Hongjun Lim

Machine Learning Research Engineer

(82) 010-9936-7525 | hongjun.lim@navercorp.com | limhj159.github.io | hongjun-lim-a60a01135

Education

KAIST (Korea Advanced Institute of Science and Technology)

Daejeon, South Korea

Master of Science, Graduate School of Culture Technology

2016–2018

- **Data Science Lab**
  - Thesis: Understanding comments on social news: analyzing comment-reading behavior and comparing textual similarity on facebook news posts
  - Advisor: Prof. Meeyoung Cha

Ajou University

Suwon, South Korea

Bachelor of Information Technology, Department of Digital Media

2012–2016

- **Undergraduate Research Student, Integrated Design Lab**
  - Project: NetSet: Interactive Visualization for Analyzing Sets in Large Networks
  - Advisor: Prof. Kyungwon Lee

Experience

NAVER Corp.

Seongnam, South Korea

Machine Learning Research Engineer, AiRSearch

Sep 2018 - Present

- Service
  - Launched recommendation systems in various domains, including news, video, and cafe.
  - *News Fairness*: Participated in the 2nd News Algorithm Review Committee as a one of news recommendation practitioners, and published news recommendation algorithms at the top-tier international conference (ICDE 22) and technical blog to hedge external risks.
- Research & modeling
  - *News Recommendation 2.0*: Designed and developed a news recommendation system 2.0 that introduces content-based personalized ranking features, social impact model, and diversification.
  - *Self-Attention Recommendation*: Researched on self-attention based models for personalized news recommendation.
  - *Session-based Recommendation*: Researched on user modeling based on the historical sessions for a short and long term recommendation using ELMo (Embeddings from Language Model).
  - *Content-based Filtering for Video Recommendation*: Built a domain-specific language model for video services and developed a content-based model.
  - *Clickbait News Detection*: Built a transformer-based model to classify articles into news and clickbait.
  - *News Quality Estimation 2.0*: Improved quality estimation (QE) which is a method for automatically assessing the quality of news articles without human intervention.
- Engineering
  - Designed and developed AiRSLab which is developed based on Python and PyTorch for reproducing and developing recommendation algorithms in a unified, comprehensive and efficient framework.
  - Designed and developed the architecture to enable ML continuous training and delivery using Airsflow and Kubeflow.
  - Designed and developed non-linear ranking framework that supports state-of-the-art ranking models based on deep learning techniques, and real time inference on C++ based serving layer.

NAVER Corp.

Seongnam, South Korea

Machine Learning Research Engineer Intern, Video Recommendation

Mar 2018 - Sep 2018

- Built a clustering ensemble based ANN (approximate nearest neighbor) model.
- Researched on video embedding with language model.
- Researched on topic modeling for video classification.

Onezero Soft, Inc. Research Institute *Seoul, South Korea*

Software Engineer *June 2015 - July 2016*

- Researched on ML-based automatic product mapping methods for managing total financial information.
- Designed and developed a system that monitors the company's products and provides new announcements or update notifications to customers using .NET Framework.
- Design, conduct, and report results from prototype that leverages modern web application frameworks.
- Developed and launched an Android application for jewelry sellers.

Additional Experience and Awards

- Instructor. Incheon National University. Data Programming. Fall 2022
- Technical Research Personnel for alternative military service at Naver R&D Center (2018-2021)
- 1st place among college students in AI R&D Challenge (Ministry of Science and ICT; 2017)

Papers

- ***Hongjun Lim**, *Yeon-Chang Lee, Jin-Seo Lee, Sanggyu Han, Seunghyeon Kim, Yeongjong Jeong, Changbong Kim, Jaehun Kim, Sunghoon Han, Solbi Choi, Hanjong Ko, Dokyeong Lee, Jaeho Choi, Yungi Kim, Hong-Kyun Bae, Taeho Kim, Jeewon Ahn, Hyun-Soung You, Sang-Wook Kim. AiRS: A Large-Scale Recommender System at NAVER News, In *proc. of the IEEE International Conference on Data Engineering (ICDE)*, 2022.
  *These authors contributed equally.
- Sungmin Cho, **Hongjun Lim**, Keunchan Park, Sungjoo Yoo, Eunhyeok Park. On the Overlooked Significance of Underutilized Contextual Features in Recent News Recommendation Models, *arXiv preprint arXiv:2112.14370*, 2021.
- Seunghyun Yoon, Kunwoo Park, Joongbo Shin, **Hongjun Lim**, Seungpil Won, Meeyoung Cha, and Kyomin Jung. Detecting Incongruity Between News Headline and Body Text via a Deep Hierarchical Encoder, In *proc. of the AAAI Conference on Artificial Intelligence (AAAI)*, 2019.
- Jaewoo Kim, Yui Ha, Seungchae Kang, **Hongjun Lim**, and Meeyoung Cha. Detecting Multiclass Emotions from Labeled Movie Scripts, In *proc. of the IEEE International Conference on Big Data and Smart Computing (BigComp)*, 2018.
- **Hongjun Lim**, Choongho Chung, Jihee Kim, Juho Kim, Sue Moon, and Meeyoung Cha. Changing News Media Landscape in South Korea, In *proc. of the Fourth Workshop on Social News On the Web (SNOW) co-located with the World Wide Web (WWW) Conference*, 2017.
- Heungseok Park, **Hongjun Lim**, Wonjae Lee, and Kyungwon Lee. NetSet: A Systematic Integration of Visualization for Analyzing Set Intersections with Network, In *proc. of the IEEE Pacific Visualization Symposium (PacificVis)*, 2017.
- Heungseok Park, **Hongjun Lim,** and Kyungwon Lee. NetSet: Interactive Visualization for Analyzing Sets in Large Networks, In *proc. of the Symposium on Visualization in Data Science (VDS) co-located with the IEEE VIS Conference*, 2015.

Patents

- Myeongill Shin, Yonghyun Lee, and **Hongjun Lim**. Method and system for managing total financial information, Korea Patent, No. 10-2016-0048820.
- **Hongjun Lim**, Heungseok Park, and Kyungwon Lee. Apparatus and Method for Interactive Visualization for Analyzing Sets in Large Networks, Korea Patent, No. 10-2015-0148286.

Skills

Languages Python, C++, SQL, Scala, Shell, Java, R

Frameworks PyTorch, Tensorflow, Scikit learn, Airflow, Kubeflow, Docker, Spark, Hive, Presto, Zeppelin