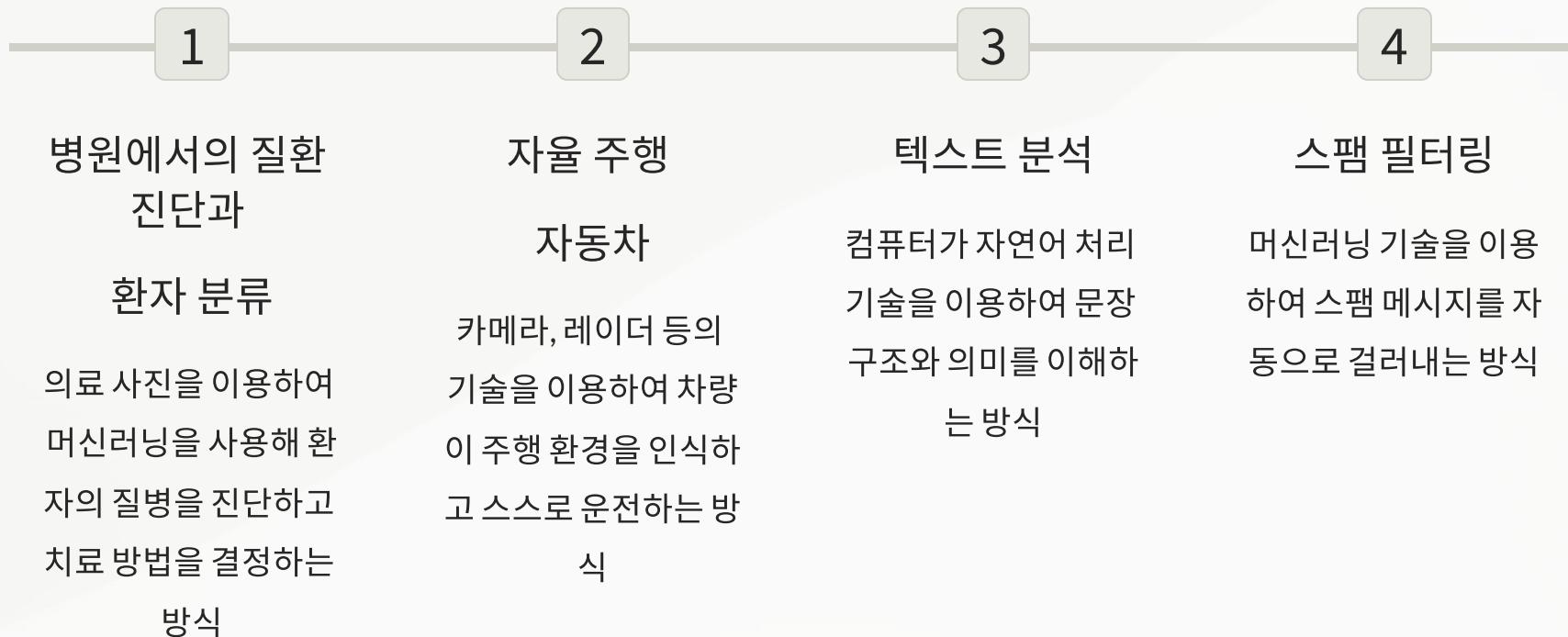


딥러닝과 캐글의 기본 용어 이해

인공지능 기술에 대해 궁금하신가요? 이번 발표에서는 구글 코랩, 머신러닝, 딥러닝, 데이터셋, 그리고 캐글 웹 사이트에 대해 알아보겠습니다.

by 5분수업 토돌이의

머신러닝과 딥러닝의 활용 사례1



머신러닝과 딥러닝의 활용 사례2



로봇

로봇의 동작을 제어하기 위해 머신러닝과 딥러닝을 사용한다. 머신러닝 알고리즘을 이용하여 로봇의 동작을 학습, 환경 변화에 적응하도록 한다.

화장품, 옷

머신러닝과 딥러닝을 사용하여 소비자의 취향과 선호도를 이해한다. 예를 들면 고객의 얼굴 이미지를 분석하여 피부 상태를 판단하여 이에 맞는 화장품을 추천한다.

반도체

머신러닝과 딥러닝을 사용하여 공정 제어, 품질 검사, 고장 진단 등 다양한 작업에 활용된다. 예를 들어 공정 데이터를 수집하여 머신러닝 모델을 학습시켜 생산 공정을 최적화하거나 결함을 예측하여 예방할 수 있다.

게임

인공지능 캐릭터를 개선하기 위해 머신러닝과 딥러닝을 사용. 머신러닝과 딥러닝을 사용하여 게임 플레이를 개선하기도 한다. 또한 GAN 등의 활용하여 게임 캐릭터, 맵, 아이템 등을 자동으로 생성 할 수 있다.

데이터셋 이해 - IRIS 데이터셋

IRIS 데이터셋

붓꽃(IRIS)의 꽃잎과 꽃 받침의 폭과 길이에 대한 측정 값으로부터 어떤 꽃인지 예측하는 머신러닝 문제

데이터

총 150개의 데이터 행이 있다.
붓꽃의 세가지 종류(세토사, 버시컬러, 버지니카)를 가지고 있다.

특징

데이터의 수가 적지만, 분류하기에 충분한 양이며, 머신러닝을 처음 적용하는 사람들이 사용하기 적합한 데이터셋

데이터셋 이해 - 유방암 데이터셋



유방암 데이터셋

위험 요소와 유방 조직의 세포학 차이 등을 이용하여 악성 종양(암)과 양성 종양(암이 아님)을 구분하는 머신러닝 문제



적용 분야

머신러닝 및 데이터 마이닝 분야에서 유방암 진단 및 분류 문제를 다루는데 널리 사용됨.



데이터셋

데이터셋은 569개의 행이 있다. 양성과 악성은 0,1로 표현되어, 0은 양성 종양, 1은 악성 종양을 나타낸다.

데이터 셋 이해 - 타이타닉 대회

생존 여부	성별	나이	객실 등급
0 - 생존하지 못한	남성	0.33세 - 80세까지	1등급 - 3등급
1 - 생존한	여성		
승선 항구	형제자매 및 배우자 수	부모자식 수	운임
Cherbourg	0명 - 8명까지 포함된 경우	0명 - 6명까지 포함된 경우	산정된 운임
Queenstown			
Southampton			

타이타닉 호에 탑승한 사람들에 대한 정보로 이루어진 데이터 셋. 해당 대회에서는 객실 등급, 성별, 나이 등과 같은 정보로 승객의 생존 여부를 예측하는 머신러닝 모델을 개발하는 것이 주된 목적.

캐글 웹 사이트와 경진 대회의 장점

캐글 웹 사이트

캐글은 데이터 분석을 주제로 한 커뮤니티 사이트로, 데이터셋을 제공하고 경진 대회를 개최한다.

실제 데이터 활용

실제 데이터를 다루는 것에서 더 많은 경험을 할 수 있고, 자신의 능력을 보여줄 수 있다.

경력 구축

캐글에 참가하고 우수한 결과를 보여주면, 이를 포트폴리오에 추가할 수 있어 회사에 입사하거나 프리랜서로 일할 때 큰 도움이 된다.



결론 및 요약

인공지능의 발달로 머신러닝과 딥러닝 기술이 널리 활용되고 있다. 이번 발표에서는 구글 코랩, 머신러닝과 딥러닝의 개념, 데이터셋, 캐글 웹 사이트, 그리고 경진 대회에 대해 알아보았다. 유용한 데이터와 머신러닝/딥러닝 기술을 활용하여 여러분들도 다양한 분야에서 진보적인 연구나 새로운 아이디어를 실현시키기를 기원한다.

질의응답

이제 질문 시간입니다! 적극적으로 참여해 주세요.