Yejin Hwang

361-800-2613 | yejincc99@gmail.com | LinkedIn | GitHub

EDUCATION

Texas A&M University, Corpus Christi — ***Expected Graduation: Dec 2025***

Master of Data Science | 4.0 GPA

Sungkyunkwan University — ***Aug 2023***

Bachelor of Culture&Technology/Food Science&biotechnology | 3.7 GPA

SKILLS

Python · R · SQL · scikit-learn · PyTorch · NLP · Time-series forecasting · Pandas · NumPy · Matplotlib · yfinance · Statsmodels

EXPERIENCE

ML Engineering Research Fellow | LAB — ***Sep 2024 - Present***

Texas A&M University-Corpus Christi — *Corpus Christi, TX*

- Built scalable time-series datasets & pipelines (50K+ records) for AI forecasting research
- Developed labeling and anomaly detection workflows for time-series and text data, improving classification accuracy by 25% and enabling more robust downstream predictive analysis

Teaching Assistant — ***Sep 2024 - Present***

Texas A&M University-Corpus Christi — *Corpus Christi, TX*

- Assisted in teaching College Algebra, Trigonometry, and Calculus courses through recitation sessions
- Held office hours at Math Learning Center, offering one-on-one support that enhanced student confidence

Clinical Data Analyst | CCE — ***Aug 2022 - Aug 2023***

Samsung Medical Center-CCE(Clinical Center for Epidemiology) — Seoul, *South Korea*

- Built 100+ patient data pipelines integrating CRF and EMS systems, reducing manual data entry time by 40% and improving data reliability for clinical research
- Managed and cleaned 1,000+ clinical datasets, coordinated workflows and protocols to ensure high-quality data for epidemiological studies
- Supported data collection and statistical analysis for rare and oncological diseases research, contributing to evidence-based insights into patient outcomes
- Collaborated with leading epidemiology researchers, including Prof. Juhee Cho (SAHIST/Johns Hopkins), gaining experience in international, cross-institutional research

Undergraduate Research Assistant | LAB — ***Mar 2022 - Jul 2022***

SKKU(Sungkyunkwan University) — Seoul, *South Korea*

- Conducted literature reviews and data handling under Prof. Jinhee Hur, Harvard Research Fellow
- Designed and prepared experimental setups following academic research protocols
- Analyzed datasets using Python, Excel, and statistical methods (e.g., regression analysis, hypothesis testing)
- Interpreted results to support conclusions and inform further experimental design

PROJECTS

Sentiment Analytics & ETL Pipeline | Individual Project — ***June 2025 - Sep 2025***

- Developed forecasting pipelines combining Reddit sentiment data with stock prices to study how online discourse influences collective behavior
- Evaluated sentiment classification with FinBERT, VADER, and DistilBERT, highlighting both the potential and limitations of NLP models in capturing human perception
- Automated SQL-based ETL from Reddit and finance APIs, enabling daily monitoring of sentiment-driven market dynamics
- Enhanced Temporal Fusion Transformer (TFT) model with sentiment features, reducing RMSE~60%(8.00→3.39) compared to no-sentiment baseline

Stock Forecasting with Transformer Models | Individual Project — ***Feb 2025 - Apr 2025***

- For 90-day prediction, reduced RMSE by 49 percent (ARIMA 83.22 to Chronos 42.24) and MAE by 46 percent (63.93→34.45)
- For 5-day prediction, reduced RMSE by 20 percent (ARIMA 11.27 to TimesFM 8.99) and MAE by 24 percent (9.64→7.37)
- Presented at Joint NMSU/UTEP Conference on Math, Computer Science & Computational Sciences
- Presented at Coastal Bend Mathematics and Statistics Conference

Activities

SKKU President's Award (Samsung Foundation) - Awarded ₩2M for biotechnology R&D project(Article) — ***Nov 2024***