BYEONGSU YU

◉ Email: byeongsu.yu@gmail.com ◉ LinkedIn: linkedin.com/in/byeongsu-yu ◉ GitHub: github.com/byeongsuyu/

PREVIOUS POSITION

AI & Data Scientist, *Samsung Fire and Marine Insurance company, Seoul, S. Korea* Jun 2023 – Oct 2024

- Developed a tree-based Retrieval-Augmented Generation (RAG) model to automate liability assessments in car accidents, leading to a patent application submission.
- Pioneered the transformation of insurance policy documents into a JSON hierarchical format with hyperlinks, enhancing the accuracy of LLM-based information retrieval and response correctness.
- Translated MT-bench questions into Korean and evaluated GAUSS, an LLM, enhancing model validation processes.
- Generated synthetic datasets for all KCD8 codes to boost the precision of AI-powered disease code recommendation models based on symptoms.

TECHNICAL SKILLS

Languages: Python, C++, SQL, R ◉ Additional Software: PyTorch, Git, GCP Vertex AI, AWS Sagemaker

EDUCATION

PhD in Mathematics, *Texas A&M University (USA),* GPA: 4/4 (expected) May 2023

Dissertation topics: Combinatorial Commutative Algebra, Computational Algebraic Geometry, Polyhedral Geometry

Selected courses: **Deep Learning** ◉ Communication & Teaching ◉ **Math. Finance** ◉ **Stochastic Calculus**

Bootcamp mini-courses: **Data Science** ◉ **Machine Learning** ◉ **Statistics** ◉ **Natural Language Processing**

Wrote **5** academic papers, invited to **14** academic talks, and taught **5** courses such as calculus and discrete mathematics.

MS in Mathematics, *Seoul National University (S. Korea),* GPA: 3.83/4.3 Aug 2017

Dissertation topics: Topological Data Analysis, Mathematical Modeling

Bachelors in Business Admin., Mathematics, and Philosophy, *Yonsei University (S. Korea),* GPA 3.4/4.3 Aug 2015

SELECTED EXPERIENCE

Data Scientist (individual contractor), *ED KPOP(Counter Culture Inc.), Seoul, S. Korea (Hybrid)* Jul - Aug 2022

- Developed a Python package based on LSTM and Autoencoder PyTorch model to evaluate (90% precision) and captioning the performance of choreography by classifying time-series data with Open Pose and AWS SageMaker.

Participant, Deep Learning Bootcamp, *Neuromatch Academy (Remote)* Jul - Jul 2022

- Completed 3-week courses for convolutional neural network, natural language processing, reinforcement learning.
- 91% test accuracy in classifying the political orientation of news articles was achieved by training neural network models with pretrained BERT and Longformer layers using PyTorch with 255,000 news articles. (link)

Participant, PhD to Industry Bootcamp, *The Erdös Institute (Remote)* Sep - Nov 2021

- Completed 11-week courses about regression, classification, and neural networks for applications in industrial data.
- Applied machine learning and time series analysis via scikit-learn to predict flight delay times of El Paso and Houston.

Research Assistant, Math. Modeling, *Seoul National University, Department of Mathematics* Aug 2015 - Aug 2017

- Applied topological data analysis to show single-nucleotide polymorphism group differences from diabetics.
- Devised a method comparing two texts' style to replicate Scientific American's experiment revealing Robert Galbraith is pseudonym of J.K. Rowling.
- Created a metric to detect fraudulent log-in attempts using keyboard rhythm for Interezen's Fraud Detection solution.

Software Engineer Intern, *38 Securities LP, Quincy, MA (Remote)* Jan 2015 - Feb 2015

- Developed a real-time algorithmic option trading application in C++ using API from Interactive Brokers.

PROJECTS

StdPairs, *created and maintained by B. Yu* (link) Jul 2021 - Present

Python package for symbolic computation of monomial ideals of affine (non-normal) semigroups in SageMath.

TDAkit, *created by Kisung You and B. Yu and maintained by Kisung You* (link) Aug 2021 – Present

CRAN package for topological data analysis involving persistent homology and persistent landscape.

SELECTED PUBLICATIONS & PREPRINT

- Standard Pairs of monomial ideals over non-normal affine semigroups in SageMath, ***B. Yu,*** **Journal of Software for Algebra and Geometry** (2022)
- Shape-Preserving Dimensionality Reduction: An Algorithm and Measures of Topological Equivalence, ***B. Yu*** *and Kisung You,* arXiv e-print:2106.02096 (2021)