Seunghan Lee

- Date of Birth: July 29, 1996
- Country: Republic of Korea
- Military Service: Completed
- Contact Info.:
  - Email: seunghan9613@yonsei.ac.kr
  - Linkedin: https://www.linkedin.com/in/seunghan-lee-0bba49169/
  - Website: seunghan96.github.io (1400+ AI-related posts)

1. Education

Undergraduate

Yonsei University, Seoul (March 2015 – August 2020)

- Majors: Business Administration / Applied Statistics
- GPA (4.5 scale): Overall 4.19 (Business 4.31, Applied Statistics 4.35)
- Additional Information:
  - Early graduation (7 semesters)
  - Summa Cum Laude (Top 1% of the cohort)

Integrated Master's & PhD Program

Yonsei University, Seoul (September 2020 – (Expected) August 2025)

- Major: Statistics and Data Science
- GPA (4.5 scale): 4.29
- Co-advisors:
  - Prof. Taeyoung Park (DSLab)
  - Prof. Kibok Lee (ML Lab)

2. Research Interests

- Deep Learning with Time Series (TS) Data
- Representation Learning, Diffusion Models
- LLM for TS, TS Foundation Models, Financial TS Modeling
- Vision-Language Models, Multimodal Learning

3. Publications (*Equally Contributed, †Co-corresponding Authors,)

10. Channel Normalization for Time Series Channel Identification (2025)

- Seunghan Lee, Taeyoung Park†, Kibok Lee†
- In *ICML*, 2025. []

9. Soft Contrastive Learning for Irregular Time Series (2025)

- Junghoon Lim, Seunghan Lee, Taeyoung Park
- Under review

8. Sequential Order-Robust Mamba for Time Series Forecasting (2024)

- Seunghan Lee*, Juri Hong*, Kibok Lee†, Taeyoung Park†
- In *NeurIPSW (Time Series in the Age of Large Models)* , 2024. [arxiv:2410.23356]

7. Partial Channel Dependence with Channel Masks for Time Series Foundation Models (2024)

- Seunghan Lee, Taeyoung Park†, Kibok Lee†
- In *NeurIPSW (Time Series in the Age of Large Models)*, 2024. Oral Presentation (Top 5) [arxiv:2410.23222]

6. Adaptive Noise Schedule for Time Series Diffusion Models (2024)

- Seunghan Lee, Kibok Lee†, Taeyoung Park†
- In *NeurIPS*, 2024. [arxiv:2410.14488]
- Outstanding Paper Award from the Journal of Korean Artificial Intelligence Association

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

- Seunghan Lee, Taeyoung Park, Kibok Lee
- In *ICLR*, 2024. [arxiv:2312.16427]
- Preliminary version: *NeurIPSW (Self-Supervised Learning: Theory and Practice)*, 2023. Oral Presentation (Top 4)

4. Soft Contrastive Learning for Time Series (2024)

- Seunghan Lee, Taeyoung Park, Kibok Lee
- In *ICLR*, 2024. Spotlight (366/7262=5%) [arxiv:2312.16424]
- Preliminary version: *NeurIPSW (Self-Supervised Learning: Theory and Practice)*, 2023.

3. Hierarchical Multi-Task Learning with Self-Supervised Auxiliary Task (2024)

- Seunghan Lee, Taeyoung Park
- In *The Korean Journal of Applied Statistics 37(5): 631-641*, 2024.

2. MAD-GL2: Multimodal Adaptive Dynamic Graph Learning with Global and Local Features for Multivariate Time Series Forecasting (2025+)

- Seunghan Lee*, Kibok Lee*, Taeyoung Park
- Under review

1. Improving Gibbs Sampler (2022)

- Taeyoung Park, Seunghan Lee
- In *Wiley Interdisciplinary Reviews: Computational Statistics 14(2): e1546*, 2022.
- Number of Citations (Google Scholar, 2025.02.16): 11

4. Careers

1. Naver Cloud (internship)

- **Period**: March 2025 –
- **Division**: HyperClova X
- **Tasks**: Multi-modal Pretraining.
    - ① Model training tasks for adding vision capabilities to a multimodal backbone
    - ② Exploration of the impact of various multimodal backbones on performance
    - ③ Efficient pretraining recipe exploration
- **Key Algorithms**: LLaVA, Qwen, DeepSeek
- **Programming Language**: Python

2. SK Telecom (internship)

- **Period**: July 2019 – August 2019
- **Division**: ICT Infra Center > Data Science > Data Analytics (AI Home)
- **Tasks**: Data analysis and service planning for SKT's air quality info service.
    - ① Analyzed air quality data and addressed measurement errors.
    - ② Designed strategies based on data analysis of parents in their 30s.
    - ③ Planned air quality reports, including data visualization.
- **Key Algorithms**: Clustering, DNN
- **Programming Language**: Python

3. Yonsei Univ. Department of Computational Science & Engineering

- **Period**: December 2019 – February 2020
- **Division**: Data Science & Deep Learning > Network Embedding
- **Tasks**: Implemented research papers and conducted related seminars.
- **Key Algorithms**: Graph Neural Network
- **Programming Language**: Python

5. Industry-Academia Collaboration

1. Amore Pacific

- Period: February 2020 – December 2022
- [2020 Project] User and purchase behavior analysis for Amore Pacific.
  - ① Analyzed purchasing behavior and conducted customer segmentation.
  - ② Predicted sales of products.
  - ③ Analyzed web log data to understand online user behavior patterns.
  - Key Algorithms: Random Forest, Logistic Regression, Hierarchical Clustering
- [2021-2022 Project] Development of Context Brand Score (CBS).
  - Developed a brand index that considers consumer sentiment on cosmetic product attributes using text data (social buzz, reviews).
  - Key Algorithms: Hierarchical BERT, Aspect-Based Sentiment Analysis.
- Programming Language: Python, R

2. Hyundai Mobis

- Period: June 2022 – December 2022
- Tasks:
  - ① Short- and long-term demand forecasting for low-circulation and eco-friendly parts of vehicles (Zero-inflated TS).
  - ② Utilizing textual data and tabular data for TS forecasting.
- Key Algorithms: GNN, GAN, Domain Adaptation, SSL with Tabular Data
- Programming Language: Python

6. Invited Talks

- 2024.08. Korean Artificial Intelligence Association (2024 Summer Conference)

7. Intellectual Property

- 2023.10: Review evaluation method using review quality metrics
- 2024.12: Universal time series analysis method and system for irregular multivariate time series

8. Programming Languages

- Proficient: Python (Pytorch, Tensorflow, Numpy, Sklearn)
- Intermediate: SQL, R
- Basic: Docker, Kubernetes

9. Instructor, Teaching Assistant (TA)

1. [Instructor] Yonsei University Atmospheric Science Department – Data Analysis

- Period: November 2020 – February 2021
- Tasks: Provided statistical and clustering analysis training for graduate students and assisted with data analysis.

2. [Instructor] KB Bank – Computer Vision / GAN

- Period: July 2021
- Tasks: Provided a lecture on computer vision and generative models

3. [Instructor] Hanwha Ocean – Data Analysis & Statistics

- Period: July 2023
- Tasks: Provided a lecture on below contents regarding statistical analysis

4. [TA] Python Programming & Web Crawling

- Period: March, September 2021,2022,2023,2024
- Tasks: Teaching assistant for Python programming and web crawling

5. [TA] SK Hynix – Bayesian Statistics & ML Course

- Period: April 2021 – June 2021
- Tasks: Conducted coding exercises for Bayesian Statistics and ML.

10. Academic Society & Club

Name	Details	Period	Etc.
ISSU	Yonsei Univ. IT Management Strategy Society	2018.03 – 2018.12	Industry-academia collaboration projects. Python mentor.
Bitamin	Big Data Analytics Club	2018.07 – 2019.06	
Data Science Lab	Yonsei Univ. Data Science Society	2019.01 – 2020.06	The 1st President

1. ISSU (Yonsei University IT Management Strategy Society)

- **Period**: March 2018 – December 2018
- **Regular Sessions**: Conducted industry-academia collaboration projects with IT companies, including Naver Papago, Naver VIBE, Send Anywhere, and Amanda.
- **Study Sessions**:
  - **1st Semester**: Contributed an article on *AI in the Arts* to *IT Chosun* (Link).
  - **2nd Semester**: Served as a Python mentor.

2. BITAMIN (Big Data Analytics Club)

- **Period**: August 2018 – March 2019
- **Regular Sessions**: Focused on machine learning and text mining.
- **Study Sessions**: Conducted three machine learning projects.

3. Data Science Lab (Yonsei University Data Science Society)

- **Period**: February 2019 – June 2020
- **The 1st President**: Organized regular sessions and study groups on ML and DL.
- **Regular Sessions**:
  - ① Studied the entire data science workflow, from data collection and preprocessing to modeling.
  - ② Conducted training sessions for society members on:
    - Data Preprocessing, Association Analysis, SVM, Boosting, CNN
- **Study Sessions**: Group projects and data science competitions.

4. OWOP (One Week One Paper)

- Period: January 2021 – June 2021
- Tasks: Conducted a research paper study group with five graduate students from the Department of Statistics and Data Science at Yonsei University.
- Topics: AI and statistics-related subjects, including CV, NLP, BNN, Recommender Systems, Deep Generative Models (VAE, Normalizing Flow).

5. Deep Learning Paper Reading Study

- Period: May 2021 – September 2021
- Tasks: Conducted a research paper study group with five graduate students from the Department of Statistics and Data Science at Yonsei University.
- Topics: Meta Learning, Continual Learning, Interpretable/Reliable Learning

11. Certifications & Awards

Certifications

Field	Certifications	Score	Date
CS/Data	ADSP (Associate Data Science Professional)	-	2018.09
	Computer Proficiency Level 1		2019.03
	SQLD (SQL Developer)		2020.06
	Big Data Analyst		2021.07
	AWS Cloud Practitioner		2021.08
Linguistics	TOEIC	965/990	2020.08
Etc.	Korean History Proficiency Test	Level 1	2012.08
	TESAT (Korean Economic Newspaper)	S-Level	2013.05

Awards

Field	Awards	Rank	Date
Academic	Academic Excellence Awards		2018.12 / 2019.06
	Summa Cum Laude Graduation		2020.08
CS	Yonsei Data Science Competition	2nd	2019.11
	Sony Pictures Audience Prediction	4th	2020.01
	Yonsei Data Science Competition	2nd	2021.12
	BK Winter Academic Conference	4th	2022.12
	Yonsei Data Science Competition	1st	2023.02
Economy	TESAT Team Excellence Award	2nd	2013.05

12. Projects

1. Fine-tuning LLM

- Period: February 2025 - March 2025
- Tasks:
  - (1) Fine-tuning LLaMA with Q-LoRA (Single-GPU)
  - (2) Fine-tuning LLaMA with Q-LoRA & FFT (Multi-GPU)
    - Axolotl, code-based
    - Distributed training: MP, FSDP, ZeRO
  - (3) Fine-tuning LLaMA for the insurance domain
    - One-cycle project: Data construction - SFT - DPO - Inference
    - Extract text from PDFs using OCR, generate additional questions with WizardLM (i.e., evolving), and build SFT & DPO datasets

2. Hands-on practice with Ollama, RAG pipeline, Vector DB, and UI/UX tools

- Period: February 2025 - March 2025
- Tasks:
  - (1) Ollama + RAG (PDF upload) + ChromaDB + Streamlit
  - (2) Ollama + RAG (Repository) + ChromaDB + Chainlit
  - (3) Ollama + RAG (Repository) + ChromaDB + Gradio

13. Competitions

1. [Dacon] 2019 KBO Baseball Player Performance Prediction

- **Period**: January 2019 – February 2019
- **Tasks**: Predicting KBO baseball player performance for the first half of 2019.
- **Key Algorithms**: MA, ARIMA, Random Forest

2. [Yonsei Data Science Competition] Embrain

- **Period**: March 2019 – May 2019
- **Tasks**: Developing a marketing strategy for highly meticulous consumers.
- **Key Algorithms**: Random Forest, NN
- **Result**: 4th

3. [Yonsei Data Science Competition] PACKUS

- **Period**: September 2019 – November 2019
- **Tasks**:
  - ① Customer segmentation and personalized strategy development.
  - ② Future sales prediction for key products.
  - ③ Created a dashboard for CRM using R Shiny.
- **Key Algorithms**: RNN, XGBoost, LightGBM, SOM, Association Analysis
- **Result**: 2nd

4. [Sony Pictures] Movie Audience Prediction

- **Period**: November 2019 – January 2020
- **Tasks**: Predicting cumulative audience numbers for the first 14 days after a movie's release.
- **Key Algorithms**: Random Forest, XGBoost, Stacking, NN
- **Result**: 4th

5. [Kakao] Melon Playlist Continuation

- **Period**: May 2019 – July 2019
- **Tasks**: Predicting missing songs and tags in a given playlist.
- **Key Algorithms**: Collaborative Filtering, Spectrogram Analysis

6. [IGA Works] CTR Prediction

- **Period**: December 2019 – February 2020
- **Tasks**: Predicting click-through probability when a user is exposed to an ad.
- **Key Algorithm**: Deep CTR

7. [Dacon] AI Competition for Psychological Tendency Prediction

- **Period**: September 2020 – November 2020
- **Tasks**: Developing an algorithm for psychological test analysis and voter prediction.
- **Key Algorithms**: AutoML, Deep CTR

8. [BigCon Test] NS Home Shopping Schedule Optimization

- **Period**: July 2020 – October 2020
- **Tasks**: Optimizing home shopping schedules to maximize sales.
- **Key Algorithms**: CatBoost, LightGBM, Bayesian Optimization

9. [Hyundai Industries] Big Data/AI Competition

- **Period**: January 2021 – February 2021
- **Tasks**: Predicting manufacturing process task durations and optimizing task allocation.
- **Key Algorithms**: Bayesian Neural Network, LightGBM
- **Result**: Finalist (18th place out of 284 teams)

10. [Yonsei Data Science Competition] Hyodol

- **Period**: September 2021 – December 2021
- **Tasks**: Customer clustering and predicting optimal engagement times for personalized care.
- **Key Algorithms**: Multi-task Learning, Self-Organizing Map
- **Result**: 2nd

11. [Conference] 2022 Winter BK Academic Conference

- **Period**: December 2022
- **Tasks**: Multivariate time series forecasting using Spatio-Temporal GNN.
- **Key Algorithms**: Spatio-Temporal GNN
- **Result**: 4th

12. [Yonsei Data Science Competition] KCB

- **Period**: December 2022 – February 2023
- **Tasks**: Multi-task learning with self-supervised learning.
    - ① **Regression**: Credit score prediction
    - ② **Classification**: Credit rating prediction
    - ③ **Time Series Forecasting**: Future revenue prediction
    - ④ **Clustering**: Customer segmentation
- **Key Algorithms**: SSL with Tabular Data
- **Result**: 1st