



이승한 Seunghan Lee

1996.07.29 (만 27세) 군필

(학사) 연세대학교 경영학과 / 응용통계학과

(석.박 통합) 연세대학교 통계데이터사이언스 대학원

T. 010-8768-8472

E. seunghan96@naver.com

B. <https://seunghan96.github.io/>

Education

학부: 연세대학교(서울)

2015.03 ~ 2020.08

- 전공: 경영학과 / 응용 통계학과
- 학점: (전체) 4.19 / (경영) 4.31 / (응통) 4.35
- 기타: **조기 졸업 (7학기) + 최우등 졸업 (학년 별 상위 1%)**

석박사 통합: 연세대학교(서울)

2020.09~

- 전공: 통계데이터사이언스 학과
- 학점: (전체) 4.29

[Lab]

지도 교수님 (공동 지도)

- 박태영 교수님 (<https://dslab-with.github.io/web/>)
- 이기복 교수님 (<https://ml.yonsei.ac.kr/>)

[연구 분야]

Deep Learning with Time Series Data

Representation Learning / Self-supervised Learning

Diffusion Models, Graph Neural Network with Time Series

Publications (*Equally Contributed)

1. Improving Gibbs Sampler (2021)

- Authors : Taeyoung Park, Seunghan Lee
- Journal : *WIREs Computational Statistics* (<https://onlinelibrary.wiley.com/doi/abs/10.1002/wics.1546>)

2. MAD-GL2: Multimodal Adaptive Dynamic Graph Learning with Global and Local Features for Multivariate Time Series Forecasting (2023)

- Authors: Seunghan Lee*, Kibok Lee*, Taeyoung Park
- Journal : *IEEE TNNLS (Transactions on Neural Networks and Learning Systems)* (under review)

3. HiSS: Hierarchical Multi-Task Learning with Self-Supervised Auxiliary Task (2023)

- Authors: Seunghan Lee*, Taeyoung Park
- Journal : *JKSS (Journal of the Korean Statistical Society)* (under review)

4. Soft Contrastive Learning for Time Series (2024)

- 저자 : Seunghan Lee, Taeyoung Park, Kibok Lee
- Conference : *ICLR, 2024*. **Spotlight (366/7262=5%)** [arXiv:2312.16424]
- Preliminary version was presented in *NeurIPS Workshop on Self-Supervised Learning: Theory and Practice, 2023*.

5. Learning to Embed Time Series Patches Independently (2024)

- 저자 : Seunghan Lee, Taeyoung Park, Kibok Lee
- Conference : *ICLR, 2024*. [arXiv:2312.16427]
- Preliminary version was presented in *NeurIPS Workshop on Self-Supervised Learning: Theory and Practice, 2023*. **Oral (4/80=5%)**

사용 가능 언어, 자격증, 수상

[사용 가능 언어 (사용 능력)]

- (능숙) Python (Pandas.Numpy / Sklearn / Tensorflow.Pytorch), R
- (평범) SQL
- (기초) Docker, Kubernetes

[자격증]

분야	자격증	점수	취득년월
컴퓨터 / 데이터 분석	ADSP 데이터 분석 준전문가	합격	2018.09
	컴퓨터 활용능력 검정시험 1급		2019.03
	SQLD (SQL Developer)		2020.06
	빅데이터 분석기사		2021.07
	AWS Cloud Practitioner		2021.08
외국어	TOEIC	965점	2020.08
기타	한국사 능력검정시험	1급	2012.08
	한국경제신문 TESAT	S급	2013.05

[수상]

분야	수상 내역	등수	수상년월
교내	성적 우수상		2018.12 / 2019.06
	최우등 졸업		2020.08
데이터 사이언스	연세대 데이터사이언스 경진대회	최우수상 (2등)	2019.11
	Sony Pictures 영화 관람객 수 예측 대회	장려상 (4등)	2020.01
	연세대 데이터사이언스 경진대회	최우수상 (2등)	2021.12
	2022 동계 BK 학술 컨퍼런스	장려상 (4등)	2022.12
	연세대 데이터사이언스 경진대회	대상 (1등)	2023.02
기타	한국경제신문 TESAT	단체 우수상	2013.05

- 기타 참가 대회 : 카카오 아레나 / 빅콘테스트 / Dacon / IGAWorks 클릭률 예측 등

1. [인턴] SK Telecom 체험형 인턴

- 기간 : 2019.07~2019.08 (2개월)
- 사업부: ICT Infra 센터 > 대직무: Data Science > 중직무: Data Analytics (부서 : AI Home)
- 업무 : SKT에서 출시한 에브리웨어 (공기 질 정보 제공 서비스)에서 데이터 분석 및 서비스 기획 진행
 - * 분석 1) 전국에 위치한 SKT 대리점에 있는 공기측정기와 공공 대기관측소의 공기 데이터 분석을 통해 SKT 대리점들에 설치된 공기측정기의 오차 보완
 - * 분석 2) 앱의 주요 타겟층이 될 30대 초.중반의 학부모의 데이터 분석을 통한 서비스 기획
 - * 분석 3) 앱의 주요 타겟층이 될 유치원 원장님들에게 발행할 공기질 Report 기획 (콘텐츠 구성 + 시각화)

2. [인턴] 연세대 계산과학공학과 대학원 연구실 인턴

- 기간 : 2019.12. ~ 2020.02 (3개월)
- 연구 분야 : Data Science & Deep Learning > Network Embedding
- Network Embedding 관련된 논문 구현 및 이와 관련된 세미나 진행
- 사용 언어 : Python

3. [산학 협력] Amore Pacific 산학협력 프로젝트

- 기간 : 2020.02 ~ 2022.12 (3년)
- 분석 주제 : 아모레 퍼시픽 유저 및 구매 내역 분석

[2020 project]

- project 1) 오프라인 / 온라인 / 옴니 고객의 구매 성향 분석 및 옴니 고객 세분화
- project 2) 방문 판매를 통한 헤어/바디 제품의 매출액 예측
- project 3) 웹 로그 데이터 분석을 통한 온라인 유저들의 성향 분석 및 장바구니 분석
- 주요 사용 알고리즘 : Random Forest, Logistic Regression, Hierarchical Clustering,
- 사용 언어 : R

[2021-22 project]

- project) 문맥적 브랜드 점수 (Context Brand Score, CBS) 개발
- 텍스트 데이터 (소셜 버즈,리뷰)를 사용하여, 소비자의 화장품 제품의 속성별 감정을 고려한 브랜드 지수 개발
- 주요 사용 알고리즘 : Hierarchical BERT, ABSA (Abstract Based Sentiment Analysis), KoBERT
- 사용 언어 : Python

4. [프로젝트] 현대 모비스 서비스 부품 수요 예측

- 기간 : 2022.06 ~ 2022.12
- 내용 : 저순환 및 친환경 부품의 수요에 대한 장/단기 예측
- 사용 알고리즘 : Graph Neural Network (GNN), Domain Adaptation with GAN, Self-Supervised Learning with Tabular Data
- 사용 언어 : Python

5. [프로젝트] 연세대학교 대기과학과 데이터 분석 및 교육

- 기간 : 2020.11 ~ 2021.02
- 대기과학과 대학원생의 통계/군집화 교육 및 분석 보조
- 보조 내용 : 기상데이터를 전처리 및 차원축소/군집화
- 사용 알고리즘 : SOM(Self-Organizing Map) & K-Means
- 사용 언어 : Python

6. [강의 조교] SK Hynix 임직원 대상 Bayesian Statistic & ML 강의 조교

- 기간 : 2021.04 ~
- 업무 : Bayesian Statistics/ML 관련 코딩 실습 강의 진행
- 사용 언어 : R

7. [교육 강사] 국민 은행 Computer Vision / GAN 교육 강사

- 기간 : 2021.07
- 내용 : Computer Vision (CNN basics, LeNet, AlexNet, ZFNet, VGG16, Inception, ResNet, DenseNet) & GAN (GAN basics, DCGAN, CGAN, LSGAN)
- 사용 언어 : Python (Tensorflow2 & Pytorch)

8. [강의 조교] Python 프로그래밍과 웹 크롤링 조교

- 기간 : 2021.08, 2022.01, 2022.03
- 내용 : 파이썬 기초 문법 및 웹 데이터 수집
- 사용 언어 : Python

9. [교육 강사] 한화 오션 데이터 분석 및 통계 교육 강사

- 기간 : 2023.07
- 내용 : 기초 통계 분석 (상관관계분석, 회귀분석, 시계열 분석)

활동 동아리 및 학회 소개

동아리/학회 명	소개	활동 기간	비고
ISSU (이슈)	연세대학교 IT경영 전략학회	2018.03~2018.12 (2학기)	다수기업과 산학협력 진행 python 멘토로 활동
Bitamin (비타민)	빅데이터 분석 동아리	2018.07~2019.06 (2학기)	-
데이터사이언스 랩	연세대학교 데이터사이언스 학회 (통계학 기반의 데이터사이언스)	2019.01~2020.06 (3학기)	1기 회장

(학회) ISSU (연세대학교 IT경영전략 학회)

- 기간 : 2018.03 ~ 2018.12
- 정규 세션) 다양한 IT 기업들과의 산학협력 진행 (네이버 파파고, 네이버 바이브, 샌드애니웨어, 아만다)
- 스터디 세션)
 - (1학기) IT 조선에 “예술분야의 AI”를 주제로 기고 (http://it.chosun.com/site/data/html_dir/2018/06/28/2018062802823.html)
 - (2학기) 파이썬 멘토로 학회원들에게 기초 파이썬 가르침

(동아리) BITAMIN (빅데이터 분석 동아리)

2018.08 ~ 2019.03

- 활동 내용 : Machine Learning & Textmining with R & Python
- 세 차례의 머신 러닝 프로젝트 진행

(학회) Data science Lab (연세대학교 데이터 사이언스 학회)

2019.02 ~ 2020.06

- 1기 회장 (2019.06~2019.12)
- 데이터 수집부터 전처리, 모델링까지 전체 과정 학습
Machine Learning & Deep Learning 관련 정규 세션 및 스터디 진행
관련된 주제로 조별로 다양한 프로젝트 진행 및 대회 참가
- 학회원들을 대상으로 세션(강의) 진행
(강의한 내용 : Data Preprocessing, Association Analysis, Support Vector Machine, Boosting, CNN)

(스터디) OWOP (One Week One Paper)

2021.01 ~ 2021.06

- 연세대학교 통계데이터사이언스학과 대학원생들(5명)과의 논문 스터디
- 주제) AI/통계 관련 자유 주제 (CV, NLP, BNN, Rec Sys ...)
- 방식) 매 주 하나의 주제(or 한 편의 주제)로 돌아가면서 발표
- 주요 발제 주제) Variational Inference, Deep Generative Model (VAE, Normalizing Flow)

(스터디) Deep Learning Paper Reading Study

2021.05 ~ 2021.09

- 연세대학교 통계데이터사이언스학과 대학원생들(5명)과의 논문 스터디
- 주제) 한 주에 4개의 주제의 논문 한편 씩 총 4편의 논문에 대해 읽은 후 발제
 - 주제 1) Meta Learning
 - 주제 2) Continual Learning
 - 주제 3) Interpretable Learning
 - 주제 4) Reliable Learning

1. [Dacon] 2019 상반기 프로야구 타자 성적 예측 대회

- 기간 : 2019.01 ~ 2019.02
- 2019년 상반기 KBO 프로야구 타자의 성적 예측하기
- 사용 알고리즘 : MA, ARIMA, Random Forest

2. [연세 데이터 사이언스 경진대회] Embrain

- 기간 : 2019.03 ~ 2019.05
- 결과 : 본선 진출
- 주제 : 간단한 자기관리 족을 위한 화장품 MKT 전략 수립

3. [연세 데이터 사이언스 경진대회] PACKUS

- 기간 : 2019.09 ~ 2019.11
- 결과 : 최우수상 (2등)
- 주제 1) PACKUS 고객군 세분화 및 맞춤형 전략 수립
주제 2) 주요 판매 상품에 대한 향후 판매량 예측
- 사용 알고리즘 : RNN, XGBoost, LightGBM, SOM(Self-Organizing Map), Association Analysis

4. [Sony Pictures] 영화 관람객 수 예측

- 기간 : 2019.11 ~ 2019.01
- 결과 : 장려상 (4등)
- 주제 : 개봉 이후 2주(14일) 간의 누적 관람객 수 예측
- 사용 알고리즘 : Random Forest, XGBoost, Stacking, Neural Network

5. [IGA Works] CTR 예측

- 기간 : 2019.12 ~ 2019.02
- 주제 : 특정 고객이 특정 광고에 노출 되었을 때 클릭을 할지에 대해 예측하기
- 사용 알고리즘 : Deep CTR

6. [Kakao] Melon Playlist Continuation

- 기간 : 2019.05 ~ 2019.07
- 주제 : 노래 플레이리스트에 수록된 곡과 태그의 일부가 숨겨져 있을 때, 해당 곡과 태그들을 예측하기

7. [Dacon] 심리성향 예측 AI 경진대회

- 기간 : 2020.09~2020.11
- 주제 : 심리학 테스트 분석 알고리즘 개발 & 이를 통한 투표 여부 예측
- 사용 알고리즘 : AutoML, DeepCTR

8. [빅콘 테스트] NS홈쇼핑 편성표 최적화

- 기간 : 2020.07~2020.10
- 주제 : 홈쇼핑 판매 데이터를 사용하여, 매출을 극대화 하는 편성표 계획 + 매출액 예측
- 사용 알고리즘 : Catboost, Light GBM, Bayesian Optimization

9. [현대중공업] 현대중공업 BIG DATA/AI 경진대회

- 기간 : 2021.01~2021.02
- 결과 : 18팀 / 284팀 (본선 진출)
- 주제 : 제조 공정 작업시간 예측 및 작업 최적 할당 알고리즘 개발
- 사용 알고리즘 : Bayesian Neural Network, LightGBM

10. [연세 데이터 사이언스 경진대회] 효돌

- 기간 : 2021.09~2021.12
- 결과 : 최우수상 (2등)
- 주제 : 효돌 고객 군집화 및 고객 별 집중 관리 시간대 예측
- 사용 알고리즘 : Multi-task Learning, Self-Organizing Map

11. [Conference] 2022 동계 BK 학술 컨퍼런스

- 기간 : 2022.12
- 결과 : 장려상 (4등)
- 주제 : Spatio-Temporal GNN을 사용한 Multivariate Time Series Forecasting 알고리즘

12. [연세 데이터 사이언스 경진대회] KCB

- 기간 : 2022.12~2023.02
- 결과 : 대상 (1등)
- 주제 : 자기지도학습(Self-supervised Learning)을 활용한 개인 및 사업자의 다중 태스크 수행
 - (회귀) 신용점수 예측
 - (분류) 신용 등급 예측
 - (시계열 예측) 향후영업금액 예측
 - 군집화
- 사용 알고리즘 : Self-supervised Learning with Tabular Data, VIME (Value Imputation and Mask Estimation), Logistic Regression