

1 <처음보는 용어집>

1. 원-핫 인코딩(One-Hot Encoding)

원-핫 인코딩은 단어 집합의 크기를 벡터의 차원으로 하고, 표현하고 싶은 단어의 인덱스에 1의 값을 부여하고, 다른 인덱스에는 0을 부여하는 단어의 벡터 표현 방식. 이렇게 표현된 벡터를 원-핫 벡터(One-Hot vector)

원-핫 인코딩을 수행하는 두 가지 과정.

첫째, 정수 인코딩을 수행. 다시 말해 각 단어에 고유한 정수를 부여.

둘째, 표현하고 싶은 단어의 고유한 정수를 인덱스로 간주하고 해당 위치에 1을 부여하고, 다른 단어의 인덱스의 위치에는 0을 부여.

2. Bag of Words

Bag of Words란 단어들의 순서는 전혀 고려하지 않고, 단어들의 출현 빈도(frequency)에만 집중하는 텍스트 데이터의 수치화 표현 방법

3. 가지치기 기법(PRUNING)

참고자료 : https://tutorials.pytorch.kr/intermediate/pruning_tutorial.html

가지치기 기법은 굉장히 많은 수의 파라미터값들로 구성된 모델과 굉장히 적은 수의 파라미터값들로 구성된 모델 간 학습 역학 차이를 조사하는데 주로 이용되기도 하며, 하위 신경망 모델과 파라미터값들의 초기화가 운이 좋게 잘 된 케이스를 바탕으로 (《lottery tickets》) 신경망 구조를 찾는 기술들에 대해 반대 의견을 제시하기도 합니다.

4. Quantization

참고자료 : <https://pytorch.org/docs/stable/quantization.html>

양자화는 부동 소수점 정밀도보다 낮은 비트폭으로 계산을 수행하고 텐서를 저장하는 기술을 말합니다. 양자화 모델은 전체 정밀도(부동 소수점) 값이 아닌 감소된 정밀도로 텐서에 대한 일부 또는 모든 작업을 실행합니다.

5. API(application programming interface) : 일종의 소프트웨어 인터페이스로, 다른 종류의 소프트웨어에 서비스를 제공한다. 즉 컴퓨터나 소프트웨어를 연결해 프로그래머가 사용할 수 있는 도구나 서비스의 역할을 한다. 한가지 목적은

시스템이 동작하는 방식에 대한 내부의 세세한 부분을 숨기는 것으로, 내부의 세세한 부분이 나중에 변경되더라도 프로그래머가 유용하게 사용할 수 있고 일정하게 관리할 수 있는 부분들만 노출시킨다.

6. Multi modal learning : 문제의 실제 현상에 대한 보다 강력한 모델을 만들기 위해 텍스트, 오디오 또는 이미지와 같은 다양한 양식의 데이터를 조합하여 사용하는 일종의 딥러닝.
7. Stemming : 형태론 및 정보 검색 분야에서 어형이 변형된 단어로부터 접사 등을 제거하고 그 단어의 어간을 분리해 내는 것을 의미한다. 이때 어간은 반드시 어근과 같아야 할 필요는 없으며, 어근과 차이가 있더라도 관련이 있는 단어들이 일정하게 동일한 어간으로 맵핑되게 하는 것이 어간 추출의 목적이라고 한다.
8. Tokenization : 정의된 범주에 속하는 의미 있는 어휘 토큰으로 변환하는 것으로, 자연어의 경우, 그 범주에는 명사, 동사, 형용사, 구두점 등이 포함되고, 프로그래밍 언어의 경우, 범주에는 식별자, 연산자, 그룹화 기호 및 데이터 유형이 포함된다. 더 나아가 어휘 토큰화는 일반적으로 어휘 문법을 기반으로 하는 반면, LLM 토큰라이저는 일반적으로 확률 기반이기에 토큰을 숫자 값으로 변환하는 두번째 단계를 수행한다.
9. Attention modules : 인간의 주의집중을 모방하여 중요한 입력 부분을 다시 참고하는 기계학습 기법으로, 트랜스포머 모델에 사용된다. Doc-product-attention, multi-head attention 기법에 많이 사용된다. 이 기법이 적용된 모델은 입력 데이터의 특정 부분에 큰 가중치를 두는 방식으로 주의를 집중한다. 경사 하강법으로 학습되며, 데이터의 어느 부분이 다른 부분보다 더 중요한지 학습하는 것은 상황에 따라 다르다. (출처: 위키피디아)
10. GAN(generative adversarial networks): 우리말로로는 '적대적 생성 신경망'이라고 번역할 수 있는 AI 기술이다. GAN은 실제에 가까운 이미지나 사람이 쓴 것 같은 글 등 여러 가짜 데이터들을 생성하는 모델이다. 이름에서 알 수 있듯 GAN은 서로 다른 두 개의 네트워크를 적대적으로 학습시키며 실제 데이터와 비슷한 데이터를 생성해내는 모델이다. GAN을 통해 생성된 데이터에는 정해진 label값이 없기

때문에 비지도 학습 기반 생성모델로 분류된다. 이 기술은 Ian Goodfellow에 의해 2014년에 처음으로 신경정보처리시스템학회에서 제안되었으며, 이후 이미지 생성, 영상 생성, 텍스트 생성 등 다양하게 응용되고 있다. 출처: <https://pseudo-lab.github.io/Tutorial-Book/chapters/GAN/Ch1-Introduction.html>

11. 지도학습: 지도 학습 알고리즘은 주어진 예시를 통해 학습되도록 하는 기계학습 방법. 비지도 학습보다 더 단순하고 일반적이며, 사람인 관여자가 답을 알고 있는 상태에서 AI를 훈련하고자 할 때 사용하고, 사용된 데이터를 분류하고 처리하며, 레이블이 지정된 데이터, 즉 분류된 데이터를 사용한다는 특징을 가지고 있다. 출처: <https://www.appier.com/ko-kr/blog/what-is-supervised-learning>
12. 컴퓨터 비전 : 이미지 처리, 기계 학습 및 AI를 병합하여 컴퓨터가 우리 주변 세계의 시각적 정보를 디지털 영상 분석을 통해 처리, 분석, 감지, 식별, 추적 및 이해할 수 있도록 하는 분야. 이를 통해 휴먼 컴퓨팅 인터페이스, 도로 안전, 의료, 산업 영상 검사 등 다양한 분야에서 활용할 수 있다. 컴퓨터 비전의 핵심은 영상 처리 기술과 딥러닝 등 인공지능 기반 알고리즘을 활용해 컴퓨터, 즉 기계에게 '현실을 보고 이에 대처할 수 있는 능력'을 부여하는 것이다. 출처: <https://soreblog.co.kr/%EC%BB%B4%ED%93%A8%ED%84%B0-%EB%B9%84%EC%A0%84-%ED%99%9C%EC%9A%A9-%EC%82%AC%EB%A1%80-%EC%86%8C%EA%B0%9C/>
13. 자연어 처리(NLP, natural language processing) : 우리가 일상 생활에서 사용하는 언어의 의미를 컴퓨터에 그대로 입력했을 때, 이러한 자연어의 의미를 분석하여 컴퓨터가 처리할 수 있도록 만들어주는 기능. 즉, 컴퓨터와 사람들이 인간의 언어를 사용해 상호작용할 수 있는 방법에 초점을 맞춘 인공지능을 말한다. 자연어 처리는 1950년도부터 시작되었으며, 초기에는 언어 문법을 그대로 컴퓨터에 적용하였으나, 1980년대 이후로는 통계 기반 접근 방식을 통해 언어의 패턴을 학습시켰고, 2010년대 이후로는 딥러닝을 통해 복잡한 인간의 언어를 빠르게 이해시킬 수 있었다. 출처: <https://www.elastic.co/kr/what-is/natural-language-processing>

문제 2. 인공지능은 크게 컴퓨터 비전, 자연어 처리 두 분야로 나뉜다. 해당 분야들의 주 기능과 어떻게 쓰이고 있는지 조사해보고, 두 분야를 응용해서 해결할 수 있는 사회 문제나 서비스에 대해 상상해보자. [주제 : 인공지능 - 토의 유도]

[작성시 유의사항]

- 컴퓨터비전과 자연어처리에 대한 기본 개념, 주 기능을 정확히 조사하였는지
- 인공지능의 컴퓨터비전, 자연어처리를 사용해서 문제를 해결했던 사례에 대해 조사하고 분석했는지
- 조사한 내용 외에 적용해볼 수 있는 인공지능 서비스를 상상하고, 구체적으로 묘사했는지

(1)컴퓨터 비전: 컴퓨터 비전(Computer Vision)은 컴퓨터를 이용해 정지 영상 또는 동영상으로부터 의미 있는 정보를 추출하는 방법을 연구하는 컴퓨터 과학의 연구 분야 중 하나다.

컴퓨터 비전의 구체적인 과제로는 객체 분류(Object Classification), 객체 지역화(Object Localization), 객체 탐지(Object detection), 이미지 분할(Image Segmentation) 등이 있다. 현재 가장 많이 활용되는 프로그래밍 라이브러리로는 Open CV가 있다.

출처: [네이버 지식백과] [컴퓨터 비전](#) (AI 용어사전)

(2)컴퓨터 비전 활용 분야

[1]머신비전:

컴퓨터 비전을 활용하여 제품의 결함 여부 실시간으로 식별. 생산 라인에 전용 카메라를 설치하여 완성된 부품 이미지를 촬영하여 정상 제품 이미지와 비교하는 방식으로 결함을 확인. 이 식별 결과를 바탕으로 유통 업체는 재고 관리를 개선.

[2]리테일 분야 고객 분석:

소매 업체들이 위험과 운영 비용을 줄이고 매장 내 트래픽 최적화 위해 컴퓨터 비전을 탐색. 리테일 분야에서의 컴퓨터 비전은 쇼핑객을 분석하여 유통 업체가 제품을 어디에 배치해야 하는지 파악하고, 재고를 보충할 시기를 결정하며, 고객 통계 자료를 더 잘 파악할 수 있도록 함으로써 매출 효과 극대화에 도움을 줌.

[3] 의료 서비스

MRI, CAT 스캔, X-ray 등의 이미지를 철저히 검사하여 의사차별 정확하게 이상 징후를 감지. 초음파 같은 3차원 이미지에 신경망을 사용하여 심박 등 시각적 차이를 감지할 수 있음.

[출처: https://m.blog.naver.com/sw_maestro/222082648250]

(3)자연어처리: 컴퓨터를 이용해 사람의 자연어 (naturallanguage, 컴퓨터에서 사용하는 프로그램 작성 언어 또는 기계어와 구분하기 위해 인간이 일상생활에서 의사 소통을 위해 사용하는 언어를 가리키는 말.)를 분석하고 처리하는 기술. 자연어 처리(NLP)에는 자연어 분석, 자연어 이해, 자연어 생성 등의 기술이 사용된다. 자연어 분석은 그 정도에 따라 형태소 분석(morphological analysis), 통사 분석 (syntactic analysis), 의미 분석(semantic analysis) 및 화용(話用) 분석 (pragmatic analysis)의 4가지로 나눌 수 있다.

자연어 이해는 컴퓨터가 자연어로 주어진 입력에 따라 동작하게 하는 기술이며, 자연어 생성은 동영상이나 표의 내용 등을 사람이 이해할 수 있는 자연어로 변환하는 기술이다.

자연어 처리는 [인공지능](#)의 주요 분야 중 하나로, 1950년대부터 [기계 번역](#)과 같은 자연어 처리 기술이 연구되기 시작했다. 1990년대 이후에는 대량의 말뭉치 (corpus) 데이터를 활용하는 [기계 학습](#) 기반 및 통계적 자연어 처리 기법이 주류가 되었으며, 최근에는 심층 기계 학습(deep learning) 기술이 기계 번역 및 자연어 생성 등에 적용되고 있다.

정보 검색, 기계 번역, 질의응답 등 다양한 분야에 응용된다.

출처 : [\[네이버 지식백과\]자연어 처리](#) [Natural Language Processing, 自然語處理] (용어로 알아보는

5G/AI/Blockchain, 한국정보통신기술협회)

(4) 자연어처리 활용사례

[1]환자와의 대화를 실시간으로 분석하여 임상 기록 생성하기

첫 번째로 소개하고자 하는 케이스는 **생성 AI를 결합한 의료진용 헬스케어 서비스인 'AWS HealthScribe(헬스 스크라이브)'**입니다. 이는 텍스트와 음성을 분석해 임상 기록을 생성하여 의료진의 문서작업을 돕는 서비스인데요. 환자와의 대화를 듣고 내용을 인식해 진료 기록을 실시간으로 작성하며 전자진료기록(EHR)에 직접 입력됩니다. 위 그림처럼, 해당 서비스는 상담 대화를 분석하여 주요 불편 사항, 현재 질병 이력, 평가, 치료 계획과 같은 섹션에 대한 요약된 임상 기록을 생성합니다. 이러한 요약을 제공하여 임상과의 의료 서기가 환자 방문의 주요 내용을 신속하게 요약하고 임상 문서용 기록을 검토, 편집, 마무리할 수 있습니다.

[2]독거노인의 긴급한 상황을 감지하는 인공지능 스피커

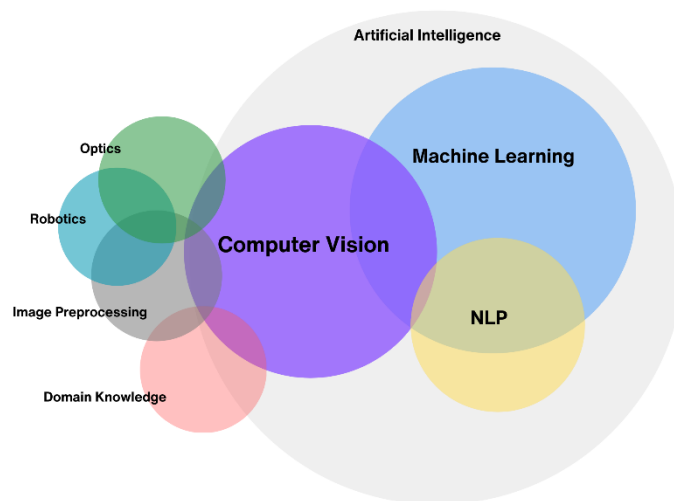
인공지능 스피커가 홀로 사는 어르신들을 구했다는 내용의 뉴스가 몇 년 전부터 많이 들려오고 있는데요. 일부 인공지능 스피커들은 '돌봄 기능'을 탑재하여 혼자 사시는 어르신이나, 거동이 불편한 분들을 위해 말벗이 되어주기도 하지만 생명을 구하는 일까지도 돕고 있습니다.

[출처:<https://aiheroes.ai/community/136>]

[컴퓨터 비전]

기본 개념

- 인공지능(AI)의 한 분야로, 기계 학습과 신경망을 사용하여 컴퓨터와 시스템을 가르쳐 디지털 이미지, 비디오 및 기타 시각적 입력에서 의미 있는 정보를 도출하고 결함이나 문제를 볼 때 권장 사항을 만들거나 조치를 취하는 것을 의미함
- 시각 데이터를 수집, 처리, 분석 및 이해하고 이 정보를 사용하여 현실 세계에 대한 의미 있는 표현, 설명 및 해석을 생성하기 위한 이론적 및 알고리즘 개발을 포함함



주 기능

- 탐지 (Detection): 이미지 내에서 여러 객체를 식별하고 위치를 파악하는 작업

- 분류 (Classification): 전체 이미지에 대한 단일 레이블 또는 범주를 예측하는 작업
- 식별 (Recognition): 단순히 탐지 및 분류 작업하는 것을 넘어 이미지 내 객체의 속성(e.g. 감정 인식, 객체 사이 관계)을 파악하는 작업
- 추적 (Tracking): 추정한 객체를 이용하여 다음 프레임(이미지)에서 객체를 위치를 다시 파악한 뒤 이미지 내 이동 경로를 계산하는 작업
- 세그멘테이션 (Segmentation): 픽셀 단위로 이미지 내 영역을 구분하여 클래스 또는 객체로 분류하는 작업

적용 사례

- 안양시 자율주행버스 '주야로'
 - ✓ 교통 상황을 정확히 파악하는 AI 영상 감지 기술 활용
 - ✓ 보행자, 역주행, 차량 정지 등 도로에서 발생 가능한 다양한 돌발상황을 영상 기반으로 감지 가능
 - ✓ <https://n.news.naver.com/mnews/article/003/0012569815?sid=103>
- SKT AI 동물 진단 서비스 '엑스칼리버'
 - ✓ 병원에서 촬영한 반려견의 근골격 및 흉부, 복부 등 엑스레이 사진에서 비정상 소견 여부와 위치 정보 등 분석결과를 수의사에게 제공함
 - ✓ <https://news.sktelecom.com/181576>

Reference

- <https://huggingface.co/learn/computer-vision-course/unit1/chapter1/definition>
- <https://paperswithcode.com/area/computer-vision>
- https://aws.amazon.com/what-is/computer-vision/?nc1=h_ls
<https://www.ibm.com/topics/computer-vision>

[자연어 처리]

기본 개념

- 인간 언어의 규칙 기반 모델링을 통계 및 머신 러닝 모델과 결합하여 컴퓨터와 디지털 디바이스가 텍스트와 음성을 인식 및 이해하고 생성할 수 있도록 함
- 인간의 언어를 이해하는 방법을 컴퓨터에게 가르치는 인공지능(AI)의 한 분야로, 음성 및 텍스트 데이터에서 특정 패턴 또는 의미를 찾아 이해하고 추론할 수 있도록 함

주 기능

- 분류 (Classification): 주어진 텍스트에 레이블이나 클래스를 부여하는 작업 (e.g. 감정 분석, 의도 분석, 뉴스 분류, 스팸 메일 분류)
- 번역 (Translation): 주어진 언어의 텍스트를 다른 언어의 텍스트로 변환하는 작업
- 요약 (Summarize): 중요한 정보와 의미를 유지하면서 긴 텍스트 문서를 더 짧고 간결한 버전으로 압축하는 작업
- 분석 (Analysis): 텍스트에서 정보를 분류, 정렬, 추출하여 패턴, 관련성, 감정, 기타 정보를 식별하는 작업
- 검색 (Retrieval): 텍스트에서 사용자 쿼리와 일치하는 정보를 찾는 작업
- 질의응답 (Question Answering): 사용자의 질문에 대한 답을 찾는 작업

적용 사례

- 비엔빛안과 의료 상담 AI 챗봇
 - ✓ 안과 전문 데이터베이스 구축하고 RAG(Retrieval-augmented Generation) 기술을 활용하여 질문에 정확하게 답변하도록 함
 - ✓ 라식, 스마일라식, 렌즈삽입술 등 시력교정술과 관련된 의료 용어를 쉽게 풀어서 설명하고, 시술 관련 주의사항과 수술 관련 질의응답, 예약 및 진료 안내 등을 제공함
 - ✓ <https://n.news.naver.com/mnews/article/346/0000075855?sid=103>
- 네이버웹툰 캐릭터챗
 - ✓ 웹툰 속 등장인물과 메시지를 주고받을 수 있는 챗봇 서비스
 - ✓ 대형언어모델 하이퍼클로바X 기반으로 동작하며 웹툰을 학습해 자연스러운 대화를 이어갈 수 있음

- ✓ <https://n.news.naver.com/mnews/article/023/0003840221?sid=105>

Reference

- <https://www.ibm.com/kr-ko/topics/natural-language-processing>
- <https://kr.mathworks.com/discovery/natural-language-processing.html>
- <https://www.nvidia.com/en-us/glossary/natural-language-processing/>
- <https://paperswithcode.com/area/natural-language-processing>
- <https://www.oracle.com/kr/artificial-intelligence/what-is-natural-language-processing/>

[컴퓨터비전+자연어처리]

제안하는 서비스

- 인간 활동 분석 기반 건강 관리 서비스
 - ✓ 시각 정보, 청각 정보를 활용하여 사용자의 이상 행동 및 현상을 감지하여 급성 질병에서 주로 나타나는 임상적 증상과 대조
 - ✓ 사용자에게 위험이 감지되면 즉시 알리고 의료서비스를 받도록 안내

핵심 분야 연구

- Human Action Recognition: <https://paperswithcode.com/task/action-recognition-in-videos>
- Emotion Recognition: <https://paperswithcode.com/task/emotion-recognition>