Yejin Hwang

+1 361-800-2613 | yejincc99@gmail.com | linkedin.com/in/yejin-data | github.com/Yejin-Hwang

Education

Texas A&M University–Corpus Christi — Corpus Christi, TX
Master of Data Science; GPA: 4.0 — *Expected Dec 2025*

Sungkyunkwan University (SKKU) — Seoul, South Korea
B.A. in Culture & Technology, Food Science & Biotechnology(Double Major) ; GPA: 3.72 — *Aug 2023*

Technical Skills

Languages:Python, R, PostgreSQL
Frameworks: Scikit-learn, TensorFlow, PyTorch, Streamlit
Tools: Git, Tableau, Excel, VS Code, Jupyter Notebook, LaTeX
Libraries: Pandas, NumPy, Scipy, Yfinance, Matplotlib, Seaborn, Transformers, Statsmodels

Professional Experience

Research Assistant — Sep 2024 – Present
Texas A&M University–Corpus Christi — *Corpus Christi, TX*

- Led research on time-series forecasting of stock data (TSLA, AAPL, NVDA) using ARIMA, TimesFM, and Chronos-T5, focusing on short-term market prediction and outlier analysis.
- Engineered key features such as lag variables, rolling statistics, volatility measures, and pattern recognition techniques to detect anomalies and enhance forecast accuracy.
- Improved model performance by 30% MAE reduction, supporting more reliable data-driven investment strategies.
- – Built reproducible Python pipelines integrating yfinance API and PyTorch for automated data ingestion, model training, and reporting, maintaining attention to detail throughout each stage.
- Presented model comparison and results at major conferences, emphasizing real-world financial implications.

Clinical Data Specialist — Aug 2022 – Mar 2024
Samsung Medical Center — *Seoul, South Korea*

- Managed large-scale patient datasets, including physical, mental health, and medication records, by working directly with cancer patients and clinical staff.
- Built and maintained ETL pipelines to ensure high-quality, reliable data for clinical research and hospital decision-making.
- Collaborated with medical professionals to identify patient needs and improve data collection processes, supporting more effective and timely healthcare delivery.
- Applied biostatistics and data analytics techniques to provide actionable insights for multidisciplinary healthcare projects.
- Contributed to a culture of evidence-based medicine by enhancing data integrity and supporting high-impact research initiatives.

Data Science Projects

Tesla Sentiment Analytics & ETL Pipeline | *Python, Pandas, yfinance, VADER, SQL*

- Developed an automated ETL pipeline integrating Tesla stock price data (via yfinance API) with Reddit-derived sentiment scores using VADER.
- Engineered time-based and moving average features, handled missing values, and merged multi-source data for downstream analytics.
- Stored final datasets in both CSV and SQLite, facilitating reproducible research and dashboard-ready data for market analysis.

Stock Forecasting with Transformer Models | *Python, PyTorch, Hugging Face, TimesFM, Chronos-T5*

- Implemented end-to-end forecasting pipelines for TSLA, AAPL, and NVDA using classical and transformer-based models.
- Conducted comparative analysis and proposed ensemble strategies to maximize forecast robustness.
- Built custom Python modules for data ingestion, feature engineering, and interactive visualization; delivered reproducible code and results.

Presentations & Research

Joint NMSU/UTEP Conference on Math, Computer Science & Computational Sciences Apr 2025
Presented forecasting model comparison using Chronos-T5 and TimesFM on TSLA data *Las Cruces, NM*

Coastal Bend Mathematics and Statistics Conference Apr 2025
Presented research on transformer vs. traditional models for stock price prediction *Corpus Christi, TX*

Awards

SKKU Dean's Award for Excellence in Global Engagement Mar 2023
Recognized for top-level academic achievement and cross-cultural research initiatives *Seoul, South Korea*