1️⃣ <처음보는 용어집>

* 케라스(Keras)

케라스는 빠른 딥러닝 실험을 목적으로 설계 되었으며 아이디어를 결과물로 최대한 빠르게 구현하는 것을 핵심으로 하며, 가장 큰 특징은 사용자 친화적이라는 점이다.

또한 단순한 인터페이스를 가지고 있어서 다른 프레임워크보다 쉽게 신경망을 구현할 수 있고, 모듈 형태로 구성되어 있어 모듈의 조합만으로 쉽게 새로운 모델을 만들어 낼 수 있다는 점이 있다.

출처: <https://terms.naver.com/entry.naver?docId=6653581&cid=69974&categoryId=69974>

* 텐서플로(TensorFlow):

구글 에서 개발한 기계 학습 엔진으로, 검색, 음성 인식, 번역 등의 구글 앱에 사용되는 기계 학습용 엔으로, 오픈소스 소프트웨어인 만큼 학생, 개발자 등 원하는 사람들은 누구나 사용할 수 있고, 스마트폰에서도 데이터 센터의 수천대 컴퓨터에서도 동작할 수도 있다고 발표하며 어떠한 제약도 받지 않고 유형하게 사용될 수 있는 기술이라 이야기하고 있다.

출처: <https://terms.naver.com/entry.naver?docId=3434677&cid=40942&categoryId=32837>

* 파이토치(pyTorch):

파이토치는 파이썬 언어를 염두에 두고 머신러닝 프로젝트를 위해 설계된 오픈 소스 라이브러리로서, 자동 차별화, 텐서 계산 및 GPU 가속을 전문으로 하고 있어서, 딥 러닝과 같은 최첨단 머신 러닝 어플리케이션에 매우 적합합니다.

출처: <https://tutorials.pytorch.kr/beginner/basics/intro.html>

* MNIST

손글씨 우편번호를 빠르게 읽기 위해 우편봉투로부터 숫자를 직접 추출해서 만든 손글씨 숫자 데이터셋 NIST를 일반화 및 표준화시킨 데이터셋임.

* 경량화

보통 많은 수의 파라미터 값을 가지는 딥러닝 모델의 특성상 용량이 큰데, 이렇게 메모리를 많이 차지하면 하드웨어 등에 탑재가 어려움. 그래서 실생활에서 딥러닝 기술을 사용하기 위해 모델의 크기를 줄이던지, 네트워크 구조 자체를 효율적으로 설계해야 하는데 그것을 경량화라고 함. (<https://velog.io/@woojinn8/LightWeight-Deep-Learning-0.-%EB%94%A5%EB%9F%AC%EB%8B%9D-%EB%AA%A8%EB%8D%B8-%EA%B2%BD%EB%9F%89%ED%99%94>)

* Quantization

학습된 딥러닝 모델이 weight 값을 저장할 때 사용하는 비트의 수를 줄여서 모델 크기를 줄임.

* Knowledge Distillation

규모가 크고 잘 학습된 네트워크를 활용해, 더 작은 네트워크를 학습해서 원래 네트워크 수준의 성능을 만든다. 큰 네트워크를 teacher 모델, 작은 네트워크를 student 모델이라고 부르고, 학습할 때 이 큰 네트워크들을 앙상블로 만들어 작은 student 모델을 학습한다.

* LLM

대형 언어 모델. 챗 GPT와 같이, 방대한 양의 텍스트 데이터에 대해 훈련된 인공 지능의 하위 집합임. 다층 신경망을 사용해 복잡한 데이터를 처리 분석, 예측함.

* 트랜스포머

입력 데이터를 처리하기 위해 self-attention 메커니즘을 사용하는 LLM 연구에서 널리 사용되는 신경망 아키텍쳐 유형

* self-attention 메커니즘

AI가 출력을 생성할 때 입력 텍스트의 특정 부분에 집중할 수 있도록 하는 알고리즘. (ex. 감정 관련 단어)

* API

영체제나 C, C++, Pascal 등과 같은 언어로 응용 프로그램을 만들 때, 윈도우를 만들고 파일을 여는 것과 같은 처리를 할 수 있도록 1,000여 개 이상의 함수로 구성되어 있다. API는 프로그래머를 위한 운영체제나 프로그램의 인터페이스로서 사용자와 직접 대하게 되는 그래픽 사용자 인터페이스나 명령형 인터페이스와 뚜렷한 차이가 있다.  
출처: [네이버 지식백과] [API](https://terms.naver.com/entry.naver?docId=4383195) [Application Program Interface] (소프트웨어 어휘다지기 - 중등)

* GPU

1990년대 초기까지만 해도 PC에 장착된 [그래픽카드](https://search.naver.com/search.naver?where=kdic&sm=ncc_clk&ie=utf8&query=%EA%B7%B8%EB%9E%98%ED%94%BD%EC%B9%B4%EB%93%9C)는 단순히 [CPU](https://search.naver.com/search.naver?where=kdic&sm=ncc_clk&ie=utf8&query=CPU)(Central Processing Unit: 중앙처리장치)의 연산 결과를 그림이나 글자 신호로 변환하여 모니터로 화면을 출력하는 어댑터(Adapter: 변환기)와 같은 부품으로 인식되고 있었다. 하지만 1990년대 중반 이후부터 PC의 멀티미디어 콘텐츠, 특히 게임이 주목을 받으면서 그래픽카드의 역할도 점차 변하기 시작했다.

게임에 입체감을 부여하고자 3D 그래픽이 본격적으로 도입되었고, 화면을 보다 현실적으로 만들기 위한 각종 광원 효과 및 질감 표현 기법이 점차 발전하기 시작했다. 이러한 작업들을 CPU 혼자서 처리하기에는 버겁기 때문에 이를 보조할 3D 그래픽 연산 전용의 프로세서, 즉 [GPU(Graphics Processing Unit)](https://search.naver.com/search.naver?sm=ncc_clk&where=kdic&ie=utf8&query=GPU%28Graphics+Processing+Unit%29&x=30&y=17)가 개발되어 그래픽카드에 탑재되기 시작했다.

출처: **[네이버 지식백과]** [GPU](https://terms.naver.com/entry.naver?docId=3572268) [Graphics Processing Unit] - 그래픽카드의 핵심 칩 (용어로 보는 IT, 김영우, IT 동아)

* 토큰화

어절이나 문장과 같이 자연언어 화자에게 익숙한 언어 단위는 컴퓨터에서 작업하는 데에 유효하지 않다. 따라서 데이터를 나누는 단위인 토큰을 설정하고, 컴퓨터에 입력할 텍스트 데이터를 토큰으로 나누는 전처리 과정이 필요하다.

토큰은 사용자가 실행하고자 하는 작업의 목적에 따라 유연하게 설정할 수 있다. 자연언어 처리 분야에서는 토큰을 단어(어절)나 문자와 같은 단위로 설정하는 것이 대체로 목적에 부합한다. 그러나 문자를 단위로 한 토큰화는 토큰 시퀀스의 길이가 지나치게 길어져 성능이 떨어질 수 있고, 단어(어절)를 단위로 한 토큰화는 어휘 집합의 크기가 지나치게 커져 성능이 떨어질 수 있다. 두 방식의 한계를 보완하기 위해 ‘서브워드(subword)’를 단위로 하는 ‘서브워드 분리(subword segmentation)’ 방식이 고안되었으며, 대표적인 서브워드 단위 토큰화 기법으로는 ‘바이트 페어 인코딩(Byte Pair Encoding, BPE)’이 있다.

출처: **[네이버 지식백과]** [토큰화](https://terms.naver.com/entry.naver?docId=6651335) (AI 용어사전)

2️⃣ <팀 미션 - 2문제>

\* 미션시, 토의방향 내용을 꼭 포함하여 작성해주세요.

**문제 1. 딥러닝 프레임워크(framework)는 검증된 많은 라이브러리와 사전 학습까지 완료된 알고리즘을 제공해 중복적인 기능 구현의 반복 작업을 줄여주고 문제해결을 위한 핵심 알고리즘 개발에만 집중하게 도와준다. TensorFlow, Keras, theano, Pytorch CNTK등 다양한 프레임워크들이 존재하지만, 강의에는 이 중 하나인 '파이토치'에 대한 필요성을 배웠다. 이번에는 파이토치 공식 홈페이지에서 튜토리얼을 보며 간단한 미니 프로젝트를 진행해 보자. [ 주제 : 파이토치 ]**

**아래 MNIST 튜토리얼 목차을 따라** [**torch.nn**](https://pytorch.org/docs/stable/nn.html) **,** [**torch.optim**](https://pytorch.org/docs/stable/optim.html) **,** [**Dataset**](https://pytorch.org/docs/stable/data.html?highlight=dataset#torch.utils.data.Dataset) **, 그리고** [**DataLoader**](https://pytorch.org/docs/stable/data.html?highlight=dataloader#torch.utils.data.DataLoader) **와 같이 잘 디자인된 모듈을 직접 활용해보고, 의견을 정리해보자.  
  
\* 튜토리얼 링크 :** [**https://tutorials.pytorch.kr/beginner/nn\_tutorial.html**](https://tutorials.pytorch.kr/beginner/nn_tutorial.html)

**[MNIST 튜토리얼 목차]  
1. MNIST 데이터 준비  
2. ()``torch.nn`` 없이) 밑바닥부터 신경망 만들기  
3. Torch 모듈 사용해 리팩토링 하기  
- torch.nn.functional, nn.Module, nn.Linear, torch.optim   
4. Dataset 을 이용하여 리팩토링하기  
5. DataLoader 를 사용하여 리팩토링하기  
6. 검증(validation) 추가하기  
7.fit() 와 get\_data() 생성하기  
8.CNN 으로 넘어가기  
9.nn.Sequential 사용하기  
10.DataLoader 감싸기  
11. GPU 사용하기**

**[작성시 유의사항]**

**- 파이토치 공식 홈페이지에 안내된 MNIST 미니 프로젝트를 성실히 수행**

**- 프로젝트 진행 과정 중 등장하는 여러 기능(데이터 변환, 코드 리팩토링, CNN정의, layer 정의, dataloader 사용 등 )들에 대해 본인이 이해한 내용이 잘 드러나게 회고 작성**

**- 수행한 프로젝트 코드 및 회고를 포함한 답안은 하나의 ipynb로 작성해 제출**

**(코드가 아닌 줄글의 경우, 주석으로 작성하거나 마크다운으로 작성하면 편리합니다.)**

**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**

* 컴퓨터 비전 (Computer Vision)
* 사람의 눈이 시각 자극 판독, 객체 구분, 인식의 과정을 거치는 것처럼, 컴퓨터 비전도 RGB 형식의 이미지를 인식하고 이해하는 방법을 학습시키는 것이다. 딥러닝을 사용해서 정확도가 올라간다. 주요 기능으로는 객체 탐지(Object Detection), 이미지 분류(Image Classification), 이미지 분할(Image Segmentation) 등이 있으며 이에 컴퓨터 비전 기술은 다양한 분야에서 활용되고 있습니다. 현재 가장 많이 활용되는 프로그래밍 라이브러리로는 Open CV가 있다.

(출처 : <https://terms.naver.com/entry.naver?docId=6653590&cid=69974&categoryId=69974>)

* 컴퓨터 비전 기능에 대한 설명

1) 객체 분류: 사전에 정의된 집합의 이미지에 적절한 이름을 할당한다.

2) 객체 식별: 객체 분류는 사전 정의된 집합에 이름을 할당하는 것이고, 객체 식별은 어떤 집합 내에서 특정 원소를 인식는 방법을 학습하는 것으로, 예를 들어 사람을 식별하는 얼굴의 특징을 학습시켜 배경 이미지에서 사람 얼굴을 식별하는 것

3) 자세 추정: 고정된 객체의 경우는, 객체의 위치와 방향을 추정한다. 움직이는 객체의 경우는, 상대적인 위치를 추정하는 것을 뜻함.

4) 인스턴스 추적: 동영상 분석에 쓰일 수 있음. 어떤 물체의 움직임을 계속 새롭게 학습 및 예측하면서 모델링함

5) 행동 인식: 이건 연속된 이미지 시퀀스가 있다고 생각하고 인식.

6) 장면 복원: 하나 이상의 이미지를 가지고, 그 장면의 기하학적 구조를 복원함. (두 사람의 시선에서 본 장면을 가지고 이미지의 대응점을 찾아 복원)

* 컴퓨터 비전 활용 서비스(현재)
* 자율 주행 자동차 : 컴퓨터 비전 기술을 활용하여 도로 상의 차선을 인식하고 신호나 표지판을 감지하여 자율 주행에 활용할 수 있습니다.
* 보안 시스템 : 얼굴 인식 기술을 활용하여 출입 통제 시스템이나 CCTV 에서 사람을 식별하고 감지할 수 있습니다.
* 이미지를 인식하여 핸드폰으로 스캔하는 서비스 (사진으로 pdf만드는 서비스, 객체 식별)
* 명함을 인식하여 연락처 관리해주는 서비스 (리멤버, 객체 식별)
* 파파고에서 특정 나라의 언어를 찍으면 다른 언어로 그 사진을 번역해주는 서비스 (객체 식별)
* 우크라이나 전쟁 무기에 적용될 수 있음 (<https://hackernoon.com/how-ai-helps-rebuild-lives-in-war-torn-ukraine>) (인스턴스 추적)

(출처: Hands-On Computer Vision with Tensorflow 2)

* 헬스케어 : 컴퓨터 비전은 의료 영상, 원격 진단 및 치료, 수술 지원 등 의료 분야에서 다양하게 활용되고 있는데요. 특히 의료 영상에서 컴퓨터 비전 알고리즘은 엑스레이, CT 스캔, MRI 이미지의 이상을 감지하고, 분류하는 데 사용되고 있습니다.
* 소매 및 E-커머스 : 컴퓨터 비전은 물체 인식, 고객 행동 분석, 재고 관리와 같은 소매 및 E-커머스 애플리케이션에서 점점 더 많이 사용되고 있는데요. 객체 인식 알고리즘은 매장 진열대에 있는 제품을 식별하고 분류하는 데 사용할 수 있으며, 고객 행동 분석은 소매업체가 고객이 제품 및 매장과 상호 작용하는 방식을 이해하는 데 도움이 되고 있습니다.

**([출처]** [인공지능의 눈으로 불리는 컴퓨터 비전이란?](https://blog.naver.com/smartnari/223056597035)|**작성자** [매니저S](https://blog.naver.com/smartnari))

* 운송: 컴퓨터 비전은 특히 자율주행 차량 및 교통관제 시스템 개발과 같은 운송 애플리케이션에서 중요한 역할을 하는데요. 자율주행 차량은 물체 인식 및 깊이 감지 등 다양한 컴퓨터 비전 기술을 사용하고 있으며, 교통관제 시스템의 경우, 컴퓨터 비전을 사용하여 교통 흐름을 모니터링하고, 사고나 혼잡을 감지하여 대응 시간을 단축하고 교통 관리를 보다 효율적으로 수행하는 데 쓰이고 있습니다.
* 제조 및 산업 : 컴퓨터 비전은 품질 관리 및 검사, 로봇 공학, 예측 유지보수 등 제조 및 산업 분야에서 다양하게 활용되고 있는데요. 품질 관리 및 검사 시스템은 컴퓨터 비전을 사용하여 결함을 감지하고 제품이 품질 표준을 충족하는지 확인하며, 예측 유지보수 시스템은 컴퓨터 비전을 사용하여 기계를 모니터링하고 고장이 발생하기 전에 잠재적 고장을 감지할 수 있습니다.

**([출처]** [인공지능의 눈으로 불리는 컴퓨터 비전이란?](https://blog.naver.com/smartnari/223056597035)|**작성자** [매니저S](https://blog.naver.com/smartnari))

* **컴퓨터 비전과 다른 기술 통합**
* 로봇 공학 : 컴퓨터 비전과 로봇 공학을 결합하면 로봇이 주변 환경을 더 잘 이해하고 탐색할 수 있어 더 정밀하고 효율적으로 작업을 수행할 수 있습니다.
* 사물 인터넷(IoT) : 컴퓨터 비전을 IoT 장치에 통합함으로써 이러한 시스템은 시각 데이터를 분석하여 가정 보안, 에너지 관리, 건강 모니터링 등 일상 생활의 다양한 측면을 모니터링하고 제어할 수 있습니다.

**([출처]** [인공지능의 눈으로 불리는 컴퓨터 비전이란?](https://blog.naver.com/smartnari/223056597035)|**작성자** [매니저S](https://blog.naver.com/smartnari))

* 증강 현실(AR) 및 가상 현실(VR) : 컴퓨터 비전은 사실적인 AR 및 VR 경험을 만드는 데 필수적인데요. 고급 컴퓨터 비전 기술은 이러한 몰입형 환경 내에서 추적, 렌더링 및 상호 작용을 개선할 수 있습니다.

**([출처]** [인공지능의 눈으로 불리는 컴퓨터 비전이란?](https://blog.naver.com/smartnari/223056597035)|**작성자** [매니저S](https://blog.naver.com/smartnari))

* 컴퓨터 비전 적용 가능 서비스

1. 자동주차 시스템

제품(시스템): 카메라 및 센서를 장착한 주차 시스템

기능: 주차장 내에 설치된 카메라와 센서를 활용하여 차량들의 위치와 움직임을 실시간으로 모니터링하고, 빈 주차 공간을 감지하여, 들어오려는 차량의 크기와 주차 공간의 적합성을 분석하고 이를 드라이버에게 최적의 주차 공간을 안내하거나, 자동으로 차량을 주차할 수 있게 할 수 있습니다. 이를 통해 주차 공간의 효율적인 활용과 주차 과정의 간편화를 도모할 수 있을 것이라 생각됩니다.

참고 아티클: AI로 주차 간단히.. 스마트주차의 도래

1. 농작물 질병 탐지 및 관리

제품(시스템): 드론 또는 카메라를 활용한 농작물 질병 탐지 시스템

기능: 드론이나 카메라로 농작물을 촬영하고, 컴퓨터 비전 기술을 활용하여 작물의 상태를 분석하여 작물의 잎이나 열매의 질병 징후를 식별하고, 신속하게 해당 질병을 탐지하게 됩니다. 이를 통해 농작물의 건강 상태를 모니터링하고, 질병 예방 및 조기 대응을 가능케 하여 작물의 수확량과 품질을 향상시킬 수 있을 것이라 예상됩니다.

참고 아티클: 드론 AI, 병충해 농작물 미리 탐지

1. 사람 얼굴 표정의 미묘한 변화를 민감하게 감지해서, 어떤 복합적인 감정이 나타나는지 알려줄 수 있음 (환자들의 사회인지 훈련에 쓸 수 있음) (객체 분류, 객체 식별, 행동 인식)
2. 홈트레이닝 서비스에서, 단순히 사람의 운동 자세의 정확도를 판독하는 것 뿐 아니라, 다음 자세 때 체형에 맞게 가이드를 더 섬세하게 줄 수 있음 (객체 분류, 객체 식별, 인스턴스 추적?, 행동 인식?)
3. 이 소프트웨어가 들어 있는 칩을 뇌에 이식하면 맹인들에게 인공 안구를 적용한 뒤 실제 뇌가 시각 정보를 받아들이는 것처럼 시각 정보 처리를 대신 담당하게 할 수 있음 (모든 기능)

* 자연어 처리 (Natural Language Processing)
* 자연어 처리는 인간의 언어를 이해하고 분석하는 기술로, 주요 기능으로는 텍스트 분류(Text Classification), 개체명 인식(Named Entity Recognition), 기계 번역(Machine Translation) 등이 있으며 이에 자연어 처리 기술은 다양한 분야에서 활용되고 있습니다.

(출처: <https://terms.naver.com/entry.naver?docId=863247&cid=42346&categoryId=42346>)

* 1990년대 이후 대량의 말뭉치 데이터를 활용해, 다음에 어떤 말이 나올지 통계적으로 예측해서 학습하는 처리 기법이 주류를 이룸. 요즘은 딥러닝 기술을 활용해 텍스트로부터 의미 있는 정보를 추출하고 활용하는 것이 더 쉬워졌음.
* 형태소 분석 -> 구문 분석 -> 의미 분석 -> 담화 분석 단계로 이루어져 있음
* 기능 : 기계 번역. 대화체 질의응답 시스템 (챗봇), 정보검색, 말뭉치 구축, 전체 텍스트 검색, 요약 등 (출처: <https://blog.cslee.co.kr/ai-research-and-technology-trends-nlp-part1/>)
* 자연어 처리 활용 분야(현재)
* 가상 개인 비서: 음성 인식과 자연어 처리 기술을 활용하여 음성 명령을 이해하고 실행하는 가상 개인 비서 서비스를 제공할 수 있습니다.
* 감정 분석: 소셜 미디어에서 사용자의 텍스트를 분석하여 긍정적인지, 부정적인지 감정을 판별하고, 이를 이용해 제품 평가나 마케팅에 적극적으로 활용할 수 있습니다.
* 자동 요약: 긴 문서나 뉴스 기사를 자동으로 요약하여 중요한 정보를 추출하거나 요약문을 생성할 수 있습니다.
* 이메일 필터링, AI 스피커, 번역, 뉴스 요약, 자동 철자 및 문법 교정, 챗봇 상담(출처: <https://enterprise.kt.com/bt/dxstory/746.do>)
* 자연어 처리 적용 가능 서비스

1. 가짜 뉴스 탐지 및 필터링

제품(시스템): 가짜 뉴스 탐지 및 필터링 시스템

기능: 자연어 처리 기술을 활용하여 온라인 뉴스와 기사를 분석하고, 가짜 뉴스를 식별하고 차단하는 기능을 작동시킬 수 있습니다. 이를 위해서는 텍스트의 문법, 어휘, 문맥을 분석하여 특정 패턴이나 신호를 감지하고, 신뢰할 수 있는 소스와 비교하여 정보의 신뢰도를 평가해야만 합니다. 또한, 이를 이용하면 소셜 미디어와 연결하여 널리 퍼지고 있는 허위 정보를 신속하게 탐지하고 경고를 제공할 수 있습니다. 이를 통해 사용자들은 신뢰할 수 있는 정보를 얻을 수 있고, 가짜 뉴스로 인한 오도방지와 피해를 예방할 수 있을 것이라 생각됩니다.

참고 아티클: 딥러닝이 가짜 뉴스와 싸울 수 있을까?

1. 제품(시스템): 감성 분석 및 감정 인식 시스템

기능: 자연어 처리 기술을 활용하여 텍스트 데이터에서 사용자의 감성과 감정을 분석합니다. 이를 위해 문장의 톤, 표현 방식, 단어 선택 등을 분석하여 긍정적인 감정, 부정적인 감정, 중립적인 감정 등을 식별하게 되는 것입니다. 이를 통해 사용자의 리뷰, 피드백, 소셜 미디어 게시물 등에서 특정 제품, 서비스, 이벤트에 대한 사용자들의 감정을 파악할 수 있게 되는 것입니다. 이 정보는 기업이 제품 개선, 마케팅 전략 수립, 고객 서비스 개선 등에 매우 공격적으로 활용될 수 있다고 생각합니다.

참고 아티클: 감성 분석을 통한 소셜 미디어 마케팅의 효과적인 활용

1. 음성 인식 기술과 결합해서 농인들과의 소통을 좀 더 원활하게 도와줄 수 있음 (실생활에서도)

1. 핵심 단어들을 추려서 한 사람의 하루 동안의 감정 일기를 자동으로 쓰고, 감정의 변화와 그와 관련된 사건들을 자동으로 빨리빨리 찾아줄 수 있음
2. 범죄수사에도 적용되어, 범죄자와의 면담 내용을 빠르게 정리해서 알리바이를 쉽게 도식화하고 범죄자의 논리 오류를 발견하는 데 도움을 줘서 수사 효율성을 높일 수 있음
3. 할머니 할아버지, 어린 아이들 등, 발음이 정확하지 않은 사람들과 소통을 도와줄 수 있음.