

Research Proposal

Project Name	멀티태스킹 학습 기법 탐색과 적용 및 성능 개선 연구
-----------------	-------------------------------

08조

202001156 김수영

202002510 송재현

지도교수: 이종률 교수님 (서명)

Document Revision History

REV#	DATE	AFFECTED SECTION	AUTHOR
1	2025/03/11	2. 연구 배경 및 관련 연구, 3. 프로젝트 수행자의 의도 작성	김수영
2	2025/03/14	2. 연구 배경 및 관련 연구, 3. 프로젝트 수행자의 의도 작성 수정 및 추가 작성 + work summary 작성	김수영
3	2025/03/14	5. 프로젝트 관련 학습 계획, 6. 연구 일정 계획 작성	김수영, 송재현
4	2025/03/14	4. 탐구 내용 및 기대 결과 작성	송재현
5	2025/03/14	4. 탐구 내용 및 기대 결과 수정 및 추가 작성	송재현
6	2025/03/15	보고서 최종 정리	김수영, 송재현

Table of Contents

목차

1.	연구 주제 이름	4
2.	연구 배경 및 관련 연구	4
3.	프로젝트 수행자의 의도	6
4.	탐구 내용 및 기대 결과	7
5.	프로젝트 관련 학습 계획	8
6.	연구 일정 계획	8

1. 연구 주제 이름

Multi-Task Learning의 최적 공유 전략 연구 확장 및 실험적 유효성 분석

2. 연구 배경 및 관련 연구

Multi-Task Learning (다중 작업 학습 이하, MTL)이란, 여러 학습 작업을 동시에 해결하면서 작업 간의 공통점과 차이점을 활용하는 기계 학습의 하위 분야이다. 1997년 발간된 Rich Caruana의 Multitask Learning에서는 다음과 같은 MTL의 특징을 제시하였다.

- MTL은 귀납적 전이(inductive transfer)이다.
- 병렬적으로 학습이 진행된다.

즉, MTL은 여러 Task들(일련의 학습 Process, Multi-Task)간 병렬적인 귀납적 전이 학습을 통해 전체 Task의 성능(Accuracy, Loss 등등)을 향상시키는 패러다임이다.

MTL 연구의 배경에는 다양한 동기가 존재한다. 생물학적 영감으로 인간이 여러 기술을 동시에 배우는 능력에서 아이디어를 얻었으며, 교육학적 관점에서는 기초 과목 학습이 여러 응용 과목 학습에 도움을 주는 것처럼 관련 작업들을 동시에 학습하는 것이 주 작업의 성능 향상에 기여할 수 있다고 본다. 기계 학습적 관점에서는 관련 작업들의 학습 신호가 유도적 편향(inductive bias)으로 작용하여 모델의 일반화 성능을 높일 수 있다고 설명한다. 영화 베스트 키드에서의 가라테 훈련 비유처럼, 겉보기에는 관련 없어 보이는 보조 작업 (예: 페인트칠, 왁스칠)이 실제 목표(가라테 기술 습득)에 필요한 기반을 다져주는 것이 MTL의 핵심 아이디어 중 하나이다

MTL 기법의 특징으로는 단연 일반화로 여러 Task를 서로 학습시키기에 하나의 통합된 모형으로 여러 Task들에 동시에 적용시킬 수 있다는 장점이 있다. 가령, 스팸 메일 분류 문제를 일반화한다 해보자. 영어권 화자가 중국어로 된 메일을 받을 경우 영어권 화자는 이를 거의 대부분 스팸 메일로 처리할 것이다. 하지만, 중화권 화자의 경우 이를 스팸 메일로 분류할 때 내용을 좀 더 살펴보고 분류할 것이다. 즉, 화자 별 스팸 메일의 분류 기준은 다르다.

하지만 스팸 메일을 판별하는 기준은 대부분의 화자가 동일한 프로세스를 거칠 것이다. 가령 광고성 url이 여럿 기입되어있거나 전화번호가 여럿 적혀있거나, 과도한 그림이나 사진이 첨부되어있는 등 말이다. 이처럼 Task들 사이에 존재하는 ‘Shared Representation’ 혹은 ‘Shared Knowledge’가 MTL의 핵심이다. 이러한 Shared Knowledge를 어떻게 찾아내고 정의하며 이용할 지가 현대 MTL 분야에서의 핵심 과제이다. 일반적으로 알려져 있는 MTL의 기법으로는 Hard Sharing과 Soft Sharing이 있다.

Hard Sharing은 초기 MTL에서 가장 흔하게 사용된 방법으로 여러 Task가 하나의 공유된 계층(Shared Layers)를 통해 입력 데이터를 처리하고 공통된 특징을 학습시키는 방법이다. 이는 각각의 작업이 유사한 특징이나 지식을 공유한다는 가정을 고려하여 개발된 학습 방법이며 그렇기에 관련 있는 작업을 함께 학습하는 것이 각각의 작업을 독립적으로 학습하는 것보다 더 나은 성능을 얻을 수 있다 기대한다. 가령, 사물을 바라볼 때 그 사물의 윤곽, 모양, 질감 등을 인식하는 여러 작업(RGB 영역, Gray Scale 영역 등)을 동시에 학습하면 복잡한 실제 세계의 사물을 더 잘 인식하고 분류할 수 있다는 것이 Hard Sharing의 접근 방식이다. 또한, Hard Sharing은 결국 하나의 모델로 여러 작업을 동시에 수행하도록 학습되기에 각각의 작업을 위한 별도의 모델을 학습하는 단일 작업 학습(Single Task Learning)에 비해 모델의 크기가 작고 계산 효율성이 높아진다. 이는 모델의 경량화와 일반화에 보다 적합하다는 특징을 지닌다.

그러나 Hard Sharing에 이용되는 작업들이 너무 다르거나 서로 관련이 없다면, 공유된 표현이 오히려 각 작업의 성능을 저하시키는 부정적 전이(Negative Transfer)가 발생할 수 있다. 이는 MTL의 각 작업들이 서로 매우 다를 경우, 작업에 대한 각각의 독립적인 모형을 사용하는 것이 더 나을 수 있다는 것을 시사한다. 또한, 공유되는 계층의 용량이 너무 작으면 각 작업에 필요한 정보를 충분히 담아내지 못할 수 있으므로 모형의 크기 또한 적절히 선정해야한다. 마지막으로 여러 작업을 동시에 학습할 때 각 작업의 학습 속도나 난이도가 다를 수 있으므로, 학습 과정에서 각 작업에 적절한 중요도를 부여하는 것이 중요하다. 즉, 모든 작업이 동일한 속도로 학습되거나 과적합(Overfitting)되지 않을 최적의 성능을 보이는 시점을 찾는 것이 중요한 것이다.

Soft Sharing은 Hard Sharing과 대조적인 접근방식이다. 여러 작업들이 모형의 parameter를 일부분만 공유하거나, 파라미터는 독립적으로 유지하되 학습 과정에서 서로에게 영향을 주는 방식을 의미한다. Hard Sharing의 경우 언급하였듯 부정적 전이(Negative Transfer)가 발생할 수 있다. 따라서, 이를 해결하기 위해 작업 별로 부분적인 공유 혹은 작업 별 독립적인 parameter를 가지게 하여 학습 도중에 정규화(regularization) 같은 메커니즘을 이용해 지식 전이(Knowledge Trasfer)가 일어나도록 유도한다. 이는 Soft Sharing에서 각 작업의 독립성을 어느 정도 보장이 중요함을 시사한다. 간접적인 정보 교환은 Soft Sharing의 핵심 메커니즘이다. 좀 더 구체적인 예시를 들어보도록 하자. Multi-Task Learnig의 학습방

법 중 하나인 Nash Multi-Task Learning은 존 내시의 게임 이론에 등장하는 개념인 bargaining solution에서 영감을 받은 학습방법이다. 이 학습방법은 각 작업 별 gradient를 계산하고 이 gradient들은 Nash Bargaining Solution이라는 합의에 의해 gradient들이 업데이트될 방향을 결정함으로써 모형을 학습시킨다. 이처럼 작업의 parameter들이 특정 메커니즘에 의해(현 예시에서는 Nash Bargaining Solution) 업데이트되며 최종적으로 모형이 학습하는 것이 Soft Sharing인 것이다. 이를 통해 앞서 언급한 부정적 전이의 가능성이 낮아지며 Hard Sharing과 달리 각 작업의 특성이 뚜렷하거나 관련성이 적을 경우에는 더 효과적일 수 있다.

이처럼 기존 MTL 연구들은 다양한 분야에서 성능 향상을 입증하였지만, 여전히 몇 가지 중요한 한계를 가지고 있다.

첫째, Hard Sharing 방식의 경우 작업 간 공유된 표현이 모든 작업에 최적화되지 않으면 성능이 저하되는 부정적 전이(Negative Transfer) 문제가 발생할 수 있다. 반면, Soft Sharing 방식은 각 작업의 독립성을 유지할 수 있으나, 학습 과정에서 정보를 효과적으로 교환할 수 있는 최적화된 메커니즘이 부족하여 아직까지는 알려져 있는 방법인 data augmentation, data regularization을 이용하는 실정이다. 이에 대한 연구들은 대부분 작업 간 공유 전략을 수동으로 설계하는 방식이 대부분이었으며, 작업 간 공유 구조를 동적으로 조정하는 최적화된 학습 전략이 부족하다는 한계가 있다.

둘째, 기존 연구들은 개별 작업에 대한 최적화된 MTL 모델을 설계하는 데 초점을 맞추었으나, MTL의 유효성을 실질적으로 검증하는 연구가 부족하다. 대부분의 실험은 기존에 알려져 있는 방법인 Soft Sharing과 Hard Sharing간 비교가 이루어지거나 특정 Single Task Learning과의 비교만을 대상으로 실험이 진행되었다. 또한, 기존의 다양한 연구에서 개발된 개별 모델들을 MTL 프레임워크 내에서 통합하여 새로운 모델을 개발하고, 이를 통해 MTL의 실질적인 효과를 평가하는 연구는 상대적으로 적었다. 따라서, 단순히 MTL의 개념을 활용하는 것을 넘어, 여러 기존 논문의 모델을 MTL을 통해 통합하여 새로운 MTL 기반 모델을 개발하고 그 유효성을 검증하는 연구가 필요하다.

3. 프로젝트 수행자의 의도

Multi-Task Learning(MTL)은 다양한 기계 학습 문제에서 효율성과 성능을 높이는 강력한 방법론으로 주목받고 있다. MTL의 핵심 개념은 여러 작업(Task) 간 공유된 표현(Shared Representation)을 학습하여 일반화 성능을 향상시키는 것이다. 이는 데이터의 양이 기하급수적으로 증가하고 이를 처리할 기술적 바탕이 마련된 현 시대에 걸맞은 학습방법이다. 또한, 가면 갈수록 모형의 절대적인 연산량이 증가하고 이를 위한 컴퓨팅 파워의 증가와 그에

따른 전력 소비 역시 덩달아 증가하는 연쇄적인 문제 또한 머신러닝 분야에서 화두에 오르고 있다. 또한, 과도한 모형의 중량화는 필요없는 연산을 요구하며 이러한 고중량 모형은 임베디드와 같은 분야에서 제대로 적용시키기 어렵다는 한계가 있다. MTL은 각 작업들을 일반화(Generalization)하여 통합된 모형을 생성할 수 있기에 모형을 경량화하고 핵심적인 특징을 학습할 수 있다는 장점이 있으며 다양한 작업에 대해 유동적으로 대처할 수 있다. MTL은 현재 머신러닝 분야가 직면한 문제인 경량화에 효과적으로 대처할 수 있다. 이는 인공지능이 일상생활에서 자연스럽게 사용되는 현대사회에 꼭 필요한 조건이다. 그렇기에 향후 머신러닝의 소양이 되는 경량화를 충실히 수행할 수 있는 MTL을 주제로 프로젝트를 수행하려는 것이다.

본 연구자는 이러한 MTL의 장점을 활용하기 위한 효율적인 공유 전략 및 학습 방법을 탐구하고자 한다. 기존에 알려져 있는 논문들을 탐색하고 여러 논문에서 사용된 모형을 Task로 삼아 통합된 모형을 만들어보며 이를 통해, MTL이 다양한 응용 분야에서 보다 최적화된 방식으로 활용될 수 있도록 기여하는 것이 본 프로젝트의 의도이다. 이후 MTL의 유효성이 입증된다면 추가적으로 MTL의 작업 간 공유 구조를 동적으로 최적화하여 성능을 더욱 더 향상시키는 것이 최종적인 연구의 목표가 될 것이다.

4. 탐구 내용 및 기대 결과

다중 작업 학습(Multi-Task Learning, MTL)에서 사용되는 대표적인 기법들은 하드 파라미터 공유, 소프트 파라미터 공유, 계층적 공유, Sparse Sharing 등이 있다. 본 연구에서는 이러한 기법들의 장단점을 비교 분석하고, 각각의 기법이 다양한 태스크에서 어떠한 영향을 미치는지 평가한다.

또한, MTL에서 발생하는 주요 문제를 식별하고자 한다. 대표적인 문제로는 과적합(overfitting)과 학습의 불안정함이 있으며, 이는 모델의 일반화 성능을 저하시킬 수 있다. 따라서 기존 기법들이 이러한 문제를 어떻게 해결하고 있는지 검토하고, 추가적인 개선이 필요한 부분을 도출하는 것이 목표이다.

기존 MTL 기법의 분석을 바탕으로, 본 연구에서는 MTL의 성능을 향상시킬 수 있는 방법을 탐색하고자 한다. 이를 위해, 다양한 모델 구조에 대해 MTL의 성능을 정량적 및 정성적으로 평가하여 비교 분석을 수행한다. 이를 통해 여러 MTL 기법들이 어떤 점에서 우수한지를 객관적으로 검증한다.

본 연구를 통해 얻은 결과를 바탕으로 학술대회에 논문을 제출하고, 좋은 성과를 거두는 것을 목표로 한다. 주요 학술대회로는 한국컴퓨터 종합학술대회(KCC), 한국인공지능학회(KAAI), 데이터마이닝 및 인공지능 학술대회(KDM/KDML) 등이 있다.

MTL 성능 향상을 정량적으로 평가하기 위해 기존 방법과 비교하여 주요 데이터셋에서 성능

을 평가하고, 자연어 처리(NLP) 태스크에서 성능 개선 여부를 확인한다. 또한, 모델의 효율성을 측정하기 위해 파라미터 수 감소율과 학습 속도 개선 여부를 분석한다.

MTL 기법의 안정성을 향상시키는 것도 중요한 목표 중 하나이다. 이를 위해 기존 MTL의 문제점인 과적합(overfitting) 및 부정적 전이(negative transfer) 현상이 감소하는지 분석한다. 다양한 작업에서의 일반화 성능을 검증하여 제안된 방법이 보다 널리 적용될 수 있는 가능성을 평가한다. 연구 결과를 바탕으로 논문을 작성하여 학계에서 채택될 가능성을 높이고, 연구의 영향력을 확대하는 것을 목표로 한다.

본 연구를 통해 MTL의 기존 문제점을 해결하고, 보다 안정적이며 효율적인 학습 방법을 제안하여 학계와 산업계에 기여할 수 있기를 기대한다.

5. 프로젝트 관련 학습 계획

학습할 내용	기간	역할 분담
Mult-Task Learning, Soft Sharing, Hard Sharing	2025/03/18 ~ 2025/03/31	김수영, 송재현
Sparse Sharing, Nash MTL	2025/04/01 ~ 2025/04/14	김수영, 송재현
실험에 사용할 논문의 데이터셋 분석 및 정량적 성능 평가	2025/04/15 ~ 2025/05/12	김수영, 송재현
연구 진행 과정에 따라 필요 시 추가적인 학습을 유동적으로 진행	2025/05/13 ~ 2025/05/28	김수영, 송재현

6. 연구 일정 계획

조사할 내용	기간	역할 분담
MTL을 적용할 논문 탐색 및 실험 준비	2025/03/18 ~ 2025/03/24	김수영, 송재현
기초적인 Mult-Task Learning 구현 및 실험	2025/03/25 ~ 2025/04/14	김수영, 송재현
실험 및 결과 분석	2025/04/15 ~ 2025/05/12	김수영, 송재현
추가 실험	2025/05/13 ~ 2025/05/28	김수영, 송재현

Related Work Summary Table

번호	연구 제목(저자)	저널/컨퍼런스(연도)	주요 내용 요약	주요 인사이트
1.	Multitask Learning - Rich Caruana	Machine Learning July 1997 volume 28, pages 41-75	MTL의 효과와 다양한 적용 사례를 중심으로 MTL의 작동원리를 서술하며 MTL의 효과를 입증하고 있다. 또한, STL에서는 학습할 수 없는 특징을 학습할 수 있다는 특징을 제시하며 MTL 적용 시 고려해야 할 다양한 이슈들 역시 논의하고 있다.	MTL은 STL보다 성능이 우수하며, 다양한 작업에서 효과적인 작업 간 공유된 표현 학습이 핵심 개념 다양한 Machine Learning 알고리즘을 적용 가능 작업 간 관계를 학습하며, 명시적인 관계 설정 없이도 내재된 연관성을 발견할 수 있음
2.	An Overview of Multi-Task Learning in Deep Neural Networks - Sebastian Ruder	arXiv:1706.05098v1 [cs.LG] 15 Jun 2017	딥러닝 기반 MTL에 대한 전반적인 개요를 제공하고 있으며, 발전 동향을 소개하고 있다. 최근에 제안된 방법들인 Uncertainty weighting, tensor factorization, Sluice Net works 등의 더 발전된 MTL을 소개하고 있다.	MTL은 Deep Learning의 분야인 NLP, 음성인식, 컴퓨터 비전 등 다양한 분야에서도 효과적임 MTL의 주요 작동 메커니즘은 암묵적 데이터 증강, 표현 편향, 정규화임. 손실 함수 가중치 조정, 텐서 분해, Sluice Networks 등 발전된 기법 활용 가능.
3.	Multi-Task Learning as a Bargaining Game - Aviv Navon, Aviv Shamsian, Idan Achituve, Haggai Maron, Kenji Kawaguchi, Gal Chechik, Ethan Fetaya	arXiv:2202.01017v2 [cs.LG] 8 Jul 2022	MTL에서 발생하는 여러 작업의 gradient 간의 충돌 문제를 해결하기 위해 게임 이론의 개념을 도입한 연구이다. 작업 별 gradient 간의 충돌을 협력적 협상 게임으로 새롭게 해석하고, 이 문제를 해결하기 위해 게임 이론의 Nash Bargaining Solution을 도입하여 Nash-MTL이라는 알고리즘을 개발하였고 이에 대한 이론적, 실험적 우수성을 입증하고 있다.	MTL에서 발생하는 작업 간 Gradient 충돌 문제를 해결하기 위해 게임 이론을 활용한 접근이 효과적임을 입증함으로써 향후 MTL 최적화 연구에 중요한 방향성을 제시