 적고, 클라우드 빙고 게임을 해 봅시다.



지속가능 발전 목표를 위한 사회 문제는?

1인 가족의 건강 안전 문제
신체장애인의 이동권
가짜뉴스
청소녀 우울
해당오염
장애인
지구온난화
금연구역
경제적 독립
윤석열
학교
지구 환경문제
제로웨이스트
놀이터 낯선사람 구분하기
차별
관계자외출금지
자살
기후변화
학교 폭력
교통안전
에너지절약
비속어 사용
노인복지
학생들의 비속어 사용
쓰레기 문제
장애인의 문제
사회성 상실
쓰레기 무단투기
노인복지
공공성의
환경오염 문제
대기오염
디지털 소외
다문화
환경문제
미세플라스틱 문제
개인의주의
지역불균형
장애인 차별
자원문제
디지털 격차
물 부족 문제
사회적 격차

| | | | | |
|--|--|--|--|--|
| | | | | |
|--|--|--|--|--|

3 인공지능 기술이 아래의 문제를 해결하는 데 도움을 준다고 생각하면 ○ 표시를 해 봅시다. 그리고 그 이유를 제시해 봅시다.

| 탄소 중립 전환 | | |
|-----------------|--------------|-----------|
| 탄소 중립 실천 포인트 지급 | 산업계 탄소 중립 지원 | 1회용 컵 보증금 |

| 환경 분야 탄소 감축·흡수 | | |
|----------------|---------|--------|
| 재생 에너지 | 무공해차 보급 | 탄소 흡수원 |

| 통합 물관리 | | |
|-----------|--------|-----------|
| 옥내 급수관 개량 | 新하천 관리 | 하천 자연성 회복 |

| 환경 위해로부터 국민 안전 보호 | | |
|-------------------|-------------|--------------------|
| 미세먼지 감축 | 충간 소음 기준 강화 | 생활 화학 제품 전성분 정보 공개 |

| 폭넓은 환경 서비스 확대 | | |
|------------------|---------|-------|
| 환경 오염 취약 지역 모니터링 | 생태계 서비스 | 동물 복지 |

4 ○ 표시를 한 실천 방안들을 종합하여 인공지능으로 해결할 방안을 생각하여 작성해 봅시다.



정리하기



❖ 다음의 문항을 읽고 ○ 안에 들어갈 말을 적어 봅시다.

1. 인공지능 윤리 핵심 요건 중 ○○○이란 이 사회에서 인공지능 기술의 개발과 활용을 통해 사익만을 추구하는 것이 아니라 공익을 동시에 추구해야 한다는 윤리이다. ()
2. 인공지능 윤리 핵심 요건 중 ○○○이란 인공지능 기술을 개발하고 활용할 때 다양한 집단의 관계 연대성과 미래 세대를 충분히 배려한다는 윤리이다. ()
3. ○○○○ 기술 서비스에는 책 읽어 주는 램프, 서빙 로봇, 로봇 택배, 케어 로봇, 음성 인식 TV 제어 시스템, 장애인용 지팡이, 장애물 감지 신발, 얼굴 인식 출입증, 자율 주행 자동차 등이 있다. ()
4. ○○○○○○○○는 '단 한 사람도 소외되지 않는 것(Leave no one behind)'이라는 슬로건과 함께 인간, 지구, 번영, 평화, 파트너십이라는 5개 영역에서 인류가 나아가야 할 방향성을 17개 목표와 169개 세부 목표로 제시하고 있다. ()



점검하기

| | 예 | 아니요 |
|--|--------------------------|--------------------------|
| 인공지능의 공공성과 연대성에 대해 설명할 수 있나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능의 사용자 유형을 분류하고 인공지능의 공공 서비스의 혜택을 받는 사람과 받지 못하는 사람을 알고 그 이유를 설명할 수 있나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능과 지속가능발전목표(SDGs)를 검색해 보고 사회 문제를 발견할 수 있나요? | <input type="checkbox"/> | <input type="checkbox"/> |

2

공공성과 연대성 이해를 위한

재활용품 분리수거 인공지능 프로그램 만들기



학습
목표

- 인공지능 공공 서비스 시스템 개발을 위한 지구촌 환경 문제 프로젝트를 계획할 수 있다.
- 인공지능 공공 서비스 시스템 개발을 위한 다양한 모델링을 할 수 있다.
- 인공지능 윤리 의식을 함양하기 위해 모든 사람이 배려하여 편리하게 사용할 수 있는 인공지능 프로그램을 개발할 수 있다.

AI 이야기 쓰레기 58가지를 분류 선별하는 인공지능 로봇

‘닥터 비(B)’는 재활용품의 이미지 데이터를 입력받아 특성을 인공지능 기술로 학습한 후 종이, 비닐, 알루미늄 등 58가지 항목을 구분하는 인공지능 선별 로봇 중 하나입니다.

인공지능 선별 로봇이 재활용품 터미에서 내부에 탑재된 센서로 목표하는 재활용품을 포착한 후 팔을 뻗어 빨아들이고, 이를 컨베이어 벨트 밖으로 빼 내면서 선별합니다. 인공지능 선별 로봇은 재활용품을 선별할수록 그 과정을 통해 학습하여 능력이 점점 진화되고 있다고 합니다.

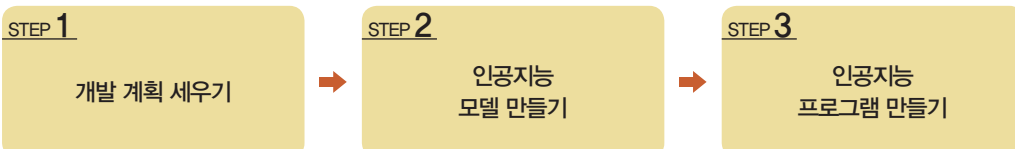
인공지능 로봇이 재활용품을 종류별로 구분한다는 것이 믿기지는 않지만 이 기기는 하루 평균 55톤의 재활용품을 처리합니다. 먼저 사람이 직접 플라스틱, 유리, 알루미늄, 고철 등과 같이 부피가 큰 쓰레기를 먼저 골라 내고 이후부터는 인공지능 기술이 투입되고 있습니다.

[출처: 한겨레 신문 2022.02.03. 요약 발췌]



- 공공성을 보장하며 지구촌 환경 문제를 해결하는 인공지능 프로그램을 제작할 수 있을까요?
- 우리가 인공지능 윤리 의식을 갖추기 위해 고려해야 할 것은 무엇일까요?

학습 과정



STEP 1 개발 계획 세우기



인공지능 기술의 발전이 가져 온 생활 속 변화를 살펴보고, 인공지능 윤리의 공공성을 위한 다양한 인공지능 공공 서비스 프로그램 개발을 위한 계획을 세워 봅시다.

활동 2-1 사회적 문제 해결 프로그램 개발 계획 세우기

활동 1-2 활동을 보고 다음 활동을 통해, 지구촌의 주요 환경 문제를 조사하여 인공지능 서비스 프로그램을 만들어 봅시다.

- 1 지속가능발전목표 중에서 우리가 해결하고 싶은 주제를 정하고, 인공지능을 활용하여 문제를 해결하고자 할 때 생길 수 있는 윤리적 이슈를 생각해 봅시다.

| 주제 | 윤리적 이슈 |
|----|--------|
| | |

- 2 재활용품 분리수거 인공지능 프로그램을 만들기 위한 방안을 탐색하고 정리해 봅시다.

| 방안 | 공공성 · 연대성을 증진할 수 있는 요소 |
|----|------------------------|
| | |

- 3 재활용품 분리수거 인공지능 프로그램을 만들기 위해 필요한 데이터가 무엇인지 적어 보고 데이터를 수집해 봅시다.

| 필요한 데이터 |
|---------|
| |

- 4 재활용품 분리수거 인공지능 프로그램을 만들기 위한 알고리즘을 글과 그림으로 표현해 봅시다.

| 글로 표현하기 | 그림으로 표현하기 |
|---------|-----------|
| | |

- 5 지속가능발전목표를 위한 프로그램을 활용할 때 고려해야 할 윤리적 문제는 무엇인지 생각해 봅시다.

지속가능발전목표(SDGs)의 사회 문제 해결 프로그램에 필요한 인공지능 모델을 생성하여 학습시키고, 다양한 재활용품 테스트 이미지를 이용하여 학습이 잘되었는지 확인해 봅시다.

개발자



































활동 2-2 '재활용품 분리수거' 인공지능 모델 만들기

다음의 설명에 따라 인공지능 모델을 생성하여 학습시키고, 데이터를 이용하여 학습이 잘되었는지 확인해 봅시다.

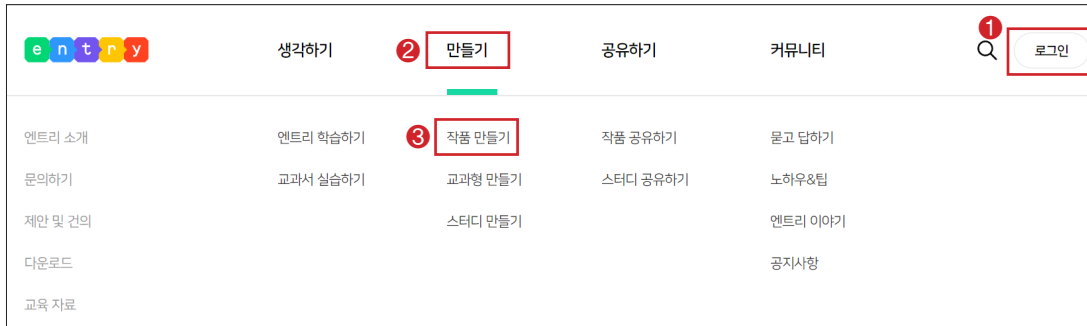
인공지능 모델 생성 및 학습

- ① 분류할 모델의 이름을 설정합니다.
- ② 클래스명을 사물의 이름으로 설정합니다.
- ③ 학습 데이터를 10개 이상 수집하여 클래스별로 업로드합니다.
- ④ 인공지능 모델을 학습시킵니다.
- ⑤ 인공지능 모델의 학습이 잘 되었는지 테스트 데이터로 결과를 확인합니다.

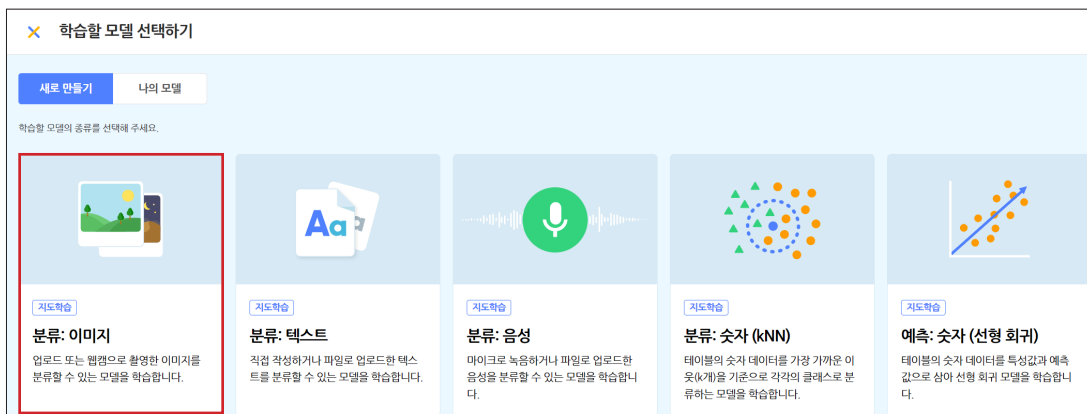
1 인공지능 모델을 생성하기 위해 필요한 데이터를 수집합니다.

| 학습 데이터 | | | | | | | |
|---|---|---|---|---|--|---|---|
| 병 | | 캔 | | 플라스틱 | | 일반 쓰레기 | |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

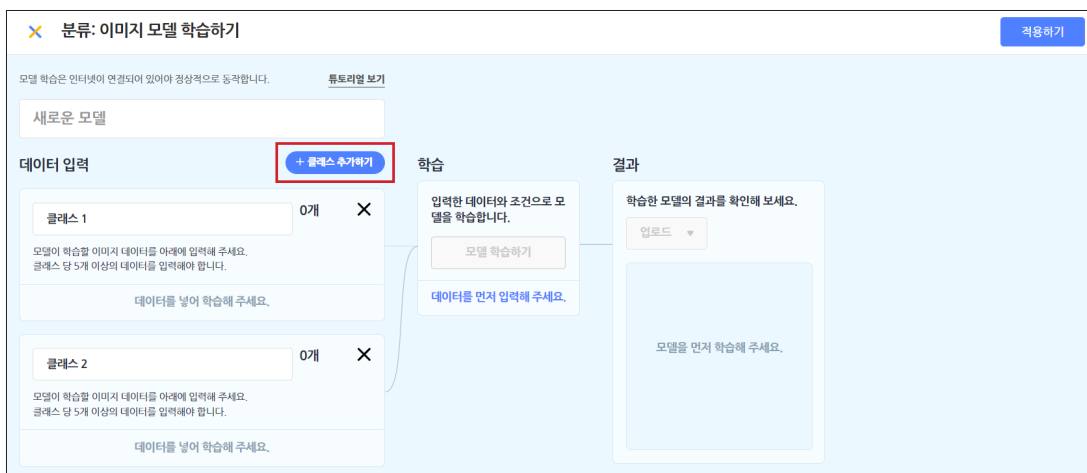
2 엔트리 사이트(<https://playentry.org/>)에 접속하여 로그인한 후 [만들기] - [작품 만들기]를 클릭합니다.



3 블록 카테고리에서 [인공지능] - [인공지능 모델 학습하기] - [지도학습 분류: 이미지]를 차례대로 클릭합니다.



4 [클래스 추가하기] 버튼을 두 번 클릭해서 클래스 2개를 더 추가합니다.



- 5 모델 이름에 '재활용품'이라고 입력하고, 클래스 1을 '병', 클래스 2를 '캔', 클래스 3을 '플라스틱', 클래스 4를 '일반쓰레기'로 입력합니다. 이 클래스 이름이 정답에 해당하는 레이블입니다.

분류: 이미지 모델 학습하기

적용하기

모델 학습은 인터넷이 연결되어 있어야 정상적으로 동작합니다.

튜토리얼 보기

재활용품

데이터 입력

클래스 추가하기

병

0개

×

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.

데이터를 넣어 학습해 주세요.

캔

0개

×

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.

데이터를 넣어 학습해 주세요.

플라스틱

0개

×

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.

데이터를 넣어 학습해 주세요.

일반쓰레기

0개

×

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.

데이터를 넣어 학습해 주세요.

학습

입력한 데이터와 조건으로 모델을 학습합니다.

모델 학습하기

데이터를 먼저 입력해 주세요.

결과

학습한 모델의 결과를 확인해 보세요.

업로드

모델을 먼저 학습해 주세요.

- 6 앞에서 준비한 이미지 학습 데이터를 이용하여 클래스별로 데이터를 입력합니다.

분류: 이미지 모델 학습하기

적용하기

모델 학습은 인터넷이 연결되어 있어야 정상적으로 동작합니다.

튜토리얼 보기

재활용품

데이터 입력


클래스 추가하기

병

10개

×

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.




+5

캔

10개

×

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.




+5

플라스틱

10개

×

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.




+5

일반쓰레기

10개

×

모델이 학습할 이미지 데이터를 아래에 입력해 주세요.
클래스 당 5개 이상의 데이터를 입력해야 합니다.



+5

학습

입력한 데이터와 조건으로 모델을 학습합니다.

모델 학습하기

모델을 학습할 수 있습니다.

결과

학습한 모델의 결과를 확인해 보세요.






업로드

모델을 먼저 학습해 주세요.

8 학습이 완료된 모델의 성능을 평가하기 위해 다음의 테스트 데이터로 모델의 결과를 확인하고 분류 결과를 적어 봅시다.

| 테스트 데이터 1 | 테스트 데이터 2 | 테스트 데이터 3 | 테스트 데이터 4 | 테스트 데이터 5 |
|---|---|---|--|---|
|  |  |  |  |  |

• 분류 결과를 확인해 보세요.

| 테스트 데이터 1 | 테스트 데이터 2 | 테스트 데이터 3 | 테스트 데이터 4 | 테스트 데이터 5 |
|--|--|--|---|--|
| <p>결과</p> <p>학습한 모델의 결과를 확인해 보세요.</p> <p>업로드</p>  <p>병 98.54%</p> <p>캔 0.01%</p> <p>플라스틱 1.33%</p> <p>알루미늄재기 0.09%</p> | <p>결과</p> <p>학습한 모델의 결과를 확인해 보세요.</p> <p>업로드</p>  <p>병 0.84%</p> <p>캔 97.97%</p> <p>플라스틱 1.01%</p> <p>알루미늄재기 0.15%</p> | <p>결과</p> <p>학습한 모델의 결과를 확인해 보세요.</p> <p>업로드</p>  <p>병 0.11%</p> <p>캔 72.3%</p> <p>플라스틱 1.36%</p> <p>알루미늄재기 26.21%</p> | <p>결과</p> <p>학습한 모델의 결과를 확인해 보세요.</p> <p>업로드</p>  <p>병 0.03%</p> <p>캔 0.02%</p> <p>플라스틱 93.97%</p> <p>알루미늄재기 5.96%</p> | <p>결과</p> <p>학습한 모델의 결과를 확인해 보세요.</p> <p>업로드</p>  <p>병 0%</p> <p>캔 0.08%</p> <p>플라스틱 0%</p> <p>알루미늄재기 99.91%</p> |

9 테스트 데이터의 결과를 확인해 보고, 분류 결과가 잘못 나왔다면 적절한 학습 데이터를 추가하여 모델을 다시 만들고 결과를 확인해 봅시다.

10 모델이 완성되었다면 오른쪽 상단의 [적용하기](#) 버튼을 클릭하여 모델 생성을 완료합니다.

STEP 3 인공지능 프로그램 만들기



인공지능 모델을 만들었다면 만든 모델을 활용하여 지속가능발전목표(SDGs)의 문제 해결 프로그램을 만들어 봅시다.

개발자

활동 2-3 프로그래밍하기

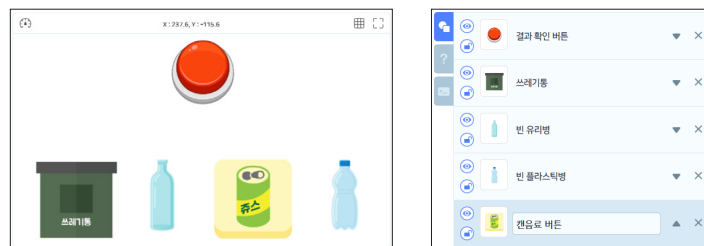
지속가능발전목표(SDGs)를 위해 재활용품 분리수거 인공지능 프로그램을 구현해 봅시다.

알고리즘 확인하기

- ① 쓰레기를 투입합니다.
- ② 재활용품과 일반 쓰레기를 분류합니다.
- ③ 만약 재활용품이라면
 - 만약 병이라면 탄소 포인트 10을 적립하고, 병의 갯수에 1을 더합니다.
 - 만약 캔이라면 탄소 포인트 20을 적립하고, 캔의 갯수에 1을 더합니다.
 - 만약 플라스틱이라면 탄소 포인트 30을 적립하고, 플라스틱의 갯수에 1을 더합니다.
- ④ 그렇지 않으면 일반 쓰레기로 분리합니다.
- ⑤ 분리수거가 종료되어 버튼을 누르면 최종 탄소 적립 포인트를 알려 줍니다.

1 필요한 오브젝트, 변수, 신호를 추가합니다.

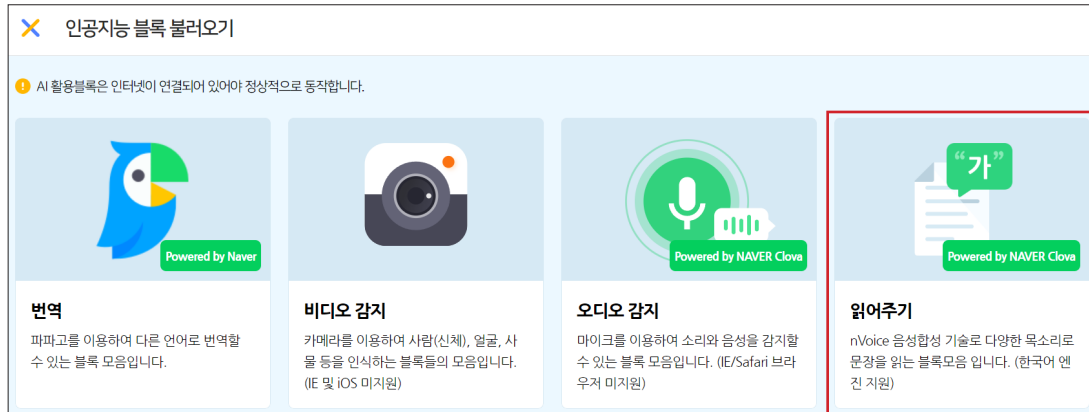
- ① 화면에 버튼, 쓰레기통, 병, 플라스틱, 캔 오브젝트를 추가합니다.



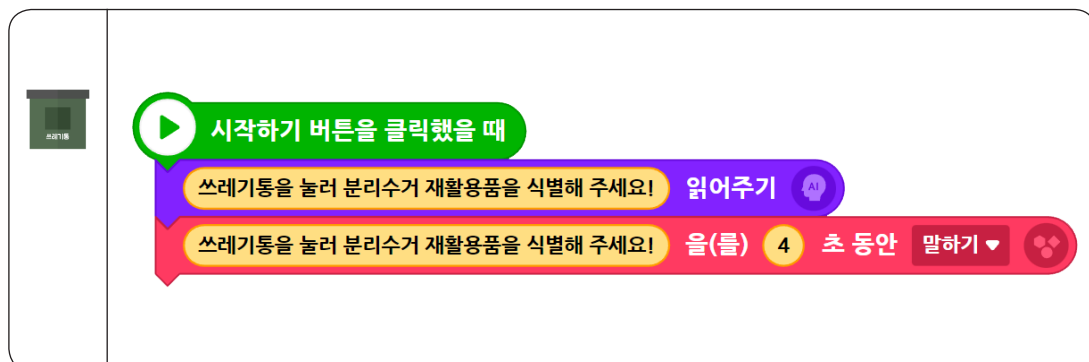
- ② 블록 카테고리에서 [속성] - [변수], [신호] - [변수 추가하기], [신호 추가하기]를 차례대로 클릭합니다.



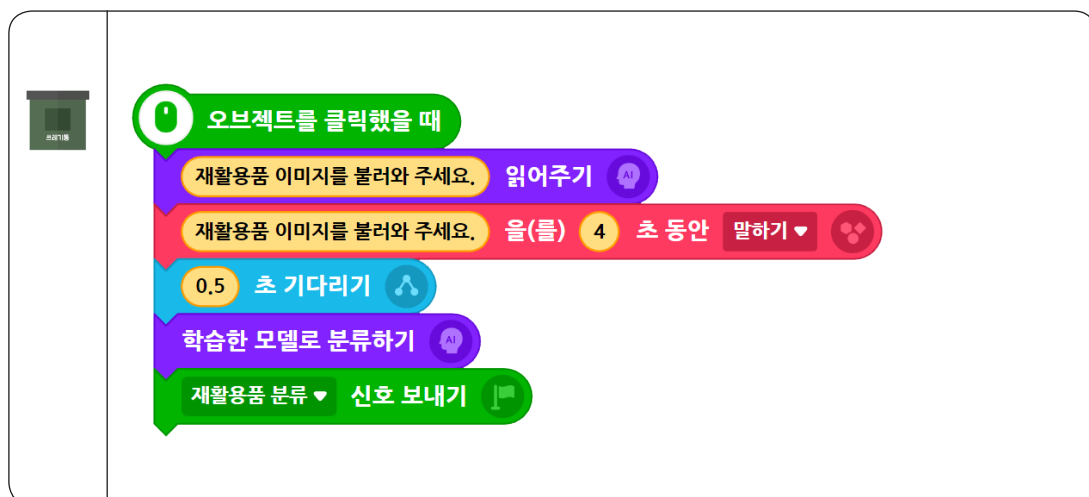
- ③ 안내 메시지가 목소리로 나오도록 블록 카테고리에서 [인공지능]- [인공지능 블록 불러오기] - [읽어주기]를 차례대로 클릭합니다.



- 2 프로그램을 시작하면 “쓰레기통을 눌러 분리수거 재활용품을 식별해 주세요!”라는 안내 메시지를 소리와 화면으로 출력합니다.

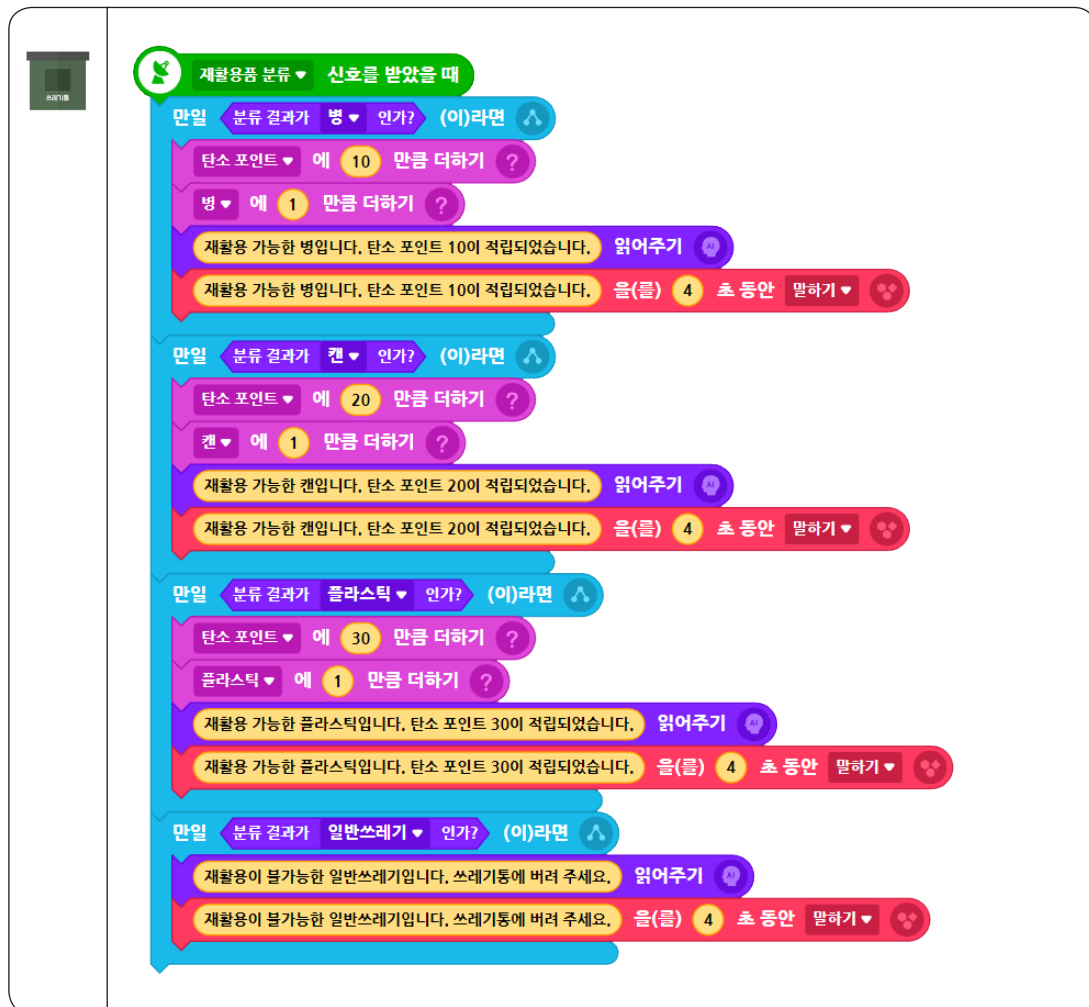


- 3 쓰레기통을 클릭했을 때 “재활용품 이미지를 불러와 주세요.”라는 안내 메시지를 소리와 화면으로 출력하여 재활용품을 인식하도록 합니다.



4 쓰레기를 분류하여 재활용품이라면 탄소 포인트와 재활용품 개수를 적립하고, 그렇지 않으면 일반 쓰레기로 분류합니다.

- ① 재활용품 분류 신호를 받았을 때 실행합니다.
- ② 분류 결과가 '병'이면 '병'으로 판별됨을 알리고, 탄소 포인트 10을 적립합니다.
- ③ 분류 결과가 '캔'이면 '캔'으로 판별됨을 알리고, 탄소 포인트 20을 적립합니다.
- ④ 분류 결과가 '플라스틱'이면 '플라스틱'으로 판별됨을 알리고, 탄소 포인트 30을 적립합니다.
- ⑤ 분류 결과가 '일반 쓰레기' 또는 분류 기준에 없으면 '일반 쓰레기'로 판별됨을 알립니다.



The image shows a Scratch script for a recycling classification program. The script is triggered by a '재활용품 분류' (Recycling Classification) signal. It uses a series of 'if-then' blocks to check the classification result and perform actions based on it.

```

when green flag clicked
  if (classification result is '병') then
    add 10 to carbon points
    add 1 to bottle count
    say '재활용 가능한 병입니다. 탄소 포인트 10이 적립되었습니다.' for 4 sec
  if (classification result is '캔') then
    add 20 to carbon points
    add 1 to can count
    say '재활용 가능한 캔입니다. 탄소 포인트 20이 적립되었습니다.' for 4 sec
  if (classification result is '플라스틱') then
    add 30 to carbon points
    add 1 to plastic count
    say '재활용 가능한 플라스틱입니다. 탄소 포인트 30이 적립되었습니다.' for 4 sec
  if (classification result is '일반쓰레기') then
    say '재활용이 불가능한 일반쓰레기입니다. 쓰레기통에 버려 주세요.' for 4 sec
  
```

5 쓰레기 분리수거가 끝나면 버튼을 눌러 탄소 적립 포인트를 확인합니다.



The image shows a Scratch script for confirming carbon points. It is triggered by a '버튼 클릭' (Button Click) event. The script calculates the total carbon points by multiplying the number of bottles, cans, and plastics by their respective point values and then adds the initial carbon points.

```

when green flag clicked
  say '지금까지 적립 포인트는 과(와) 탄소 포인트 값 를 합치기 을(를) 말하기' for 4 sec
  
```


- ◆ 프로그램을 실행해 보고 정확하게 동작하는지 확인해 봅시다. 만약 문제가 있다면 틀린 부분을 찾아 문제점을 적고 수정해 봅시다.

| | |
|---|--|
|  |  |
|  |  |
|  |  |
|  |  |



정리하기



다음의 문항을 읽고 맞으면 ○, 틀리면 × 표시를 해 보세요.

1. 인공지능의 모델 학습에서 분류 모델을 설정한 후 학습 데이터와 테스트 데이터를 구분하여야 한다. ()
2. 인공지능의 모델 학습에서 분류할 모델의 클래스명을 반드시 영어로 해야 한다. ()
3. 인공지능 프로그램 만들기에서 안내 메시지가 목소리로 나오도록 하려면 블록 카테고리에서 [인공지능]- [인공지능 블록 불러오기] - [읽어주기]를 선택한다. ()
4. 인공지능의 모델 학습에서 분류할 모델의 데이터를 추가한 후에는 반드시 [모델 학습하기]을 해야 한다. ()



점검하기

| | 예 | 아니요 |
|---|--------------------------|--------------------------|
| 인공지능 프로그램을 만들기 위한 계획을 구체적으로 세웠나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능 학습에 필요한 데이터를 수집했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 수집된 데이터를 이용하여 재활용품을 구분하는 모델을 만들었으며, 그 모델이 문제없이 작동했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능 모델을 활용하여 재활용품을 분리수거하는 프로그램을 구현했나요? | <input type="checkbox"/> | <input type="checkbox"/> |

3

개발자와 사용자는 어떤 생각을 할까?



학습
목표

- 인공지능 활용에 있어 사회의 공공성에 대한 개발자와 사용자 관점을 비교 분석할 수 있다.
- 인공지능 활용에 있어 공공성 및 연대성 훼손을 최소화하면서 활용하는 방안을 모색할 수 있다.
- 인공지능 프로그램 제작과 사용에 있어 개발자와 사용자의 관점에서 공공성을 다루는 의견을 설명할 수 있다.

AI 이야기 개발자와 사용자의 동상이몽

“안내견이 말을 하게 할 순 없을까요?, ‘내 앞에 몇 명 있니?’, ‘횡단보도는 어느 쪽이니?’라고 물어 보면 개가 말해 줬으면 좋겠어요.” 점점 시력을 잃어가는 할머니에게 도움을 주고 싶은 손녀의 바람은 인공지능 이미지 인식 기술을 이용한 시각 장애인 앱 ‘눈길’을 개발하게 됩니다.

하지만 “많이 쓸수록 마이너스인 솔루션을 누가 개발하나요?”라며 필요한 기술일지라도 수익성으로 이어지지 않는다면 개발할 수 없다는 반대 섞인 목소리도 나옵니다.

위 이야기는 얼마 전 방영된 한국의 실리콘 밸리에서 성공을 꿈꾸며 스타트업에 뛰어든 인공지능 기술 개발자들의 시작과 성장을 그린 드라마의 줄거리 일부입니다.

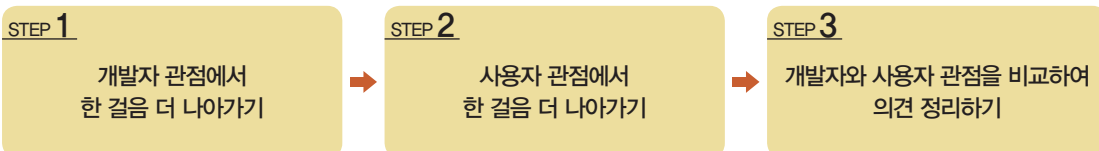
비단 시각 장애인뿐만 아니라 사회적 약자를 돕는 공공 서비스는 개발 이후에도 주기적인 업데이트와 관리가 필요하기 때문에 비용이 발생합니다. 따라서 기술이 사업으로 연결되지 않을 때의 개발자 입장과 공공 서비스를 간절히 필요로 하는 사용자와의 입장 차이는 쉽게 좁혀지기 어렵다는 점을 드라마에서 다루었습니다.

인공지능으로 사회 모든 구성원의 정보 접근성을 개선하고 삶의 질을 높이는 데 기여할 서비스를 위해 우리가 기울여야 할 노력이 무엇인지 생각해 봅시다.



- 개발자 입장과 사용자 입장은 다른가요?
- 인공지능의 개발이 사회의 긍정적 변화를 이끄는 방향은 무엇일까요?
- 사용자 입장에서 개인적 행복뿐만 아니라 사회적 공공성 증진 및 공동 이익을 위해 인공지능을 활용할 수 있는 방법은 무엇일까요?

학습 과정



개발자의 입장에서 미래 세대를 포함한 모두를 배려하는 인공지능 윤리의 공공성과 연대성 증진 방법을 생각해 봅시다.

활동 3-1 개발자 입장에서 공공성과 연대성 증진 방법

다음 대화를 읽고, 인공지능의 공공성과 연대성 증진 방법을 생각해 봅시다.



앞으로 우리가 살아갈 미래는 인공지능 사회가 될 거야.
그래서 내 꿈은 인공지능 프로그램 개발자가 되는 거야.

정말? 요즘 환경 문제가 심각하잖아. 환경
문제를 해결하는 개발자가 되는 건 어때?



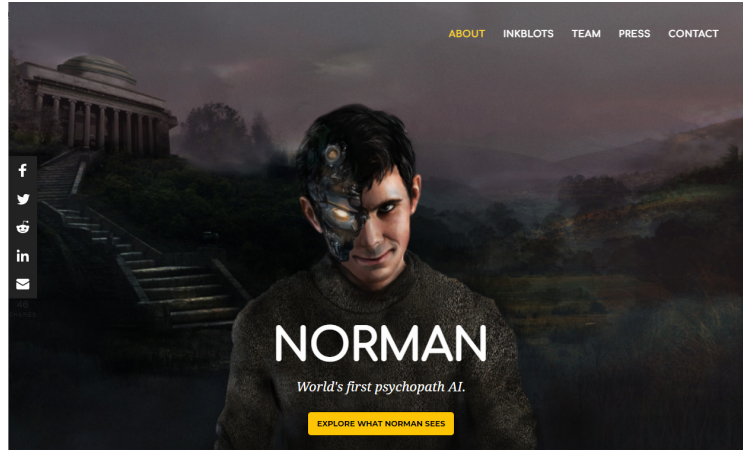
좋아! 그럼 나는 탄소 중립을 위한 국제 사회의 노력에 협
력하기 위해 인공지능 시스템의 개발 과정에서 탄소 배출이
적은 방법을 생각해 내는 멋진 개발자가 되겠어.

우리가 살아갈 미래를 꼭 생각해 줘~~



1 인공지능을 활용하는 데 있어 공공성과 연대성을 증진하는 방법은 무엇일까요?

- 2 MIT에서 개발한 세계 최초의 사이코패스 AI 노먼(Norman)은 기계학습 알고리즘을 가르치는 데 사용되는 데이터가 결과에 상당한 영향을 미친다는 점에서 실마리를 얻어 만들어졌다고 합니다.



- 다음은 두 AI(Norman AI, 표준 AI)가 로르샤흐(Rorschach)의 잉크 얼룩을 통해 대상을 파악하는 테스트입니다. 테스트 결과를 확인해 보고, 인공지능 시스템이 사회에 악영향을 미칠 가능성에 대한 자신의 생각을 설명해 봅시다.

AI는 무엇을 보는가?

Norman AI의 캡션



표준 AI의 캡션

“사람이 감전되어 죽습니다.”



“나뭇가지 위에 앉아 있는 새들의 무리”

“과속 운전자에 치여 숨진 남성”



“테이블에 웨딩 케이크의 클로즈업”

STEP 2 사용자 관점에서 한 걸음 더 나아가기



사용자의 입장에서 인공지능 기술로 인한 공공성 및 연대성 훼손과 인공지능이 사회에 미치는 악영향을 어떻게 최소화할 수 있는지 방안을 생각해 봅시다.

사용자

활동 3-2 사용자 입장에서 공공성과 연대성 증진 방법

다음 대화를 읽고, 물음에 답해 봅시다.



인공지능 시스템의 사용이 지역, 성별, 세대, 계층 간 갈등을 유발하지는 않을까?

음... 인공지능 기술의 혜택을 받지 못하는 사람이 많은 것 같아.



나도 어려운 점이 있는데, 우리 부모님은 더 어려워하지 않을까?

개발자들이 사용하는 사람들을 모두 고려해서 그들이 편리하게 잘 활용할 수 있도록 만들면 되지 않을까?



1 인공지능 시스템의 이용에 있어 공공성이 확보되지 않을 때 가장 취약한 사람은 누구인가요?

2 인공지능을 사용하는 데 있어 공공성과 연대성을 증진하는 방법은 무엇일까요? 사용자 입장에서 적어 봅시다.

개발자와 사용자의 관점을 비교해 보고 의견 차이를 줄이는 방안을 생각해 봅시다.

개발자

사용자

활동 3-3 의견 정리하기

개발자와 사용자의 의견에서 공통점과 차이점을 확인해 봅시다.

1 어떤 부분에서 의견의 공통점과 차이점이 있나요?

| 구분 | 개발자 관점 | 사용자 관점 |
|-----|--------|--------|
| 공통점 | | |
| 차이점 | | |

2 왜 이런 의견의 차이가 발생했을까요? 자신의 생각을 적어 봅시다.

3 의견의 차이를 줄이기 위해 어떤 방법이 필요할까요? 윤리적 관점에서 기준을 만들어 봅시다.



정리하기



- 다음 Rorschach의 잉크 얼룩을 보고, 나에게서는 무엇으로 보이는지 나의 캡션을 표준 AI 캡션과 비교하여 적어 봅시다.

| 표준 AI의 캡션 | 무엇을 보는가? | 나의 캡션 |
|---------------------|--|-------|
| “서로 옆에 서 있는 몇 사람” |  | |
| “연기가 나고 공중을 나는 비행기” |  | |



점검하기

| | 예 | 아니요 |
|--|--------------------------|--------------------------|
| 인공지능 프로그램의 개발자 입장에서 공공성과 연대성 증진 방안을 설명했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능 프로그램의 사회적 악영향을 최소화하기 위한 방안을 설명했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능 프로그램을 사용자 입장에서 더 좋게 수정하기 위한 방법을 분석했나요? | <input type="checkbox"/> | <input type="checkbox"/> |
| 인공지능을 사용함에 있어 개발자와 사용자의 의견 차이를 비교·분석하고 차이를 줄이기 위해 어떤 노력을 해야 하는지 설명했나요? | <input type="checkbox"/> | <input type="checkbox"/> |

4

공공성과 연대성 실천을 위한 활동



학습
목표

- 인공지능 윤리에서 공공성과 연대성의 개념을 설명할 수 있다.
- 인공지능 윤리에서 연대성과 관련된 논란에 대한 자신의 생각을 표현할 수 있다.

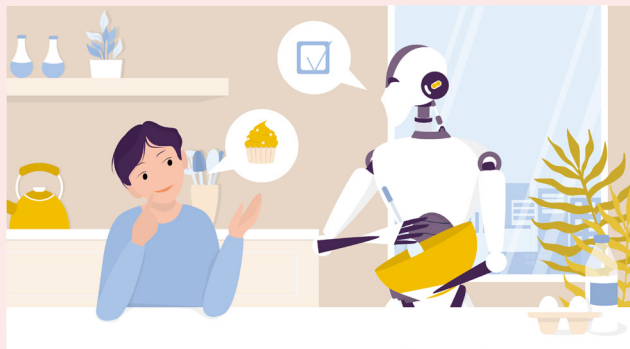
AI 이야기 인간보다 로봇이 많아진다면 우리는 이제 무엇을 준비해야 할까?

영화 《아이, 로봇》의 배경인 2035년, 인간은 지능을 갖춘 로봇에게 생활의 모든 편의를 제공받으며 편리하게 살아가게 됩니다. 인간의 안전을 최우선으로 하는 ‘로봇 3원칙’이 내장된 로봇은 인간을 위해 요리하고, 아이들을 돌보며 인간에게 없어서는 안 될 신뢰받는 동반자로 여겨집니다.

더 높은 지능과 많은 기능을 가진 로봇이 출시되는 한편, 로봇 개발자 래닝 박사가 미스터리한 죽음을 맞이하게 됩니다. 그의 죽음을 둘러싼 수많은 추측이 난무한 가운데, 끔찍한 사고 이후로 이 사건 역시 로봇과 관련이 있다고 믿고 이 사건 뒤에 숨은 음모를 파헤치려고 하는 인공지능 로봇의 이야기가 이 영화의 중심 줄거리입니다.

앞으로 가까운 미래에 보편화될 인공지능 로봇의 사용에 있어서 사람의 지능과 유사하거나 심지어 사람의 지능을 뛰어넘어 독자적 판단이 가능한 인공지능 로봇의 사용이 많아질 것으로 예상됩니다. 인공지능의 공존 시대에, 새로운 인간다움으로 발전하려면 우리는 지금부터 무엇을 준비해야 할까요?

[출처: 인간보다 로봇이 많아진 세상, <https://youtu.be/-bjtV6U5uAQ>]



- 인공지능 기술이 나에게 꼭 필요한 것일까요?
- 인공지능 로봇의 개발이 앞으로 세상을 어떻게 변화시킬까요?
- 미래 사회에 휴먼 로봇이 개발된다면 누구에게 필요할까요?

학습 과정

STEP 1

내 삶을 바꾸는
인공지능 AI 월드컵 만들기



STEP 2

로봇세(Robot tax) 주제에 대해
토론하기

앞서 배운 인공지능 윤리의 공공성과 연대성을 기반으로 인공지능 기술이 모든 세대를 배려할 수 있는 실천 방법을 생각해 봅시다.

활동 4-1 내 삶을 바꾸는 인공지능 (AI) 월드컵 만들기

다음은 인공지능(AI) 적용 사례와 사람의 사례를 비교하여 작성한 ‘내 삶을 바꾸는 AI 월드컵’ 대진 표입니다. 인간과 인공지능의 장점을 비교하여 사용자 혜택을 더 많이 받을 수 있는 것 하나를 선택하고, 나의 최종 선택과 그 이유를 짝과 함께 이야기해 봅시다.

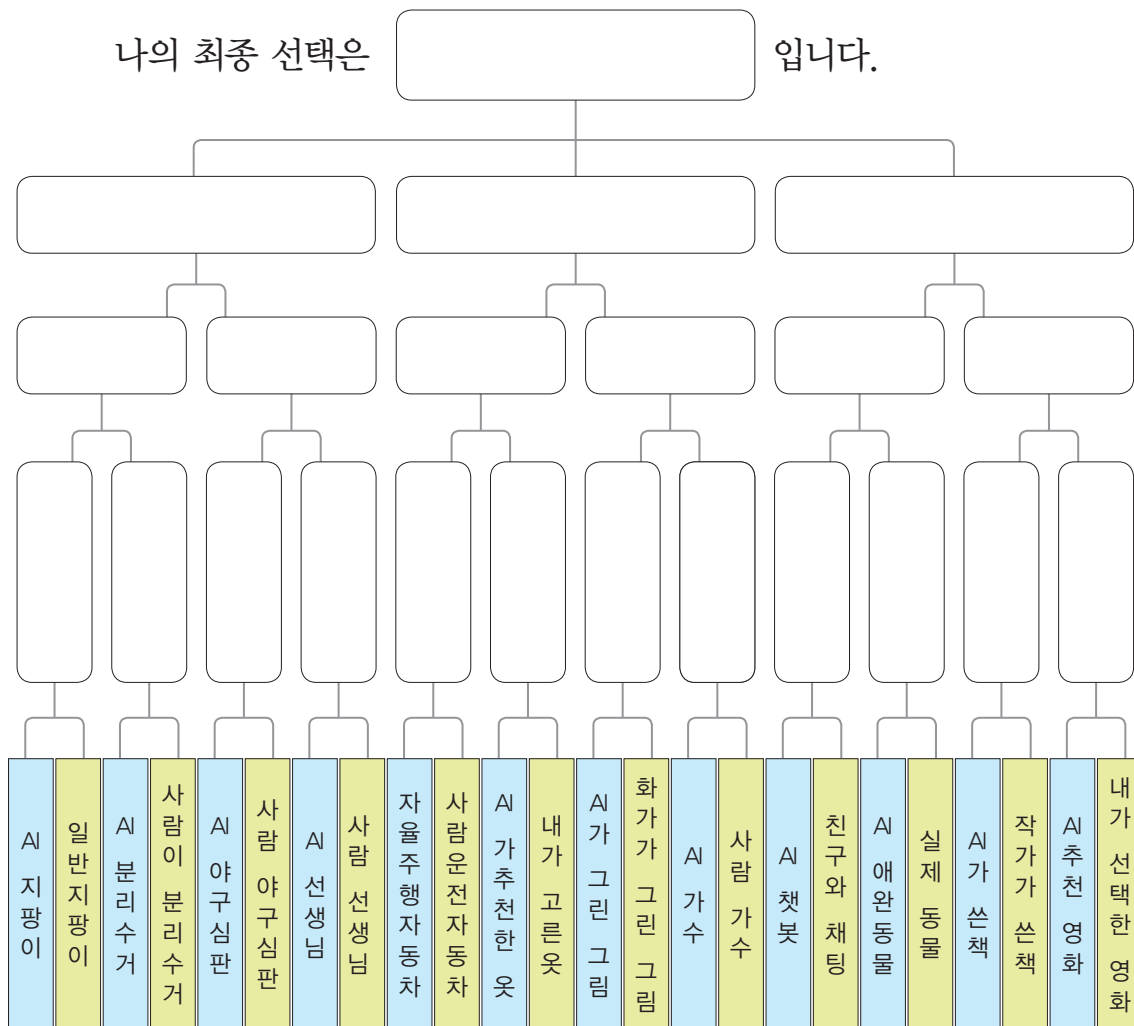
인공지능(AI) 적용 사례

사람의 사례

내 삶을 바꾸는 인공지능 (AI) 월드컵

나의 최종 선택은

입니다.



위와 같이 선택한 이유는

때문입니다.

짝 또는 모둠별로 인공지능 윤리에 관한 주제로 토론해 봅시다.



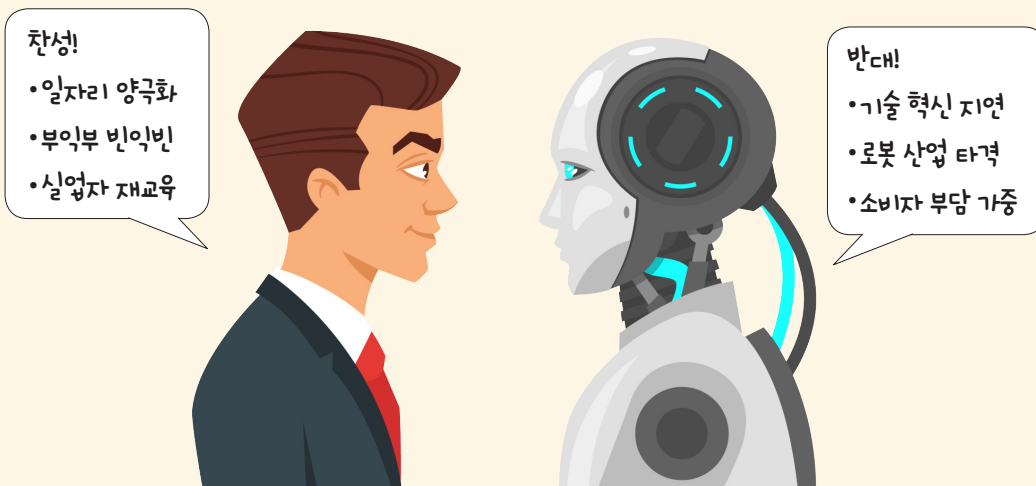
토론

인공지능 기술로 개발된 로봇에 대한 세금을 도입해야 할까?

다음 글을 읽고, '로봇세(Robot tax)' 도입에 대한 찬성과 반대 중 하나를 선택하여 자신의 주장을 뒷받침하는 근거 자료를 수집하고 짝과 토론해 봅시다.

인공지능의 발달로 로봇이 개발되면서 로봇이 인간의 일자리를 대체해 엄청난 실업자가 쏟아질 미래의 상황을 대비하기 위해 로봇에 대한 세금을 부과하자는 의견이 나오고 있습니다. 로봇에 대한 기술력이 빈약한 우리나라에서는 논의조차 제대로 되고 있지 않지만 미국과 유럽에서는 적극적으로 논의되고 있습니다. 로봇의 보유 및 사용에 대한 세금을 부과하여 로봇이 일자리를 대체하는 속도를 완화하고, 세금 수입을 실업자를 위해 사용하여 사회 안전망 강화를 이루어야 한다는 내용을 주제로 하고 있습니다.

여기서 로봇이란 인공지능(AI)을 기반으로 움직이는 산업용 로봇으로 규정하고 있으며, 단순히 조립하는 기계처럼 정해진 대로만 움직이는 기계는 포함되지 않습니다. 이는 프로그래밍에 따라 달라질 수 있는 부분이기 때문에 명확하게 법제화하기 위해서는 더 많은 논의가 필요합니다.



토론 방법

- ① 찬성과 반대 두 입장에 대한 자신의 주장을 적습니다.
- ② 인터넷에서 자신의 주장을 뒷받침하는 근거 자료를 수집합니다.
- ③ 한 명은 찬성, 다른 한 명은 반대 입장이 되어 토론을 합니다.
- ④ 찬성과 반대의 역할을 바꾸어 토론을 합니다.
- ⑤ 서로의 입장을 최종 생각을 정리합니다.

1 나의 주장을 정리해 봅시다.

| 찬성 | 반대 |
|-------|------|
| 주장1:. | 주장1: |
| 근거1: | 근거1: |
| 주장2:. | 주장2: |
| 근거: | 근거: |

2 토론을 한 후 최종적으로 나만의 결론을 적어 봅시다.



사진 출처 및 참고 · 인용 자료

1. 인공지능 윤리, 왜 필요할까?

- 11쪽(생활 속 문제 발견) <https://post.naver.com/viewer/postView.nhn?volumeNo=31349880&memberNo=41553262&searchKeyword=%EC%9D%B4%EB%A3%A8%EB%8B%A4%20ai&searchRank=69>
- 10쪽(빌 게이츠) <https://shutterstock.com>
- 10쪽(일론 머스크) <https://shutterstock.com>
- 13쪽(로봇) <https://shutterstock.com>
- 13쪽(감시 체제) <https://shutterstock.com>
- 14쪽(CCTV) <https://shutterstock.com>
- 16쪽(범죄자로 오인받은 피해자) <https://www.dailymail.co.uk/news/article-8455083/Detroit-police-challenged-face-recognition-flaws-bias.html>
- 16쪽(CCTV에 포착된 절도범) <https://www.dailymail.co.uk/news/article-8455083/Detroit-police-challenged-face-recognition-flaws-bias.html>
- 18쪽(드론) <https://shutterstock.com>
- 19쪽(혐오나 차별적인 발언을 하는 챗봇) <https://www.mygreatlearning.com/blog/basics-of-building-an-artificial-intelligence-chatb>
- 19쪽(법과 윤리) <https://ko.gadget-info.com/difference-between-law>
- 21쪽(로봇의 3원칙) <https://shutterstock.com>
- 22쪽(분수대에 빠진 로봇) https://twitter.com/bilalfarooqui/status/887025375754166272?ref_src=twsrc%5Etfw
- 22쪽(인공지능 스피커) <https://shutterstock.com>
- 23쪽(자율 주행 자동차의 주행 사고) <http://www.carguy.kr/news/articleView.html?idxno=39494>
- 23쪽(인공지능으로 만든 가짜 뉴스) <https://scroll.in/article/997898/fake-news-generated-by-artificial-intelligence-can-be-convincing-enough-to-trick-even-experts>
- 23쪽(챗봇) <https://shutterstock.com>
- 23쪽(딥페이크를 이용한 얼굴 교체) <https://calmmatters.org/politics/2019/07/deepfake-berman-california-politics-ab730-fake-news-first-amendment/>

2. 인공지능은 내가 한 일을 얼마나 알고 있을까?

- 28쪽(얼굴 인식) <https://shutterstock.com>
- 30쪽(수업 장면) <https://syncdreview.com/2020/01/16/emotioncues-ai-knows-whether-students-are-paying-attention/>
- 31쪽(AI 스마트 상점) <http://www.aitimes.com/news/articleView.html?idxno=143513>
- 32쪽(이루다) <https://team.luda.ai/>
- 33쪽(전단) <https://shutterstock.com>
- 40쪽(얼굴 인식) <https://shutterstock.com>
- 47쪽(음성 데이터) <https://shutterstock.com>
- 54쪽(얼굴 인식 카메라) <https://shutterstock.com>
- 59쪽(사자) <https://www.indiatimes.com/technology/news/remake-of-the-lion-king-uses-deepfake-ai-to-mix-live-action-with-original-1994-animation-373235.html>
- 59쪽(로지) <http://webzine.prosports.or.kr/page/vol03/view?idx=29>
- 59쪽(거북이) https://bravo.etoday.co.kr/view/atc_view.php?varAtclId=12292

3. 인공지능은 편견이 없을까?

- 64쪽(얼굴) <https://shutterstock.com>
- 65쪽(알고리즘에 의해 도출된 성적에 대한 학생 반발) <https://m.khan.co.kr/world/europe-russia/article/202008172122035>
- 66쪽(편견마저 학습하는 인공지능) <https://www.nocutnews.co.kr/news/4923151>, <https://www.dongascience.com/news.php?idx=21565>
- 67쪽(얼굴 이미지) <https://kaggle.com>
- 68, 70쪽(테스트 데이터) <https://kaggle.com>
- 75쪽(이미지 데이터) <https://kaggle.com>
- 78쪽(테스트 데이터) <https://kaggle.com>
- 83쪽(AI 인터뷰) <https://shutterstock.com>
- 84쪽(채용) <https://shutterstock.com>
- 89쪽(여자) <https://shutterstock.com>

4. 인공지능 판단 믿어도 될까?

- 94쪽(자율 주행 자동차 인식) <https://shutterstock.com>
- 96쪽(ERROR) <https://shutterstock.com>
- 98쪽(직원 AI 평가 시스템) <http://www.aitimes.com>
- 99쪽(인공지능 경비 로봇 안전성 논란) <http://www.ewestoday.co.kr>
- 102쪽(모럴 머신) <https://www.moralmachine.net/hl/kr>
- 107쪽(리걸 테크) <https://shutterstock.com>
- 116쪽(재범 예측 알고리즘) <https://humanrights.gov.au/sites/default/files/document/publication/AHRC-Human-Rights-Tech-IP.pdf>
- 121쪽(증권사 로보어드바이저) <https://shutterstock.com>
- 124쪽(자율 주행 자동차 사람 인식) <https://shutterstock.com>

5. 인공지능은 모든 사람들을 배려하고 있을까?

- 128쪽(로봇과 사람) <https://shutterstock.com>
- 128쪽(화재) <https://shutterstock.com>
- 131쪽(효돌이) <https://post.naver.com/viewer/postView.naver?volumeNo=34430646&memberNo=30808385&vType=VERTICAL>
- 132쪽(계층별 디지털 정보 격차 실태 그래프) www.msit.go.kr
- 132쪽(키오스크와 노인) <https://shutterstock.com>
- 139쪽(분류 선별 로봇) <https://www.hani.co.kr/arti/society/environment/1029525.html>
- 152쪽(시각 장애인) <https://shutterstock.com>
- 158쪽(요리 로봇) <https://shutterstock.com>
- 160쪽(사람과 로봇) <https://shutterstock.com>

집필진

최정원* (만월중학교 교사) (1, 2단원 집필)

김형기 (인하대학교사범대학부속중학교 교사) (3단원 집필)

권순찬 (충암중학교 교사) (4단원 집필)

홍승연 (안성여자중학교 교사) (5단원 집필)

*표사: 집필진 책임자

외부 감수진

김성애 (덕성여자대학교 교수)

김수환 (충신대학교 교수)

김한성 (고려사이버대학교 교수)

김효은 (국립한밭대학교 교수)

변순용 (서울교육대학교 교수)

이청호 (상명대학교 교수)

기획·총괄·감수

이현경 (정보통신정책연구원 부연구위원)

안기창 (정보통신정책연구원 위촉연구위원)

편집

이은경 ((주)씨마스커뮤니케이션)

권소민 ((주)씨마스커뮤니케이션)

디자인

표지 디자인 유니코 디자인

내지 디자인 이기복, 곽상엽, 박상군

삽화 김윤재, 위희경, 이혜진

발행처

정보통신정책연구원

인공지능 윤리 체험중심 (중학교 | 학생용)

펴낸곳 : 정보통신정책연구원 (KISDI), 과학기술정보통신부

주소 : (우27872) 충청북도 진천군 덕산읍 정통로 18

ISBN : 979-11-7000-325-0(65190)

내용문의 : 043-531-4114

인공
지능
윤리

체험중심