

소프트웨어야 놀자 AI 교육 시리즈
프로젝트 - 5

나만의 번역기 만들기



들어가며

OECD(경제협력개발기구)의 교육2030에서 제시한 미래 아이들을 위한 교육의 목표는 '개인과 사회의 웰빙(Well-being)'이라고 합니다. 이전의 교육목표는 개인과 사회의 성공에 초점이 맞추어져 있었는데 미래는 더불어 잘 살아가는 방법을 알려주는 것이 중요하다고 보는 것입니다. 잘 사는 것은 아이들이 스스로 삶의 가치를 찾고 자신뿐만 아니라 다른 사람들과 함께 살아가는 사회를 만드는 것입니다. 그래서, 학생의 자기 주체성(Student agency)이라는 개념을 도입하여 학생들의 자율성을 강조합니다.

미래사회에서 우리 아이들은 인공지능과 함께 살아가게 될 것입니다. 2016년 알파고 쇼크로 촉발된 인공지능 발전의 가속화는 이미지 인식, 음성 인식, 번역, 의료, 법률 등 다양한 분야에서 인간의 능력을 따라잡고 있습니다. 인공지능은 어디까지 발전하게 될까요? 미래 아이들은 인공지능과 어떻게 살아가게 될까요? 인간의 능력과 인공지능의 능력이 비교되는 세상에서 학생들은 자신의 가치를 찾고 자율성을 실천할 수 있을까요?

본 교재는 인공지능과 살아갈 아이들에게 인공지능에 대한 기본적인 역량을 길러줄 수 있는 내용으로 구성되어 있습니다. 미래사회에서 아이들이 자율성을 가지고 잘 살아가기 위해서는 인공지능의 발전과 한계를 알아야 합니다. 인공지능이 만능이라는 무한한 기대감과 인공지능이 인간을 지배할 것이라는 막연한 두려움은 모두 무지에서 기인한 것입니다. 막연한 기대감이나 두려움을 없애려면 인공지능에 대해 제대로 알아야 합니다.

산업사회에서 3R's(읽고 쓰고 셈하기)가 리터러시(Literacy)였던 것처럼 인공지능 시대에는 인공지능에 대한 리터러시가 필요하게 될 것입니다. 초중등 인공지능 교육은 인공지능 개발자를 양성하는 교육이 아닙니다. 국어, 수학, 과학 등을 가르치는 것이 세상에서 살아갈 가장 기본적인 지식과 기능을 갖추도록 하는 교육인 것처럼 미래세대 아이들에게는 인공지능에 대한 이해와 활용이 가장 기본적인 교육이 될 것입니다. 특히, 미래시대는 인공지능을 활용하거나 인공지능과 협업하는 시대가 될 것이므로 어렸을 때부터 바른 관점과 태도를 가질 수 있도록 해주어야 합니다.

이 교재는 아이들에게 인공지능이 무엇인지? 어떤 원리로 작동하는지? 어떤 한계를 가지고 있는지? 어떻게 활용 가능한지에 대한 내용을 가르치고자 만들었습니다. 인공지능에 대한 올바른 교육은 다양한 문제 상황에서의 적합한 활용과 함께 살아가는 사회에 기여하는 태도를 길러줄 수 있습니다. 우리는 이 교재를 통해서 아이들에게 인공지능의 원리를 이해하고 다양한 관점에서의 성찰을 통해 인공지능을 올바르게

활용할 수 있는 능력을 길러주고 싶습니다. 엔트리(Entry)를 통해 다양한 프로젝트를 만들어보고 인공지능의 가능성과 한계를 탐색하고 성찰할 수 있도록, 재미있고 유익한 내용으로 구성하였습니다. 특히, 단순한 인공지능 기능의 사용을 지양하고 그 원리를 이해하고 다양한 활용을 탐색할 수 있도록 구성하였습니다.

본 교재가 교육현장에서 널리 활용되어 미래세대 아이들을 위한 교육에 도움이 되면 좋겠습니다.

- 저자 일동

집필진

김현철 교수	고려대학교 정보대학 컴퓨터학과 교수이며 인공지능, 기계학습 그리고 초중등 SW/AI교육 분야의 연구를 하고 있다.
김수환 교수	고려대에서 컴퓨터교육 박사학위를 받고, 현재 총신대학교에서 컴퓨터 교육을 가르치고 있다. 초중등 SW/AI교육의 정착과 발전을 위해 교육과정 개발, 교사연수 운영 등 다각도로 힘쓰고 있다.
장연주 교사	고려대학교 컴퓨터학과 박사과정
최승윤 교사	고려대학교 컴퓨터학과 박사과정

감수자

주재걸 교수	현재 한국과학기술원(KAIST) AI 대학원에 부교수로 재직 중이며, 컴퓨터 비전 및 자연어 처리 등의 분야에서 딥러닝 기법에 관한 연구를 수행하고 있다.
유정수 단장	현재 교육부 인공지능 교육 정책 자문단 단장, 대통령직속 4차 산업혁명 위원회 제도혁신위원회 위원으로 활동하고 있다.

발행처 네이버 커넥트재단
집필진 김현철 교수(고려대학교), 김수환 교수(총신대학교), 장연주 교사(고려대학교), 최승윤 교사(고려대학교)
감수자 주재걸 교수(한국과학기술원), 유정수 단장(교육부 인공지능 교육정책 자문단)

참고사이트 소프트웨어야 놀자 : <https://www.playsw.or.kr>
엔트리 : <https://playentry.org>
Hello AI World(인공지능 개념 영상) :
https://playsw.or.kr/artificial/video?currentTab=ai_data_video_hello_ai



이 저작물은 크리에이티브 커먼즈 저작자표시-비영리 4.0 국제 라이선스에 따라 이용할 수 있습니다. 단, 크리에이티브 커먼즈 라이선스 조건에도 불구하고, 강사들이 교육 목적으로 (공/사립학교, 지역아동센터 등 공공기관에서 진행되는 수업만 해당) 사용하는 경우에만 영리 목적의 사용이 허용되며, 출판사 또는 학원 등 영리기관에서의 사용은 금지됩니다.

학습목표

관련 인공지능 영역 : 자연어 처리(NLP)/번역, 딥러닝



- ① 컴퓨터가 텍스트를 인식하는 방법을 알아봅시다.
- ② 엔트리 번역 기능을 이용하여 인공지능 번역 프로그램을 만들어봅시다.
- ③ 번역 인공지능 기술이 가져올 수 있는 사회적 영향에 대해 이야기를 나눠봅시다.



교사용 수업 가이드

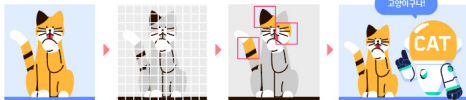
수업 가이드를 다운 받으신 후, 내용 참고하셔서
수업을 진행하실 수 있도록 구성하였습니다.

1) 수업 가이드 & 수업 내용

컴퓨터는 이미지를 어떻게 감지할까요?

수업용 ppt 11~16p1차시 | 2차시 | 3차시 | 4차시

여기에 고양이 이미지 데이터가 있습니다. 이걸 보고 고양이 인지 아닌지를 사람은 쉽게 구분할 수 있습니다. 컴퓨터는 고양이인지 아닌지를 어떻게 알 수 있을까요?



제목

수업용 ppt 페이지 및 차시

해당 수업내용

2) 수업용 ppt

교사용 가이드와 수업용 ppt를 활용해 수업을 진행해보세요.

3) 학생용 워크시트 목차

교사용 교재 뒤에는 학생용 워크시트가 포함되어 있습니다.
해당 수업에 맞는 워크시트를 인쇄하여 수업 진행해주세요.

워크시트		페이지
원리 이해하기	컴퓨터가 텍스트를 인식하는 방법 & 음성합성 알아보기	워크시트 2-7p
실습하기	나만의 번역기 만들기 (1)	워크시트 8p / 엔트리 실습 링크
심화하기	나만의 번역기 만들기 (2)	워크시트 9p / 엔트리 실습 링크
생각해보기	번역 및 음성합성 인공지능이 우리 생활에 미치는 영향	워크시트 10-13p

엔트리 실습 링크 : <https://blog.naver.com/nvplaysw/222108959770>

프로젝트 전개

1개 프로젝트 당 4차시로 진행

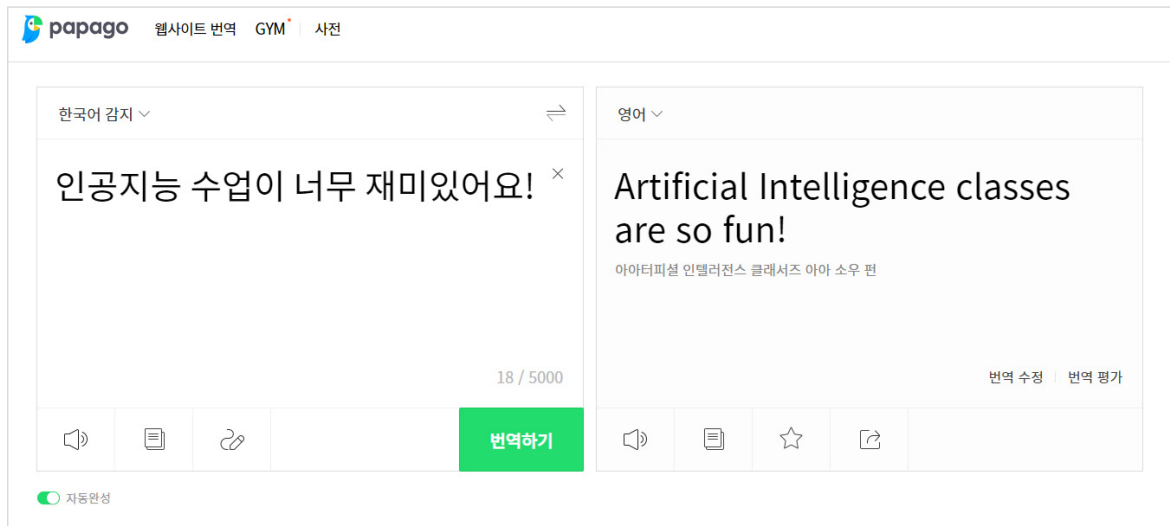
차시	관련 AI 영역	교수·학습활동	시간	준비물/자료
1차시	자연어 처리 (NLP)/번역, 딥러닝	<input type="checkbox"/> 파파고 번역기 체험하기 - 파파고의 번역 기능 사용해보기 <input type="checkbox"/> 컴퓨터는 텍스트를 어떻게 인식할까요? - 텍스트 데이터 처리 - 자연어 처리 <input type="checkbox"/> 음성합성이란 무엇일까요? - 음성합성이란? - 음성합성 체험하기	40'	컴퓨터 언플러그드
2차시		<input type="checkbox"/> 나만의 번역기 만들기(1) - 외국어를 한국어로 바꿔주는 번역기 프로그램 만들기	40'	엔트리
3차시		<input type="checkbox"/> 나만의 번역기 만들기(2) - 외국어를 한국어로 바꿔주는 번역기 프로그램 발전시키기	40'	엔트리
4차시	사회적 영향	<input type="checkbox"/> 번역 및 음성 합성 인공지능이 우리 생활에 미치는 영향 - 번역 인공지능이 우리 생활에 미치는 영향 - 음성 합성 인공지능이 우리 생활에 미치는 영향	40'	활동지

파파고 번역기 체험하기

수업용 ppt 1~7p

1차시 | 2차시 | 3차시 | 4차시

네이버 번역기 파파고를 사용해본 적이 있나요?



파파고는 여러 나라의 언어를 다른 언어로 번역해 주는 번역 프로그램입니다. 스마트 폰 앱으로도 제작되어 있어 편리하게 사용할 수 도 있습니다. 파파고는 현재 15가지 언어를 번역할 수 있고 웹사이트 페이지 번역도 가능합니다. 그리고 필요할 때는 번역한 내용을 음성으로 나타내주는 기능도 있습니다. 그렇다면 컴퓨터는 어떻게 텍스트의 의미를 인식하고 다른 언어로 번역할 수 있을까요? 또 번역한 내용을 어떻게 음성으로 나타낼 수 있을까요?

이번 주제에서는 컴퓨터가 텍스트를 처리하는 방법을 배우고 자연어 처리가 무엇인지 배우도록 하겠습니다. 또, 텍스트를 소리로 표현할 수 있게 해주는 음성 합성에 대해 알아봅니다. 아울러, 엔트리의 번역 기능과 읽어주기 기능을 이용하여 나만의 번역기를 만들어보겠습니다. 마지막으로 다른 나라 언어로 번역해 주는 인공지능 기술과 음성 합성 인공지능 기술이 가져올 수 있는 사회적 영향에 대해 서로 이야기해 보겠습니다.

인공지능 체험하기 →

파파고를 직접 체험해봅시다.

파파고 : <https://papago.naver.com>



컴퓨터는 텍스트를 어떻게 인식할까요?

수업용 ppt 8~13p

1차시 | 2차시 | 3차시 | 4차시

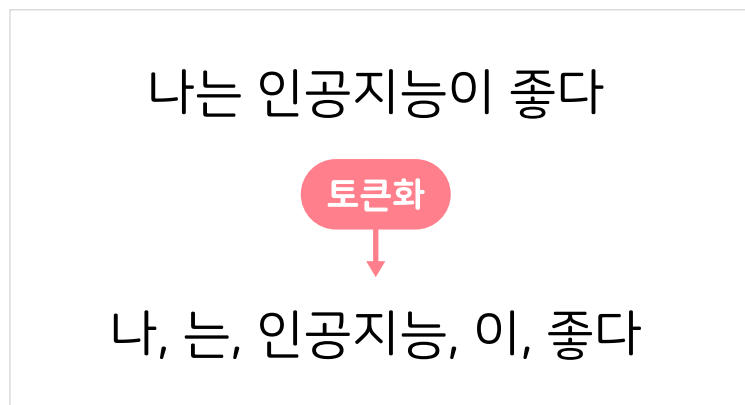
1. 텍스트 데이터 처리

컴퓨터가 텍스트를 인식하고 다른 나라 언어로 번역하기 위해서는 우선 컴퓨터가 이해할 수 있도록 텍스트 데이터를 숫자로 바꾸는 처리 과정이 필요합니다. 그렇다면 컴퓨터는 어떻게 텍스트를 처리할 수 있을까요?

1-1. 토큰화

토큰(Token)은 텍스트 데이터에서 의미를 가지는 최소한의 덩어리를 말합니다. 그리고 텍스트 데이터를 토큰 단위로 나누는 과정을 토큰화라고 합니다. 단어 단위로 토큰을 나누는 것을 '단어 토큰화'라고 하고 문장 단위로 토큰을 나누는 것을 '문장 토큰화'라고 합니다. 토큰은 처리하는 목적에 따라 달라질 수 있습니다.

토큰화를 할 때는 언어의 특성을 고려해야 합니다. 특히 우리나라 말의 경우 '은, 는, 이, 가'와 같은 조사가 있고 영어보다 띄어쓰기가 잘 지켜지지 않는다는 특징이 있기 때문에 이와 같은 특성을 고려해 토큰화하는 것이 중요합니다.



'나는 인공지능이 좋다.'라는 예시문을 토큰화하면 '나', '는', '인공지능', '이', '좋다'의 토큰으로 나눌 수 있습니다. 여기서는 조사도 구분하여 토큰으로 나눌 것을 확인할 수 있습니다.

컴퓨터는 텍스트를 어떻게 인식할까요?

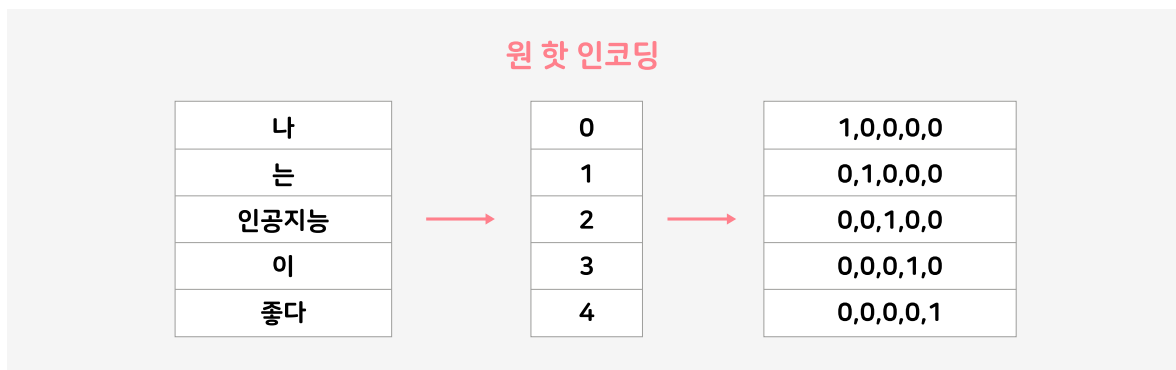
수업용 ppt 14~18p

1차시 | 2차시 | 3차시 | 4차시

1. 텍스트 데이터 처리

1-2. 임베딩

토큰화가 끝나면 이제 임베딩의 단계를 거칩니다. 임베딩이란 하나의 단어를 숫자로 바꾸는 과정으로 여기서는 토큰화된 단어들을 컴퓨터가 이해할 수 있는 숫자로 변환시키는 과정을 말합니다. 우리는 그동안 컴퓨터가 이미지 데이터와 소리 데이터를 컴퓨터가 이해할 수 있는 숫자 0과 1로 바꾸는 과정에 대해 배웠습니다. 텍스트 데이터도 이와 마찬가지로 0과 1로 바꾸는 과정이 필요합니다. 임베딩하는 방법은 다양한데 이번 시간에는 그 중 하나인 원 핫 인코딩에 대해 알아보도록 하겠습니다.



먼저 처리하는 텍스트의 토큰을 모읍니다. 이렇게 모인 토큰의 집합을 '단어 사전'이라고 표현합니다. 단어 사전의 토큰들은 서로 중복이 되지 않게 구성되고 단어 사전의 크기는 사전에 들어 있는 토큰의 개수입니다. 예시문의 단어 사전에는 '나', '는', '인공지능', '이', '좋다'가 들어있습니다. 따라서 이 단어 사전의 크기는 5입니다.

이제 단어 사전에 있는 토큰에 고유한 숫자를 부여합니다. 이때의 숫자는 0부터 시작합니다. 처리할 텍스트 데이터가 양이 많은 경우 자주 사용하는 단어의 빈도수를 고려해 숫자를 부여하기도 합니다.

숫자를 정한 후 이제 해당하는 위치에 숫자 1을 입력합니다. 그리고 나머지 부분은 0을 입력합니다. 예시문의 '인공지능' 토큰은 이제 (0,0,1,0,0) 이라는 숫자로 변경된 것을 확인할 수 있습니다.

원 핫 인코딩은 사용하기 매우 쉽다는 장점이 있지만 토큰의 개수가 많아지면 변경된 숫자의 길이가 매우 길어지는 문제가 생기는 단점도 있습니다. 예를 들어 단어 사전에 있는 토큰의 수가 100개일 때를 생각해 본다면 0과 1이 총 100개나 되는 엄청 긴 숫자로 변형될 것입니다.

컴퓨터는 텍스트를 어떻게 인식할까요?

수업용 ppt 19~26p

1차시 | 2차시 | 3차시 | 4차시

2. 이미지에서 특징 추출(Feature Extraction)하기

<보기>

희석이는 햄버거를 너무 좋아해서 매일 햄버거를 먹습니다.

Q. 보기의 문장을 토큰화해보세요.

Q. 토큰들을 모아 단어 사전을 만들어보세요.

Q. 원 핫 인코딩으로 토큰들을 숫자로 바꾸어보세요.

컴퓨터는 텍스트를 어떻게 인식할까요?

수업용 ppt 27~30p

1차시 | 2차시 | 3차시 | 4차시

2. 자연어 처리

텍스트 데이터를 숫자로 바꾸었다고 해서 컴퓨터가 이 데이터가 어떤 내용을 담고 있는지 인식하지는 못합니다. 그저 컴퓨터가 알아볼 수 있는 형태로 텍스트를 바꾼 것이고 여기까지는 인공지능의 영역이 아닙니다. 컴퓨터가 텍스트를 인식한다는 것은 텍스트 속에서 특징(feature)을 추출하고 그것을 분석해서 그 텍스트가 어떤 의미인지 인식할 수 있어야 합니다.

텍스트 데이터에서 특징을 추출하는 것은 매우 어려운 일입니다. '약이 쓰다'와 '글을 쓰다' 두 문장을 보면 공통적으로 '쓰다'라는 단어를 사용하고 있습니다. 하지만 그 의미는 다르게 사용되고 있습니다. 우리는 문맥이나 상황을 고려해 이 텍스트가 어떤 의미를 담고 있는지 쉽게 알 수 있습니다. 하지만 컴퓨터가 이런 것들을 모두 특징들을 고려하는 것은 매우 힘든 일입니다. 최근의 인공지능 기술(딥러닝)이 발전하면서 인공지능은 이러한 텍스트 데이터에서 특징을 스스로 찾아내고 조합하여 분석해 텍스트 데이터를 인식할 수 있게 되었습니다.

이렇게 컴퓨터가 인간의 언어를 인식하고 목적에 맞게 처리할 수 있도록 하는 것을 자연어 처리(Natural Language Processing, NLP)라고 합니다. 자연어 처리는 인공지능의 한 분야로 여기서 자연어는 우리가 일상생활에서 말하고 듣고 읽고 쓰기를 하는 언어를 말합니다. 따라서 자연어 처리를 이해하기 위해서는 인공지능에 대한 이해뿐만 아니라 언어에 대한 이해도 가지고 있어야 합니다.



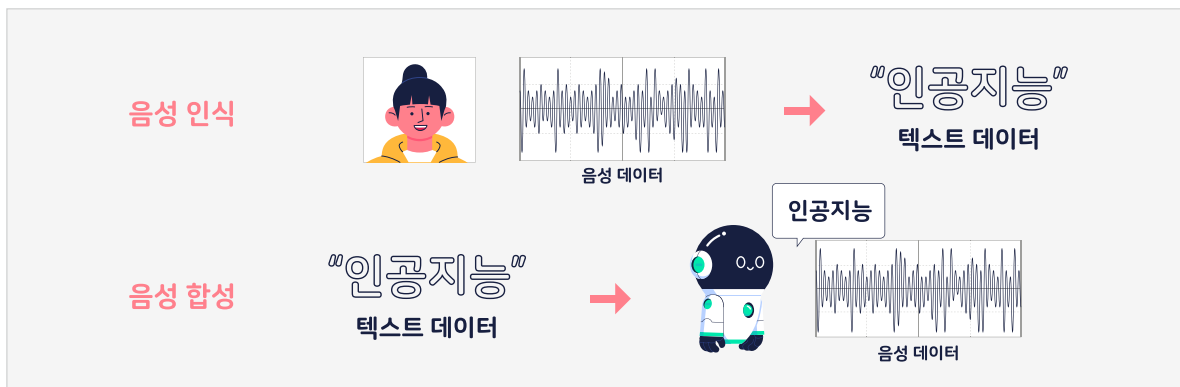
자연어 처리에는 다양한 분야가 있습니다. 먼저 텍스트를 인식하여 이 텍스트가 긍정적인 내용인지 부정적인 내용인지 분류해내는 것을 감정 분석이라고 합니다. 이러한 감정 분석은 고객들의 리뷰를 관리하는 데 자주 사용됩니다. 또한, 자연어 처리는 문서를 자동으로 요약해주는 데 사용되기도 하고, 오류를 자동으로 수정해 주는 데에도 사용됩니다. 컴퓨터를 이용해서 한 언어에서 다른 언어로 바꾸는 번역도 자연어 처리의 한 분야입니다.

음성 합성이란 무엇일까요?

수업용 ppt 31~35p

1차시 | 2차시 | 3차시 | 4차시

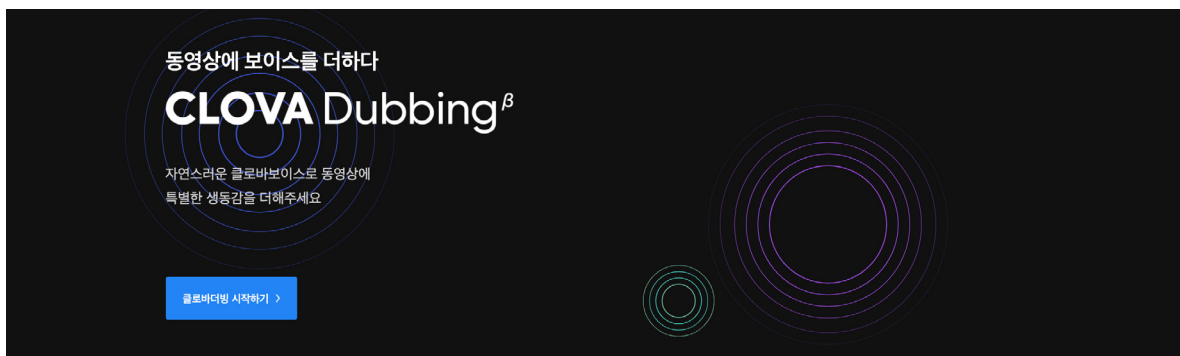
1. 음성 합성이란?



사람이 직접 녹음을 할 때에는 말의 내용이 바뀔 때 마다 새로 녹음해야 하는 문제가 있었습니다. 음성 합성 기술이 발달함에 따라 덕분에 사람이 일일이 녹음을 할 필요가 없어졌습니다. 인공지능 기술이 발달하면서 기계가 내는 목소리가 더욱더 자연스러워지게 되었습니다. 현재에는 자동 응답 서비스, 버스나 지하철 안내방송, 책이나 뉴스 기사 읽어주기 등등 다양한 분야에서 사용되고 있습니다.

2. 음성 합성 체험하기

음성 합성 기술을 이용해 만들어진 네이버의 클로바 더빙을 체험해봅시다.



클로바 더빙 체험하기 : <https://clovadubbing.naver.com>

음성 합성이란 무엇일까요?

수업용 ppt 36~39p

1차시 | 2차시 | 3차시 | 4차시

※ 학습한 내용을 정리해봅시다.

Q. 컴퓨터는 텍스트를 어떻게 인식할까요?

텍스트를 (), ()의 숫자 형태로 처리한 뒤,
데이터에서 ()을 추출하여 어떤 텍스트인지 인식합니다.

Q. 컴퓨터가 인간의 언어를 인식하고 목적에 맞게 처리할 수 있도록 하는 것을
()라고 합니다.

Q. 텍스트 데이터를 사람의 목소리로 바꾸는 인공지능 기술을 ()이라고 합니다.

나만의 번역기 만들기 (1)

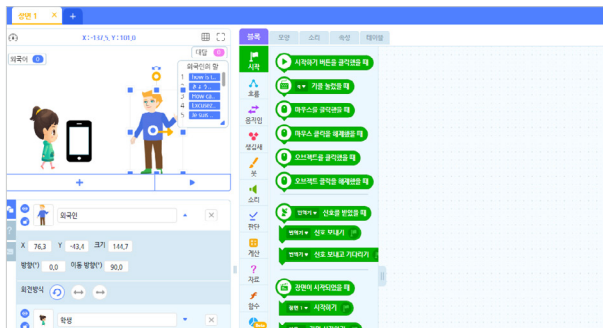
수업용 ppt 40~43p

1차시 | 2차시 | 3차시 | 4차시

1. 만들어볼까요?

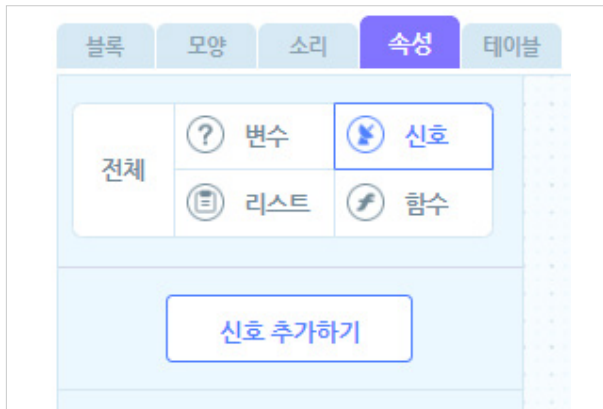
지금부터 엔트리의 번역 기능과 읽어주기 기능을 사용하여 나만의 번역기를 만들어보겠습니다.

<STEP1> 3가지 신호를 만들고, '외국어' 변수 추가하기

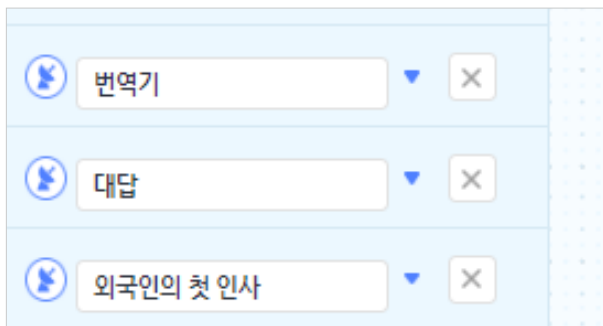


1) 엔트리 공유 주소

<http://naver.me/xtbMUj9I>에 접속하여, 코드보기를 클릭합니다.



2) 먼저 프로젝트에 사용할 신호를 추가해보겠습니다. 속성 탭에서 신호를 클릭하고 '신호 추가하기'를 클릭하세요.

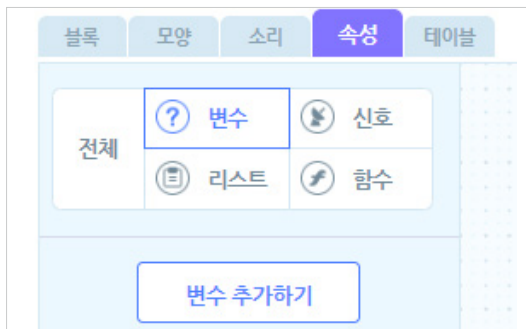


3) '번역기', '대답', '외국인의 첫인사' 3가지 신호를 만들어주세요. 신호의 이름은 상황에 맞게 변경할 수 있습니다. 하지만 꼭 기억하기 쉬운 이름으로 사용해주세요.

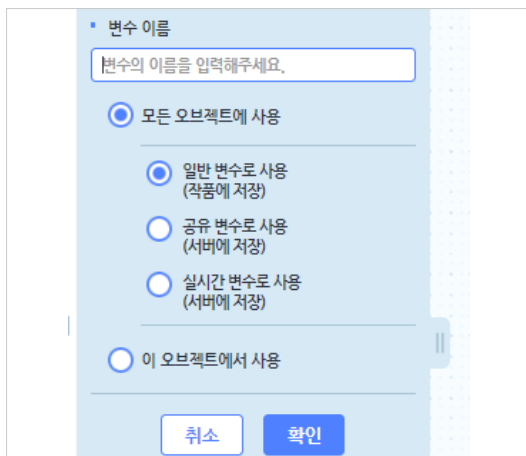
나만의 번역기 만들기 (1)

수업용 ppt 44p

1차시 | 2차시 | 3차시 | 4차시



4) 이번엔 변수 탭을 클릭하여 변수를 추가해보도록 하겠습니다. '변수 추가하기'를 클릭하세요.



5) 변수 이름은 '외국어'로 지정해 주세요. 이 변수는 모든 오브젝트에서 사용할 변수이므로 '모든 오브젝트에 사용'을 선택하고 확인을 클릭하세요.

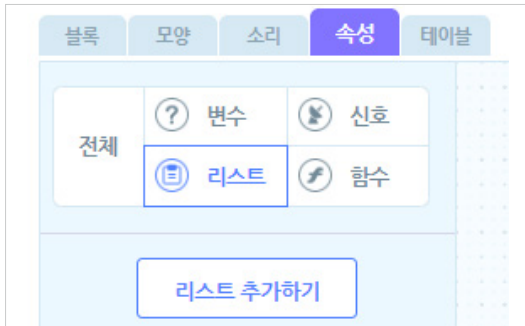


6) 다음과 같이 변수가 추가된 것을 확인할 수 있습니다.

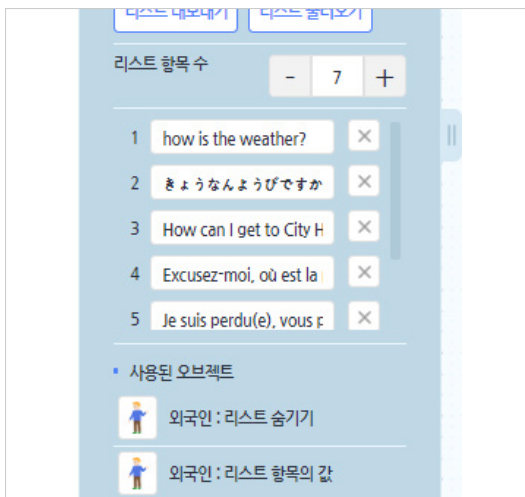
나만의 번역기 만들기 (1)

수업용 ppt 45~46p

1차시 | 2차시 | 3차시 | 4차시

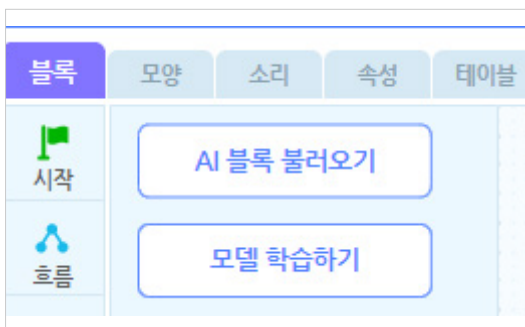


7) 이번에는 리스트를 확인하겠습니다. 리스트는 미리 작성되어 있으니 확인만 하면 됩니다. 리스트 탭을 클릭하세요.



8) 미리 등록된 '외국인의 말' 리스트를 확인해보세요. 총 7개의 외국어 문장이 들어 있는 리스트입니다. 여기에서 리스트 항목을 추가하여 더 다양한 외국어 문장을 추가할 수 있습니다.

<STEP2> '번역', '읽어주기' 블록 추가하고, 외국인 오브젝트 블록 조립하기



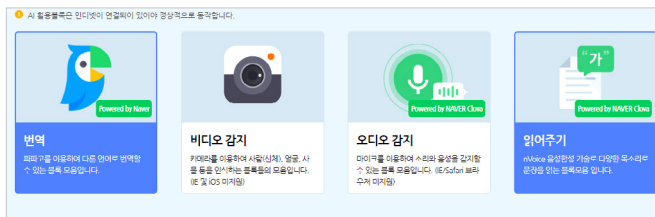
9) 이번에는 엔트리 AI 블록을 불러오도록 하겠습니다. 블록 탭에서 'AI 블록 불러오기'를 클릭해 주세요.

나만의 번역기 만들기 (1)

수업용 ppt 46~50p

1차시 | 2차시 | 3차시 | 4차시

<STEP 2> '번역', '읽어주기' 블록 추가하고, 외국인 오브젝트 블록 조립하기



10) '번역'과 '읽어주기' 블록을 선택하고 추가하기를 클릭하세요.



11) 이제 블록을 조립해보도록 하겠습니다. 먼저 **외국인 오브젝트** 블록을 조립하도록 하겠습니다. 시작하기 버튼을 클릭하면 프로그램이 작동할 수 있도록 '시작하기 버튼을 클릭했을 때' 블록을 사용합니다.



12) 리스트의 내용을 숨기기 위해 프로그램이 '리스트 숨기기' 블록을 사용합니다.

13) 7개의 리스트 항목 중에서 임의의 내용을 선택할 수 있도록 다음과 같이 블록을 조립합니다.



14) 임의로 선택한 리스트 항목을 '외국어' 변수에 저장될 수 있도록 블록을 조립합니다.



나만의 번역기 만들기 (1)

수업용 ppt 51~54p

1차시 | 2차시 | 3차시 | 4차시

15) '외국어' 변수에 저장한 외국어 내용을 외국인 오브젝트가 말할 수 있도록 말하기 블록을 추가해 조립합니다.



16) 학생 오브젝트가 작동될 수 있도록 우리가 만들어 놓은 신호를 이용해 신호를 보냅니다.

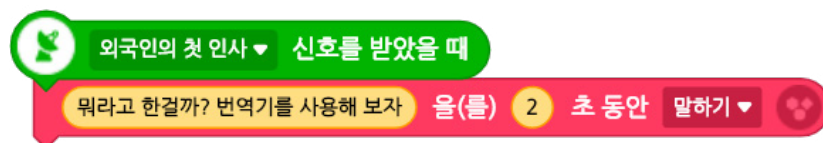


<STEP3> 외국인의 첫인사 신호를 받으면 작동하는 학생 오브젝트 블록 조립하기



17) 이제 **학생 오브젝트 블록**을 조립해보겠습니다. 외국인의 첫인사 신호를 받으면 작동할 수 있도록 다음 블록을 사용합니다.

18) '뭐라고 한걸까? 번역기를 사용해 보자'라고 2초동안 말하도록 블록을 조립합니다.



나만의 번역기 만들기 (1)

수업용 ppt 55~58p

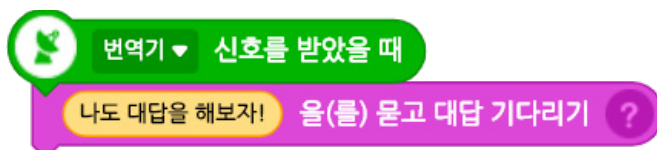
1차시 | 2차시 | 3차시 | 4차시

19) '번역기를 눌러주세요'라고 2초간 말하도록 블록을 조립합니다.

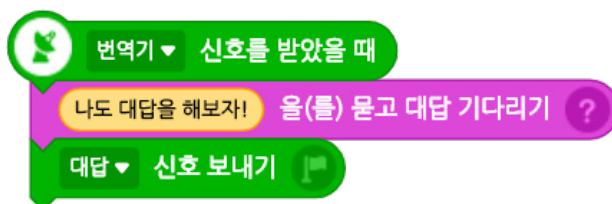


20) 이후에 번역기에서 번역이 끝나면 학생 오브젝트가 다시 작동할 수 있도록 블록을 조립하겠습니다. 먼저 번역 '번역기 신호를 받았을 때' 블록을 준비합니다.

21) 외국인의 말에 대답을 할 수 있도록 '묻고 대답 기다리기' 블록을 사용합니다. 이 블록을 이용하면 내가 하고 싶은 말을 작성할 수 있습니다.



22) 외국인 오브젝트에게 하고 싶은 말을 작성하면 번역기 오브젝트에 대답 신호를 보냅니다.



나만의 번역기 만들기 (1)

수업용 ppt 59~62p

1차시 | 2차시 | 3차시 | 4차시

<STEP4> 번역기 오브젝트 블록 조립하기 (1)



오브젝트를 클릭했을 때

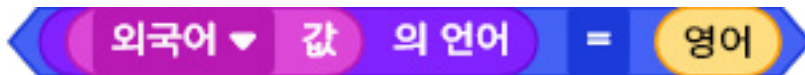
23) 이제 **번역기 오브젝트** 블록을 조립하겠습니다. 먼저 학생 오브젝트가 '번역기를 눌러주세요'라고 말한 뒤 번역기 오브젝트를 클릭하면 번역기 오브젝트가 작동되도록 구성하도록 하겠습니다. '오브젝트를 클릭했을 때' 블록을 준비해주세요.



엔트리 의 언어

24) 이 블록은 '엔트리'라고 되어 있는 부분에 문자를 입력하면 이 문자가 어느 나라 언어인지 감지해주는 블록입니다. 이 블록을 이용해 외국인 오브젝트가 어느 나라 말을 했는지 감지하고 그 언어에 맞게 번역을 하도록 구성하겠습니다.

25) 다음과 같이 블록을 조립하면 변수에 저장된 외국어가 영어인지 확인할 수 있습니다.



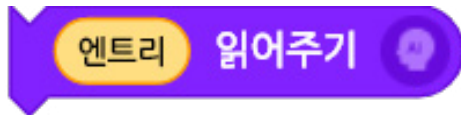
26) '만약 ~라면' 블록을 이용해 외국인 오브젝트가 한 말이 영어이면 동작할 수 있도록 블록을 조립합니다.



나만의 번역기 만들기 (1)

수업용 ppt 63~65p

1차시 | 2차시 | 3차시 | 4차시



27) 이 블록은 '엔트리'라고 되어 있는 부분에 문장을 넣으면 그 문자를 소리 내어 읽어주는 블록입니다. 이 블록을 이용해 번역기가 번역한 내용을 소리로 나타낼 수 있습니다.

28) 외국인 오브젝트가 한 말이 영어일 때 영어를 한국어로 번역해 주고 번역한 내용을 읽어줄 수 있도록 다음과 같이 블록을 조립합니다.



29) 같은 방법으로 감지한 외국어의 값이 일본어일 때 내용을 번역해 주고 번역한 내용을 읽어주는 블록을 조립합니다.



나만의 번역기 만들기 (1)

수업용 ppt 66~67p

1차시 | 2차시 | 3차시 | 4차시

<STEP5> 번역기 오브젝트 블록 조립하기 (2)

1 오브젝트를 클릭했을 때



30) 같은 방법으로 스페인어, 프랑스어, 독일어 블록도 조립합니다.

1 오브젝트를 클릭했을 때

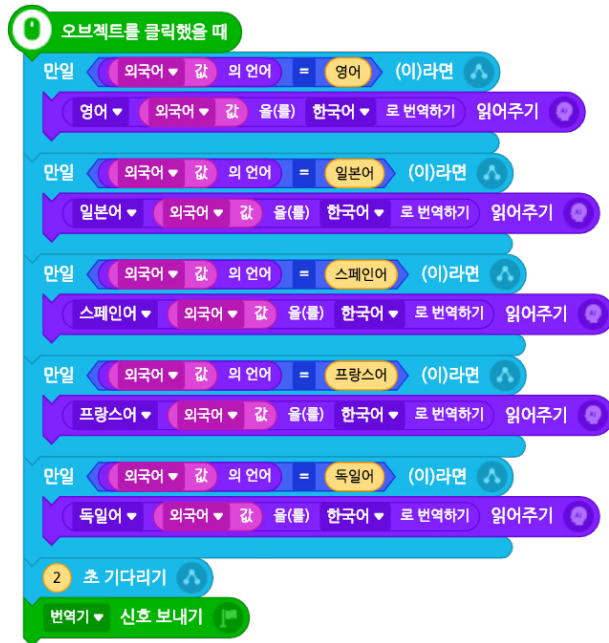


31) 원활한 동작을 위해 '2초 기다리기' 블록을 추가합니다.

나만의 번역기 만들기 (1)

수업용 ppt 68~70p

1차시 | 2차시 | 3차시 | 4차시



32) 번역이 끝나면 학생 오브젝트에게 번역기 신호를 보냅니다.

<STEP6> 학생 오브젝트가 대답을 입력하면 동작하도록 블록 추가하기



33) 학생 오브젝트가 대답을 입력하면 동작할 수 있도록 블록을 추가하겠습니다. '대답' 신호를 받으면 동작할 수 있도록 다음 블록을 이용합니다.

34) 이번엔 우리가 입력한 한국어를 외국인이 한 말의 언어로 번역할 수 있도록 블록을 조립합니다.



나만의 번역기 만들기 (1)

수업용 ppt 71~72p

1차시 | 2차시 | 3차시 | 4차시



35) 같은 방법으로 일본어, 스페인어, 프랑스어, 독일어 부분의 블록도 조립합니다.

36) '시작하기' 버튼을 클릭하여 프로그램이 잘 동작하는지 확인합니다.



나만의 번역기 만들기 (1)

수업용 ppt 73p

1차시 | **2차시** | 3차시 | 4차시

프로그램에 대해 이야기해봅시다.

Q. 이번 시간에 만든 작품은 엔트리의 어떤 인공지능 관련 기능을 사용하였나요?

Q. 작품을 만들면서 어떤 생각이 들었나요?

나만의 번역기 만들기 (2)

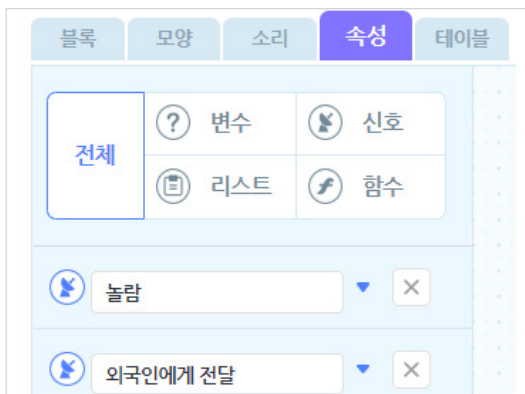
수업용 ppt 74~77p

1차시 | 2차시 | 3차시 | 4차시

2. 더 나아가기

<심화> '좋은 하루 보내세요'라는 말을 하도록 프로그램 발전시키기

오른쪽과 같이 '좋은 하루 보내세요'라는 말을 할 수 있도록 프로그램을 발전시켜봅시다.



1) 먼저 신호를 추가합니다. 여기서는 '놀람', '외국인에게 전달'로 신호의 이름을 정하였습니다.

나만의 번역기 만들기(2)

수업용 ppt 78~81p

1차시 | 2차시 | 3차시 | 4차시



2) 번역기 오브젝트가 번역을 마치면 '외국인에게 전달 신호'를 보낼 수 있도록 블록을 추가



3) 외국인 오브젝트가 '외국인에게 전달' 신호를 받으면 작동할 수 있도록 합니다.



4) 신호를 받으면 놀란 상태를 표현합니다.



5) 번역기 오브젝트에 놀람 신호를 보냅니다.

나만의 번역기 만들기(2)

수업용 ppt 82~84p

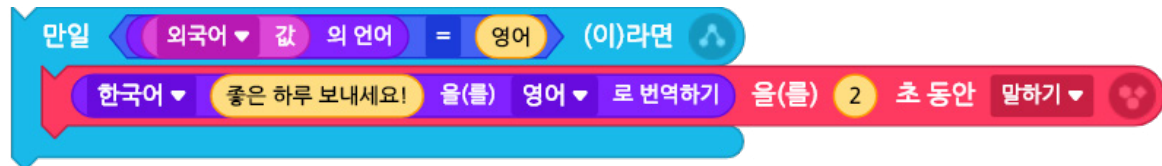
1차시 | 2차시 | 3차시 | 4차시



놀람 ▼ 신호를 받았을 때

6) 번역기 오브젝트가 놀람 신호를 받았을 때 동작하도록 다음 블록을 사용합니다.

7) 외국인이 말한 언어가 영어일 때 한국어 '좋은 하루 보내세요!'를 영어로 번역해주는 블록을 조립합니다.



8) 같은 방법으로 일본어, 스페인어, 프랑스어, 독일어 블록도 조립합니다.



나만의 번역기 만들기(2)

수업용 ppt 85~87p

1차시 | 2차시 | 3차시 | 4차시

9) '시작하기' 버튼을 클릭하여 프로그램이 잘 동작하는지 확인합니다.



더 발전시킬 부분이 있는지 찾아보고 자유롭게 프로그램을 변형해봅시다.

Q. 프로그램을 어떻게 바꿔보고 싶나요?

Q. 어떤 블록을 사용해야 하나요?

변형한 프로그램을 친구들과 서로 공유해봅시다.

Q. 가장 인상적인 프로그램은 무엇인가요?

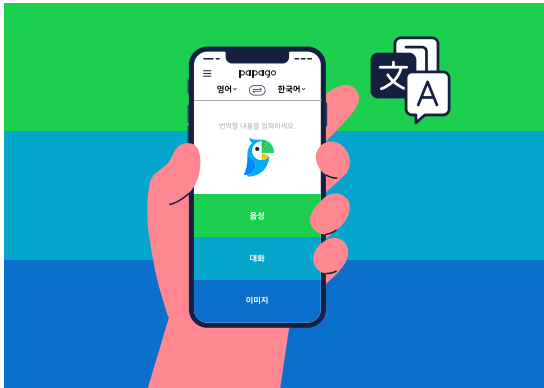
Q. 그 이유는 무엇인가요?

번역 및 음성 합성 인공지능이 우리 생활에 미치는 영향

수업용 ppt 88~92p

1차시 | 2차시 | 3차시 | 4차시

1. 번역 인공지능이 우리 생활에 미치는 영향



번역 인공지능 기술이 사용되는 사례를 살펴보고, 또 어떻게 쓰일 수 있을지 생각해봅시다.

영상 시청하기 :

<https://tv.naver.com/v/2628710>

Q. 번역 인공지능 기술을 어떻게 활용하고 있나요?

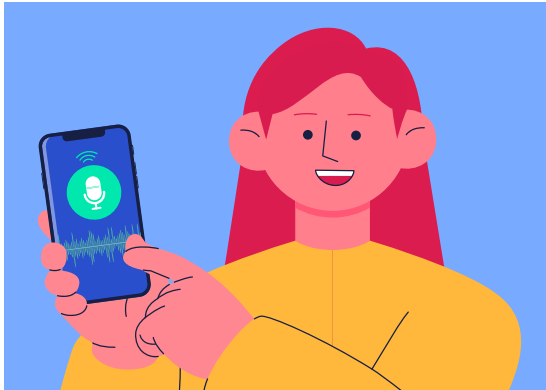
Q. 제시된 사례 이외에 번역 인공지능 기술을 어떻게 사용할 수 있을지 생각해 보고 만화로 나타내어봅시다.

번역 및 음성 합성 인공지능이 우리 생활에 미치는 영향

수업용 ppt 93~97p

1차시 | 2차시 | 3차시 | 4차시

2. 음성 합성 인공지능이 우리 생활에 미치는 영향



음성 합성 인공지능 기술이 사용되는 사례를 살펴보고,
또 어떻게 쓰일 수 있을지 생각해봅시다.

영상 시청하기 :

<https://tv.naver.com/v/10664305>

Q. 음성 합성 인공지능 기술을 어떻게 활용하고 있나요?

Q. 제시된 사례 외에 음성 합성 인공지능 기술을 어떻게 사용할 수 있을지 생각해봅시다.

Q. 음성 합성 인공지능 기술사용에 따른 부정적인 면은 무엇이 있을까요?

Q. 이러한 부정적인 면을 어떻게 예방할 수 있을까요?

번역 및 음성 합성 인공지능이 우리 생활에 미치는 영향

수업용 ppt 98p

1차시 | 2차시 | 3차시 | 4차시

이번 주제에서 배운 내용을 생각하며, 스스로 평가해봅시다.

컴퓨터가 텍스트를 인식하는 방법을 설명할 수 있다.	☆ ☆ ☆ ☆ ☆
음성 합성 인공지능에 대해 설명할 수 있다.	☆ ☆ ☆ ☆ ☆
엔트리 번역 기능을 이용하여 인공지능 번역 프로그램을 만들 수 있다.	☆ ☆ ☆ ☆ ☆
번역과 음성 합성 인공지능 기술이 가져올 수 있는 사회적 영향에 대해 알 수 있다.	☆ ☆ ☆ ☆ ☆

참고자료

- 아서 줄리아니. (2017). 강화학습 첫걸음. 한빛미디어.
- Bishop, C. M. (2006). Pattern recognition and machine learning. springer.
- 오가와 유타로(2018). PyTorch를 활용한 강화학습/ 심층강화학습 실전 입문. 위키북스.
- 김세훈. (2015). OCR 프로그래밍. 한빛미디어.
- 피터 플래치. (2016). 머신러닝-데이터를 이해하는 알고리즘의 예술과 과학. Bjpublic.
- 바라가브 스리니바사 디지칸. (2019). 자연어 처리와 컴퓨터 언어학. 에이콘.
- 김현철. (2014). 데이터로 표현하는 세상. 고려대학교 출판문화원.
- 김현철. (2019). 정보적 사고에서 인공지능까지. 한빛아카데미.
- 크리스슈나 바브사. (2019). 자연어 처리 쿡북 with 파이썬. 에이콘.
- 재클린 카질. (2017). 파이썬을 활용한 데이터 길들이기. 인사이트.
- 타니구치 타다히로(2017). 그림과 수식으로 배우는 통통 인공지능. 제이펍.
- 다다 사토시(2017). 처음 배우는 인공지능. 한빛미디어.
- 드미트리 지노비예프(2017).모두의 데이터 과학 with 파이썬.길벗
- Young, S. (1996). A review of large-vocabulary continuous-speech. IEEE signal processing magazine, 13(5), 45.
- Aubert, X. L. (2002). An overview of decoding techniques for large vocabulary continuous speech recognition. Computer Speech & Language, 16(1), 89-114.
- Chan, W., Jaitly, N., Le, Q., & Vinyals, O. (2016, March). Listen, attend and spell: A neural network for large vocabulary conversational speech recognition. In 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 4960-4964). IEEE.
- Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.
- Alpaydin, E. (2020). Introduction to machine learning. MIT press.
- Abdel-Hamid, O., Mohamed, A. R., Jiang, H., Deng, L., Penn, G., & Yu, D. (2014). Convolutional neural networks for speech recognition. IEEE/ACM Transactions on audio, speech, and language processing, 22(10), 1533-1545.
- Kubat, M. (2017). An introduction to machine learning. Springer International Publishing AG.
- Mitchell, T. M. (1999). Machine learning and data mining. Communications of the ACM, 42(11), 30-36.
- Luger, G. F. (2005). Artificial intelligence: structures and strategies for complex problem solving. Pearson education.
- Szeliski, R. (2010). Computer vision: algorithms and applications. Springer Science & Business Media.
- Badue, C., Guidolini, R., Carneiro, R. V., Azevedo, P., Cardoso, V. B., Forechi, A., ... & de Paula Veronese, L. (2020). Self-driving cars: A survey. Expert Systems with Applications, 113816.
- Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.
- Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., ... & Zhang, X. (2016). End to end learning for self-driving cars. arXiv preprint arXiv:1604.07316

소프트웨어야 놀자 AI 교육 시리즈
프로젝트 - 5

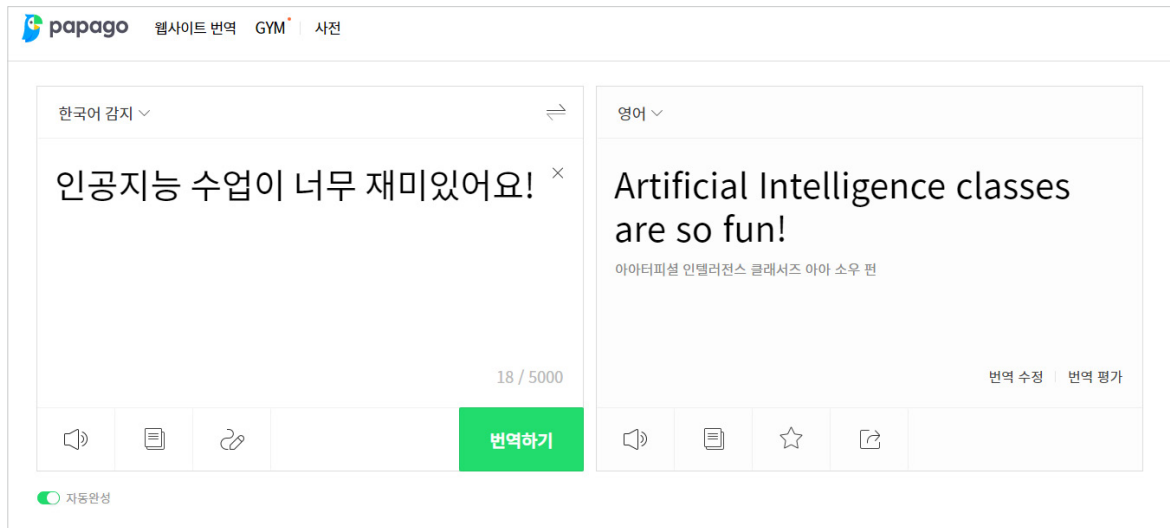
나만의 번역기 만들기

학생용 워크시트

학생용 워크시트 - 5. 나만의 번역기 만들기

체험해볼까요?

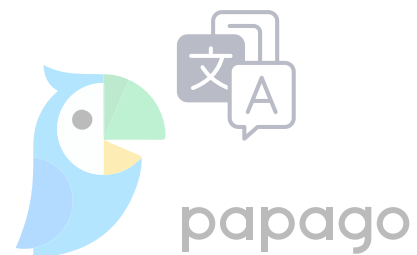
* 네이버 번역기 파파고



파파고 : <https://papago.naver.com>

▶ 네이버 번역기 파파고를 사용해본 적이 있나요? 파파고는 여러 나라의 언어를 다른 언어로 번역해 주는 번역 프로그램입니다. 스마트 폰 앱으로도 제작되어 있어 편리하게 사용할 수 도 있습니다. 파파고는 현재 15가지 언어를 번역할 수 있고 웹사이트 페이지 번역도 가능합니다. 그리고 필요할 때는 번역한 내용을 음성으로 나타내주는 기능도 있습니다. 네이버 번역기 파파고의 기능을 직접 체험해 봅시다.

▶ 어떤 인공지능 기술이 들어있을까요? 자유롭게 이야기해 봅시다.



학생용 워크시트 - 5. 나만의 번역기 만들기

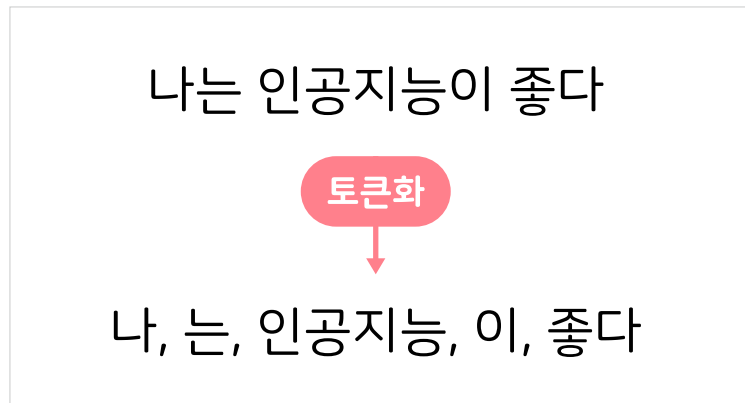
원리를 알아볼까요?

1. 텍스트 데이터 처리

컴퓨터가 텍스트를 인식하고 다른 나라 언어로 번역하기 위해서는 우선 컴퓨터가 이해할 수 있도록 텍스트 데이터를 숫자로 바꾸는 처리 과정이 필요합니다. 그렇다면 컴퓨터는 어떻게 텍스트를 처리할 수 있을까요?

1) 토큰화

토큰(Token)은 텍스트 데이터에서 의미를 가지는 최소한의 덩어리를 말합니다. 그리고 텍스트 데이터를 토큰 단위로 나누는 과정을 토큰화라고 합니다.



'나는 인공지능이 좋다.'라는 예시문을 토큰화하면 '나', '는', '인공지능', '이', '좋다'의 토큰으로 나눌 수 있습니다. 여기서는 조사도 구분하여 토큰으로 나눌 것을 확인할 수 있습니다.

학생용 워크시트 - 5. 나만의 번역기 만들기

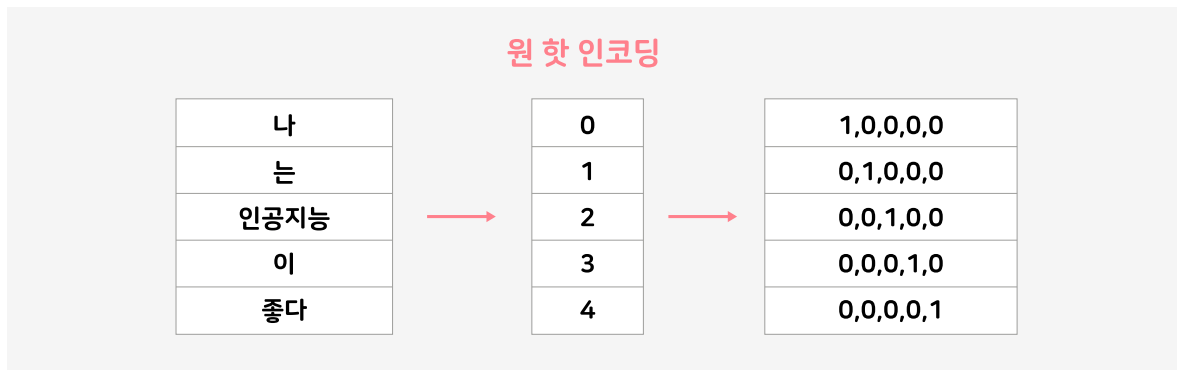
원리를 알아볼까요?

1. 텍스트 데이터 처리

2) 임베딩

토큰화가 끝나면 이제 임베딩의 단계를 거칩니다. 임베딩이란 하나의 단어를 숫자로 바꾸는 과정으로 여기서는 토큰화된 단어들을 컴퓨터가 이해할 수 있는 숫자로 변환시키는 과정을 말합니다.

우리는 그동안 컴퓨터가 이미지 데이터와 소리 데이터를 컴퓨터가 이해할 수 있는 숫자 0과 1로 바꾸는 과정에 대해 배웠습니다. 텍스트 데이터도 이와 마찬가지로 0과 1로 바꾸는 과정이 필요합니다. 임베딩하는 방법은 다양한데 이번 시간에는 그중 하나인 원 핫 인코딩에 대해 알아보도록 하겠습니다.



먼저 처리하는 텍스트의 토큰을 모읍니다. 이렇게 모인 토큰의 집합을 '단어 사전'이라고 표현합니다. 단어 사전의 토큰들은 서로 중복이 되지 않게 구성되고 단어 사전의 크기는 사전에 들어 있는 토큰의 개수입니다. 예시문의 단어 사전에는 '나', '는', '인공지능', '이', '좋다'가 들어있습니다. 따라서 이 단어 사전의 크기는 5입니다.

이제 단어 사전에 있는 토큰에 고유한 숫자를 부여합니다. 이때의 숫자는 0부터 시작합니다.

숫자를 정한 후 이제 해당하는 위치에 숫자 1을 입력합니다. 그리고 나머지 부분은 0을 입력합니다. 예시문의 '인공지능' 토큰은 이제 (0,0,1,0,0) 이라는 숫자로 변경된 것을 확인할 수 있습니다.

다음 보기의 문장을 원 핫 인코딩해 봅시다.

<보기>

희석이는 햄버거를 너무 좋아해서 매일 햄버거를 먹습니다.

Q. 보기의 문장을 토큰화해보세요.

Q. 토큰들을 모아 단어 사전을 만들어보세요.

Q. 원 핫 인코딩으로 토큰들을 숫자로 바꾸어보세요.

학생용 워크시트 - 5. 나만의 번역기 만들기

원리를 알아볼까요?

2. 자연어 처리

텍스트 데이터를 숫자로 바꾸었다고 해서 컴퓨터가 이 데이터가 어떤 내용을 담고 있는지 인식하지는 못합니다. 그저 컴퓨터가 알아볼 수 있는 형태로 텍스트를 바꾼 것이고 여기까지는 인공지능의 영역이 아닙니다. 컴퓨터가 텍스트를 인식한다는 것은 텍스트 속에서 특징(feature)을 추출하고 그것을 분석해서 그 텍스트가 어떤 의미인지 인식할 수 있어야 합니다.

텍스트 데이터에서 특징을 추출하는 것은 매우 어려운 일입니다. '약이 쓰다'와 '글을 쓰다' 두 문장을 보면 공통적으로 '쓰다'라는 단어를 사용하고 있습니다. 하지만 그 의미는 다르게 사용되고 있습니다. 우리는 문맥이나 상황을 고려해 이 텍스트가 어떤 의미를 담고 있는지 쉽게 알 수 있습니다. 하지만 컴퓨터가 이런 것들을 모두 특징들을 고려하는 것은 매우 힘든 일입니다. 최근의 인공지능 기술(딥러닝)이 발전하면서 인공지능은 이러한 텍스트 데이터에서 특징을 스스로 찾아내고 조합하여 분석해 텍스트 데이터를 인식할 수 있게 되었습니다.

이렇게 컴퓨터가 인간의 언어를 인식하고 목적에 맞게 처리할 수 있도록 하는 것을 자연어 처리(Natural Language Processing, NLP)라고 합니다. 자연어 처리는 인공지능의 한 분야로 여기서 자연어는 우리가 일상생활에서 말하고 듣고 읽고 쓰기를 하는 언어를 말합니다. 따라서 자연어 처리를 이해하기 위해서는 인공지능에 대한 이해뿐만 아니라 언어에 대한 이해도 가지고 있어야 합니다.



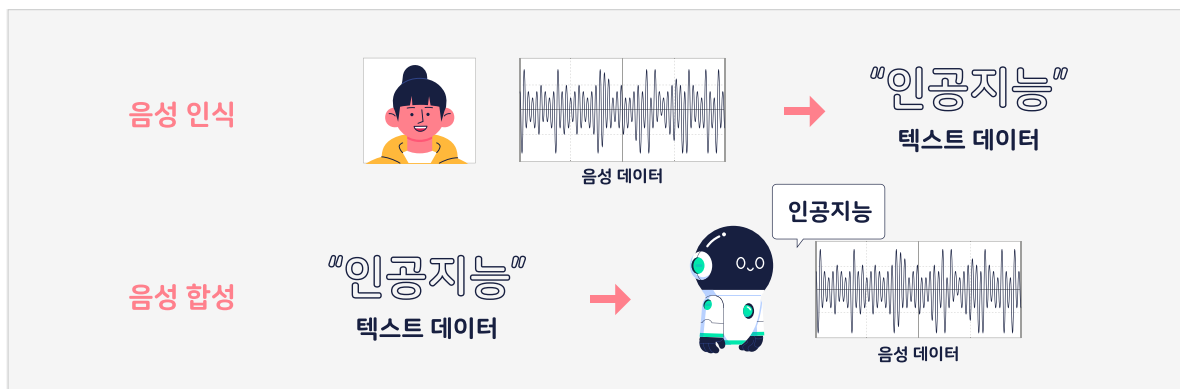
자연어 처리에는 다양한 분야가 있습니다. 먼저 텍스트를 인식하여 이 텍스트가 긍정적인 내용인지 부정적인 내용인지 분류해내는 것을 감정 분석이라고 합니다. 이러한 감정 분석은 고객들의 리뷰를 관리하는 데 자주 사용됩니다. 또한, 자연어 처리는 문서를 자동으로 요약해주는 데 사용되기도 하고, 오류를 자동으로 수정해 주는 데에도 사용됩니다. 컴퓨터를 이용해서 한 언어에서 다른 언어로 바꾸는 번역도 자연어 처리의 한 분야입니다.

학생용 워크시트 - 5. 나만의 번역기 만들기

원리를 알아볼까요?

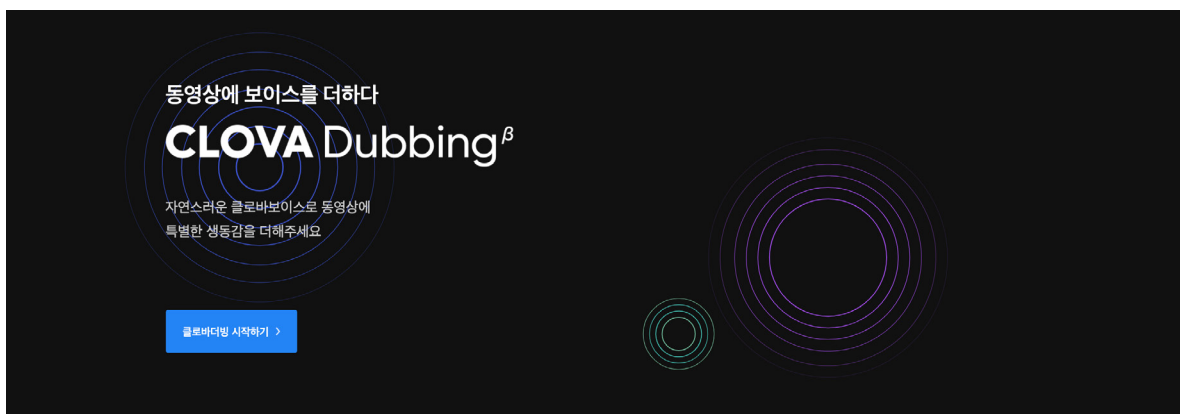
3. 음성 합성

음성 합성이란 텍스트를 소리로 만들어 내는 기술입니다. 지난 시간에 배운 음성 인식이 사람의 목소리를 텍스트로 변환하였다면 음성 합성은 반대로 텍스트를 사람의 목소리로 바꾸는 것입니다.



사람이 직접 녹음을 할 때는 말의 내용이 바뀔 때마다 새로 녹음해야 하는 문제가 있었습니다. 음성 합성 기술이 발달함에 따라 사람이 일일이 녹음을 할 필요가 없어졌고 기계가 내는 목소리가 더욱더 자연스러워지게 되었습니다. 현재에는 자동 응답 서비스, 버스나 지하철 안내방송, 책이나 뉴스 기사 읽어주기 등등 다양한 분야에서 사용되고 있습니다.

음성 합성 기술을 이용해 만들어진 네이버의 클로바 더빙을 체험해 봅시다.



클로바 더빙 체험하기 : <https://clovadubbing.naver.com>

학생용 워크시트 - 5. 나만의 번역기 만들기

원리를 알아볼까요?

※ 학습한 내용을 정리해봅시다.

Q. 컴퓨터는 텍스트를 어떻게 인식할까요?

텍스트를 (), ()의 숫자 형태로 처리한 뒤, 데이터에서 ()을 추출하여 어떤 텍스트인지 인식합니다.

Q. 컴퓨터가 인간의 언어를 인식하고 목적에 맞게 처리할 수 있도록 하는 것을 ()라고 합니다.

Q. 텍스트 데이터를 사람의 목소리로 바꾸는 인공지능 기술을 ()이라고 합니다.

학생용 워크시트 - 5. 나만의 번역기 만들기

만들어볼까요?

지금부터 엔트리의 번역 기능과 읽어주기 기능을 사용하여 나만의 번역기를 만들어보겠습니다.

1) 엔트리 공유주소 <http://naver.me/56lKE8FC> 에 접속한 후, 코드보기를 클릭합니다.

2) 엔트리 실습 링크 <https://blog.naver.com/nvplaysw/222108959770> 를 열고 실습해 봅시다.



- step.1 3가지 신호를 만들고, '외국어' 변수 추가 하기
- step.2 '번역', '읽어주기' 블록 추가하고, 외국인 오브젝트 블록 조립하기
- step.3 외국인의 첫인사 신호를 받으면 작동하는 학생 오브젝트 블록 조립하기
- step.4 번역기 오브젝트 블록 조립하기 (1)
- step.5 번역기 오브젝트 블록 조립하기 (2)
- step.6 학생 오브젝트가 대답을 입력하면 동작하도록 블록 추가하기

3) 프로그램에 관해 이야기해봅시다.

Q. 이번 시간에 만든 작품은 엔트리의 어떤 인공지능 관련 기능을 사용하였나요?

Q. 작품을 만들면서 어떤 생각이 들었나요?

학생용 워크시트 - 5. 나만의 번역기 만들기

발전시켜볼까요?

<심화> '좋은 하루 보내세요'라는 말을 하도록 프로그램 발전시키기

1) 오른쪽과 같이 '좋은 하루 보내세요'라는 말을 할 수 있도록 프로그램을 발전시켜봅시다.



2) 더 발전시킬 부분이 있는지 찾아보고, 자유롭게 프로그램을 변형해봅시다.

Q. 프로그램을 어떻게 바꿔보고 싶나요?

Q. 어떤 블록을 사용해야 하나요?

3) 설계한대로 프로그램을 만들어 봅시다.

4) 변형한 프로그램을 친구들과 서로 공유해봅시다.

Q. 가장 인상적인 프로그램은 무엇인가요?

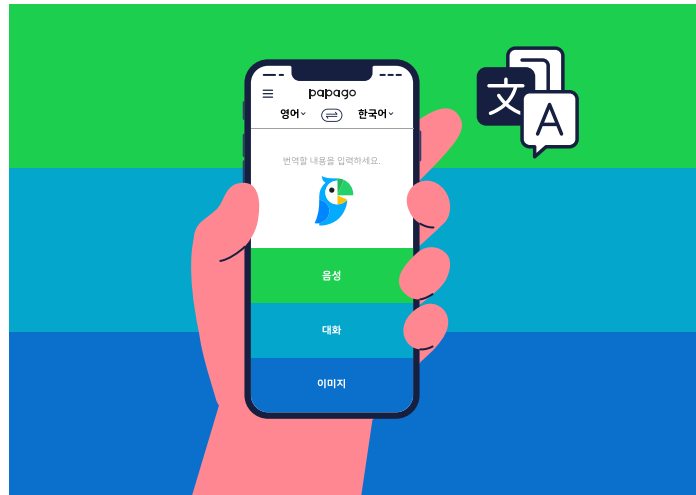
Q. 그 이유는 무엇인가요?

학생용 워크시트 - 5. 나만의 번역기 만들기

생각해볼까요?

1. 번역 인공지능이 우리 생활에 미치는 영향

번역 인공지능 기술이 사용되는 사례를 살펴보고, 또 어떻게 쓰일 수 있을지 생각해봅시다.



영상 시청하기 : <https://tv.naver.com/v/2628710>

Q. 번역 인공지능 기술을 어떻게 활용하고 있나요?

학생용 워크시트 - 5. 나만의 번역기 만들기

생각해볼까요?

1. 번역 인공지능이 우리 생활에 미치는 영향

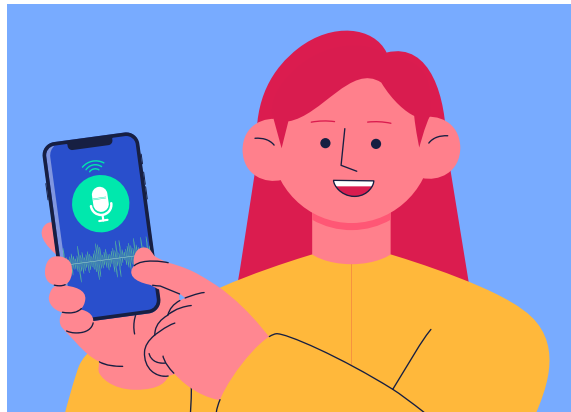
Q. 제시된 사례 이외에 번역 인공지능 기술을 어떻게 사용할 수 있을지 생각해보고 만화로 나타내어 봅시다.

학생용 워크시트 - 5. 나만의 번역기 만들기

생각해볼까요?

2. 음성 합성 인공지능이 우리 생활에 미치는 영향

음성 합성 인공지능 기술이 사용되는 사례를 살펴보고, 또 어떻게 쓰일 수 있을지 생각해봅시다.



영상 시청하기 : <https://tv.naver.com/v/10664305>

Q. 음성 합성 인공지능 기술을 어떻게 활용하고 있나요?

Q. 제시된 사례 외에 음성 합성 인공지능 기술을 어떻게 사용할 수 있을지 생각해 봅시다.

Q. 음성 합성 인공지능 기술사용에 따른 부정적인 면은 무엇이 있을까요?

Q. 이러한 부정적인 면을 어떻게 예방할 수 있을까요?

학생용 워크시트 - 5. 나만의 번역기 만들기

★ 셀프 평가

이번 주제에서 배운 내용을 생각하며, 스스로 평가해봅시다.

컴퓨터가 텍스트를 인식하는 방법을 설명할 수 있다.	☆ ☆ ☆ ☆ ☆
음성 합성 인공지능에 대해 설명할 수 있다.	☆ ☆ ☆ ☆ ☆
엔트리 번역 기능을 이용하여 인공지능 번역 프로그램을 만들 수 있다.	☆ ☆ ☆ ☆ ☆
번역과 음성 합성 인공지능 기술이 가져올 수 있는 사회적 영향에 대해 알 수 있다.	☆ ☆ ☆ ☆ ☆

참고사이트

소프트웨어야 놀자 : <https://www.playsw.or.kr>

엔트리 : <https://playentry.org>

Hello AI World(인공지능 개념 영상) : https://playsw.or.kr/artificial/video?currentTab=ai_data_video_hello_ai