**[문항3] MOOC with IC-PBL**

**자기주도학습과제**

|  |  |
| --- | --- |
| **양식** | - 글자크기 10 / 맑은고딕 / 줄 간격 워드 1.0  - 반드시 양쪽 정렬로 작성할 것 EMB00002d681984  - 제시된 양식을 임의로 변경하지 말 것  - 수강 과목의 언어로 작성할 것 |
| **분량** | 주어진 양식에 맞춰 수행 |
| **충실도** | 욕설, 속어, 과제와 관계없는 내용, 표절한 내용을 작성할 경우에는 0점 처리  - 표절은 정도와 관계없이 0점 처리.  [예: 인용 양식 없이 인터넷 기사내용 및 통계자료 발췌 등]  - 고의적인 띄어쓰기, 줄 바꿈 밑 가운데 맞춤 시, 0점 처리 |

**MOOC Self Directed Learning (1)**

|  |  |
| --- | --- |
| **전공** | 컴퓨터소프트웨어학부 |
| **학번** | 2017029889 |
| **이름** | 심승현 |

|  |  |  |  |
| --- | --- | --- | --- |
| **사이트** | K-MOOC | | |
| **카테고리** | **컴퓨터공학, 인공지능/딥러닝** | **강사** | 오혜연 (KAIST) |
| **강의명** | **인공지능과 기계학습** | **강의시간** | 3시간 x 6주 |

|  |
| --- |
| **강의 선정 배경 [5줄이상]** |
| 최근 딥러닝과 인공지능을 이용하여, 바이오인포매틱스 분야에 적용하는 프로젝트를 진행하고 있다. 이러한 상황에서, 해당 강의는 2018년 강의이고, Transformer 모델이 상용화되기 전의 내용을 담고 있다. Transformer 모델 같은 경우, 압도적인 성능으로 인하여 Bert 등의 후속 모델들도 내부에 Transformer의 구조를 담고 있는 것을 볼 수 있는데, 이와는 다른 방향의 연구를 수강함으로써 딥러닝과 인공지능에 관한 다양한 관점을 얻을 수 있을 것이라 기대하였다. 또한, 현재 진행 중인 프로젝트에도 도움이 되지 않을까 기대하였다. |

|  |
| --- |
| **강좌 자기주도학습 계획(안) [10줄이상]** |
| 계절학기임을 고려하여 하루 1강좌이상 수강한다.  낮에는 인턴 등의 활동이 있으므로, 6주의 강의를 월/화/수/목/금/토 에 수강한 후,  다시 빠르게 들으며 요점 정리와 과제 작성을 토/일 간 수행한다.  또한, AI과 인공지능이 결코 가벼운 주제가 아니기 때문에, 해당 내용에 관한 면밀한 공부와, 이의 응용 방안 등에 대한 철저한 논증과 실험이 필요하기에, 과제를 제출하고 나서도 강의에서 소개되는 내용을 실제로 어떻게 적용할 수 있을지 꾸준히 고민한다.  매일 8시~11시 간 강의를 수강하고, 일단은 진도를 나가는 것을 최우선 목표로 둔다.  AI과 인공지능에 대한 기본 개념은 알고 있고, 해당 강좌에서도 그러한 개념들에 관하여는 따로 설명을 하지 않기 때문에, 강의를 들으며 헷갈리는 개념이 있다면 검색 등을 활용하면서 듣는다.  계절학기이기 때문에 시간이 여의치 않을 수 있으므로, 진도를 최우선 과제로 두고, 과제 제출 이후에 이를 실제로 적용하는 부분에 관한 논증을 이어나간다.  해당 강의의 내용을 실제로 실습해 볼 환경으로는, 구글 Colab 과 Vscode 를 활용하여 실습을 진행하고, 간단한 예제의 경우 윈도우 기본 파워셸에서 구동한다.  강의를 시작하기 전에, Tensorflow, Keras 등의 인공 신경망 라이브러리에 관한 기본적인 구조를 학습 후 학습에 임하도록 한다. |

**MOOC Self Directed Learning (2)**

|  |
| --- |
| **주요 강의 화면 [화면캡쳐]** |
|  |

|  |  |
| --- | --- |
| **강의내용**  **요약** | 다중 퍼셉트론에 관한 기본적인 개념을 담고 있는 슬라이드이다.  어떠한 뉴런이 존재하면 그 뉴런에는 입력값과 출력값이 존재하고, 그러한 입력값과 출력값이 데이터에 곱하거나 더해져서 (일반적으로는 Matrix 형태를 띄면서 연산을 하고, 이는 GPU가 딥러닝 모델에서 훨씬 많이 쓰이는 이유이다) 다음 뉴런에 전달되며, 그 과정에서 Weignt(곱해지는 값) 과 Bias(더해지는 값) 을 Optimizer 혹은 다른 Object 를 통해서 조절해주는 것이 다중 신경망을 활용한 딥러닝 모델이다.  다중 신경망 모델의 경우 입력을 여러 뉴런에서 동시에 받고, 출력 또한 여러 레이어로 해주고 있고, 일반적으로 딥러닝 프로젝트의 목료는 이러한 값을 우리가 실제로 예측한 값(정답)과 차이를 줄여나가는 것이고, 컴퓨터과학에서 드물게 실험적인 분야라고 볼 수 있다. |
| **학습노트**  **[주요한 점]** | 하나의 레이어에 여러 뉴런들로 이루어 져 있음  Fully-connected Neural Network 같은 경우에는, 이러한 뉴런들이 이전/다음 레이어의 모든 뉴런과 연결되어 값을 모든 뉴런에서 반영할 수 있도록 해 줌.  이러한 과정에서, 입력과 출력은 모든 뉴런이 가지고 있는 Weight 와 Bias 에 의해서 값의 변화가 결정되고, 이를 변형시켜 목표값을 찾는 것이 딥러닝의 일반적인 목표 |

**MOOC Self Directed Learning (3)**

|  |
| --- |
| **주요 강의 화면 [화면캡쳐]** |
|  |

|  |  |
| --- | --- |
| **강의내용**  **요약** | 은닉층 Gradient (기울기) 에 관한 연산을 담고 있는 화면이다.  은닉층이라 함은, 앞서 이야기한 다중 신경망 네트워크에서 맨 처음 레이어와, 맨 끝 레이어를 제외한 레이어들을 Hidden Layer 이라고 한다. 이러한 Hidden Layer 에 대하여, 모든 뉴런들이 서로 다르게 가지는 Weight 와 Bias 값을 수정해주는 절차를 거쳐야 하는데, 얼마나 수정해줄지를 정하는 과정이 Gradient Computation 이다. 딥 러닝에서는 기본적으로 Gradient Descent (경사 하강법) 을 활용하고, Loss Function 에서 구해진 Loss 를 딥러닝 과정중에 활용된 Variable에 대해서 미분해주는 절차이다. 그림에서 색 화살표로 표현된 부분이 입력 레이어까지 전달되고, 이를 용어로 하면 역전파, Back Propagation 이라고 한다. RNN 등의 순환 신경망에서는 BPTT 라는 기법을 사용하지만 원리는 같다. |
| **학습노트**  **[주요한 점]** | Back Propagation 과정의 입력값은 Loss 와 Gradient(Loss, variables) 이다. Loss를 variables 라는 변수에 대해서 미분한것인데, 각 레이어를 통과하면서 Loss를 계속 업데이트 해주며 미분값도 계속 변하게 된다. 이를 수정하는 절차는 Optimizer 등이 Gradient Apply Policy 에 따라 기울기를 수정하게 되고, 이러한 과정에는 기본적인 편미분을 사용하여 계산한다. |

**MOOC Self Directed Learning (4)**

|  |
| --- |
| **주요 강의 화면 [화면캡쳐]** |
|  |

|  |  |
| --- | --- |
| **강의내용**  **요약** | 딥러닝을 활용하여 실제 언어 모델을 만들 때, 어떠한 단어가 숫자로 표현되는 과정을 담은 것이다. 이 화면을 선정한 이유는, 딥러닝의 대표적인 사용처가 NLP, Natural Language Problem 에서 딥러닝이 주로 사용되는데, 이러한 문제에서 Word Representation 이 필수적인 과정이기 때문이다. 어떠한 단어가 있으면, 오메가(w)로 표현된 것이 어떠한 word dictionary 에서 index (ID) 를 표현하게 되고, 이를 어떠한 Dimension 의 vector 로 표현하게 되는 것이 C(w)로 표현된 단어 표현값이다. 현재 통용되는 Transformer 나 다른 모델에서도 어떠한 단어를 저렇게 숫자로 풀어서 표현하게 되고, 몇 차원으로 표현할 것인지는 보통 단어의 압축률과 연결된다. 이 vector 를 활용해서 NN에서 행렬 연산을 하게 된다. |
| **학습노트**  **[주요한 점]** | The 라는 단어는 String type 이고, 이러한 String type 이 Int64 type 인 Omega(w) 로 바뀌게 된다. 또한, 이 Int64 type를 C라는 함수가 받아서 float32 type의 array를 가지게 되는 Representation을 가지게 되고, 이것이 각 뉴런의 weight matrix, bias value 등과 행렬 연산을 거치게 된다. 그리고 이러한 행렬에 대한 결과 행렬의 미분값이 NN의 결과로 구해지고, Back Propagation 을 통해서 Gradient Computing / Applying 을 수행하게 된다. |

**MOOC Self Directed Learning (5)**

|  |  |
| --- | --- |
| **MOOC 강의 Reflection [느낀점]** | |
| Transformer 이전의 AI/DL 에 관한 관점을 볼 수 있을 것으로 기대했고, 강의를 수강하고 나서 보면 Transformer 이나 Bert 등의 현대적 딥러닝 모델의 Basic Idea 들을 소개하는 강의인 느낌이다. 꽤나 무거운 주제였음에도 불구하고, 기본적으로는 한국어 강의였기 때문에 따라가기 어렵지는 않았고, 기본적인 강의 타겟층 자체가 AI/DL 에 관심이 있는 공학도들을 대상으로 진행되는 강의였던 것 같다. 이 강의에서는 DL 모델이 등장하게 된 지점, DL 모델의 정의가 무엇인지, DL 모델의 발전과정과 그러한 발전 과정에서 포커스가 어느 쪽으로 맞추어 져 있었는지, 그리고 그런 모델들을 사용해서 NLP등과 같은 문제에 어느 정도 성능을 보였는지, 또한 자연어 모델을 강의에서 소개한 이론을 통해 빌드 해봄으로써, 현대까지도 주요한 문제로 다루어지고 있는 NLP문제와 연결할 수 있었다.  다만 아쉬운 점은, 코드 등을 동반한 실습 강의라기보다는, 기본적인 이론과 그것을 수학을 통해서 설명하는, 학부의 딥러닝 및 응용 이론강의와 진행 방식이 비슷하였고, 18시간에 달하는 시간 동안 코드 한 줄 없이 이러한 이론만을 수강하는 것에 대하여 조금 아쉬운 부분이 있었다. 하지만 이 강의를 주위 DL에 관심있는 사람들에게 기본적인 개념을 쌓을 수 있는 강의로 추천해줄 수 있을 것 같고, 강의 초반부의 최대 마진 분류기와 같은 내용들은 딥러닝을 공부하는 나로서도 거의 들어보지 못했던 문제였기 때문에 정말 흥미롭게 들을 수 있었다. | |
| **나의 학업**  **직무 관련 시사점**  **[기업가정신]** | 현재 진행하고 있는 딥러닝 관련 프로젝트의 기초 개념을 더욱 더 튼튼히 함으로써, 앞으로의 프로젝트 과정에 도움이 조금 더 될 것이라는 점과, 딥러닝/인공지능 분야의 시작과 발전 과정을 이해하여 딥러닝 모델들의 포커스를 이해할 수 있었고, 앞으로 이러한 모델들이 어떤 식으로 발전해나갈 지 이해하는 것에도 도움이 될 것 같다. |
| **후 계획** | 딥러닝의 기본 개념들을 다룬 이후, 강의에서 응용 개념으로 등장한 최대 마진 머신, 확률론, 언어 모델 등의 사례를 활용하여 현재 진행하고 있는 프로젝트의 딥러닝 머신에 응용할 수 있는 방안이 있는지 모색할 계획이다. 또한, 확률론 등은 심도 있게 공부한다면 앞으로의 진로에 큰 도움이 될 지식이므로 공부를 이어나갈 계획이다. |

**MOOC with IC-PBL (6)**

|  |  |  |
| --- | --- | --- |
| **지속가능발전목표(Sustainable Development Goals, SDGs)란 무엇인가?** | | |
| SDGs 는 심화되는 환경문제로 인해 논의되었던, 지속 가능한 발전 목표에 대한 논의이고, 유엔 총회에서 결의된 부분이다. 세부 분야로써는 빈곤, 질병, 교육, 성평등, 난민, 분쟁, 기후 변화, 환경 오염, 기술, 소비, 생산, 법, 경제 등의 키워트를 통해 사회 전반적인 분야를 아우르고 있다. 이러한 키워드를 통해, 환경친화적인 발전을 도모해야 한다는 내용이다. 예를 들어, 중국 등의 개발도상국이 뒤늦은 개발로 인하여 에너지 소비가 막대한데, 이를 일방적으로 환경 문제를 언급하며 제재하기에는 선진국의 사회적 책임이 없지 않다. 이에 따라 선진국부터 환경적인 발전을 도모하며 이산화탄소 배출, 온실가스 배출 등에 관하여 감축 결의가 있었고, 이는 점차 우리의 발전 방향이 자연을 이용하는 방안으로 지금까지 발전해 왔지만, 우리는 서서히 자연과 공존하는 방향으로 꺾어야 할 시점이 왔음을 시사하고 있다. 실제로, 최근 환경 시계, 인간이 막을 수 없는 급속한 기후변화가 오기까지의 시간을 나타내는 환경 시계를 보면 시간이 얼마 남지 않은 것을 알 수 있다. 따라서 더 이상 문제가 심각해지기 전에, 선진국부터 모범을 보여 우리 사회가 점차 친환경적인 발전을 도모할 수 있도록 준비해야 한다. 최근 스타벅스 등의 카페나 음식점 등에서 일회용품의 사용을 줄여나가고 있고, 또한 더욱 나아가서 일회용품을 사용하더라도 재활용이 용이한 형태로 제공하는 것이 그러한 일환으로 볼 수 있다. | | |
|  | | |
| **지속가능발전목표와 관련된 나의 학업 혹은 관심분야의 문제점 및 현재 상황** | | |
| **지속가능발전목표**  **항목:**  **8**  **내용:**  컴퓨터공학이 이끄는 무인화, 그리고 일자리 | 컴퓨터공학은 지속 가능한 발전 목표와 가장 동떨어진 분야라고 생각할 수 도 있을 것 같다. 기본적으로 컴퓨터공학의 발전이 사회적인 의미에서는, 실업자 수의 증가 등을 초래한다고 생각했었기 때문이다. 이러한 과정에서 컴퓨터공학의 발전으로 인해, 관리 인력 등의 수요가 점점 줄어듦에 따라 일어날 수 있는 사회 문제가 있다. 실제로 공장 등은 관리 인력 (공장 프로그램을 다루는 사람) 한두명만 있어도, 별다른 관리 인력이 필요가 없다. 과거 수백명의 노동자들이 공장을 이루었던 것을 생각해보면, 현재 수많은 로봇들과 그 관리자 한두명만이 공장을 운영하는 것은 분명 사회적 문제를 야기할 것 같아 보인다. 실제로 수많은 분야에서 무인화가 이루어지고 있다. 점점 올라가는 인건비(한국은 인건비가 굉장히 저렴한 나라지만)에 따라, 무인 편의점, 무인 아이스크림 할인점, 무인 자전거 대여, 심지어는 무인 카페까지, 로봇으로 대체할 수 있는 분야는 최대한 사람을 밀어내고자 하는 일이 일어나고 있고, 이는 분명 실업 등의 문제를 야기할 것 같다. |

**MOOC with IC-PBL (7)**

|  |
| --- |
| **문제점을 개선하고 해결할 수 있는 내용 및 방법** |
| 로봇세를 걷어야 한다. 과거에는 공장 등의 운영 수익의 꽤 많은 수의 노동자들에게 돌아갔지만, 현재는 이러한 공장 운영으로 얻는 수익이 로봇 등의 구매와 관리자에게만 돌아간다. 이러한 비상식적인 수익 구조는 여러 명의 실업자를 만들면서, 차세대 IT 회사들에게 막대한 수익을 안겨주고 있다. 예를 들어 현재 카카오가 논란이다. 카카오가 자사의 소프트웨어를 무기로 삼아. 굉장히 방대한 분야의 진출을 도모하였으며, 실제로 그 중 굉장히 많은 분야에서 기존 사업자들을 밀쳐내고 독과점의 행태를 보이고 있다. 카카오게임즈, 카카오택시, 카카오뱅크 등 이러한 차세대 IT 회사의 소프트웨어가 가진 파급력은 엄청나다. 하지만 카카오는 IT기업의 특성상 많은 인원을 고용할 필요가 없다. 개발자들과 관리자 몇 명만 있으면 돌아가는 게 카카오다. 그에 따라 카카오가 독과점 시장을 형성하여 벌어들이는 막대한 수익이 노동자들에게 돌아가지 않고 자사의 덩치만 키우고 있는 것이다. 이러한 현상을 방지하기 위하여 IT 기업들에 막대한 양의 세금을 부과하여 그로 인해 발생하는 실업자들을 지원할 방법을 강구하여야 한다. 지금이야 카카오, 네이버 등 몇몇 기업에 불과하지만, 소프트웨어가 가지는 파급력을 고려한다면, 조만간 꽤나 많은 회사들이 이러한 형태를 취할 것이고, 그때는 실업률을 사회적으로 감당하기 힘들어 질 것이다. 이미 미국에서는 역대급으로 많은 퇴사자가 나오는 등, 일 안하는 세대가 오는 것 아니냐는 예측이 팽배하다. 이것 또한 사회의 변화이기에, 많은 사람들이 이러한 문제 속에서 또한 공존할 수 있는 방안을 강구해야 할 것이다. |
|  |
| **방법의 기대효과 및 시사점** |
| 앞서 이야기했듯, 카카오 등 대형 IT 회사는 소프트웨어가 가지는 파급력 덕분에, 꽤나 많은 분야에서 독과점의 시장을 형성하며, 특유의 낮은 인건비 등으로 경쟁 회사들을 밀어내고 있다. 하지만 이러한 회사들은 몇몇 개발자들과, 시스템 관리자들만 있으면 돌아가는 회사이고, 추가 고용을 할 이유가 없는 회사이기에, 양질의 일자리는 만들어지지 않고 있다. 이는 현재까지 수많은 기업들의 사회적 책임인, 양질의 일자리 창출에 크게 위배되는 행위이고, 근본적으로 사회의 시스템을 흔들 것이다. 하지만 이는 또한 사회의 흐름이기 때문에, 이것을 억지로 거슬러 고용을 촉구하는 것은 또한 많은 부작용 등을 초래할 수 있고, 국가 경쟁력 등에 문제가 될 수 있다. 그래서 이런 대형 IT 회사들에 대하여 사회적 책임의 일환으로, 고용 대신 세금을 요구한다면, 그러한 세금을 올바르게 집행한다는 가정 하에, 그로 인하여 생기는 많은 실업자들을 지원하고, 사회적으로 의미 있는 어떠한 활동들을 정부 차원에서 새로운 일자리로 포장하여 지원금 등을 지급하는 방안을 통해 이러한 문제를 해결할 수 있을 것이라 생각한다. |