

Doheon Kim

peterjhms1922@gmail.com | peterdokim.com | linkedin.com/in/dokim19 | github.com/peterdokim

EDUCATION

- University of California, San Diego** San Diego, CA
Bachelor of Science in Computer Science, Minor in Business; GPA: 3.40 September 2021 – December 2027

EXPERIENCE

- Kearney** Seoul, South Korea
Research Analyst Intern April 2026 – June 2026
 - Built a multi-platform data pipeline for YouTube, TikTok, Instagram, Facebook, and X that extracts captions or transcribes audio, converting video posts into text for a client's brand-monitoring analysis
 - Designed a two-stage Gemini classifier that filters out posts unrelated to six tracked smartphone brands, then tags each remaining post by content type and topic
 - Validated the classifier on 293 hand-labeled posts at 87.7% filtering accuracy, adding rules to correct overused tags
 - Scaled the classifier to 324,697 news articles with cheaper Gemini models while maintaining 93.3% precision
 - Automated the analysts' event calendar with a daily Cloud Run pipeline that scans 67 news sources, merges duplicate stories across languages, and writes one-line summaries into their Excel file
- Republic of Korea Army (KATUSA), 2-23 Infantry Battalion, 4th Infantry Division** South Korea
Sergeant; Expert Infantryman Badge October 2024 – April 2026
 - Served as rifleman and the U.S. battalion commander's translator, training at the Korea Combat Training Center
- K2 Investment Partners** Seoul, South Korea
Analyst Intern June 2023 – August 2023
 - Drafted the market and competitor sections of an investment memo on a design-tool startup competing with Figma
 - Wrote research memos on Korean ICT firms abroad, telecom open-innovation programs, and Seoul venture funds
 - Evaluated valuation models with DART and EDGAR peer financials, cross-checking against the research memos
- Acelab** Seoul, South Korea
Software Engineering Intern June 2022 – August 2022
 - Converted a Jupyter notebook for autonomous-driving detection tests into a PyQt app that plots ROS bag data
 - Added the app's first quantitative metric by measuring Mobileye camera and lidar position error against 1 cm-accurate GPS ground truth, aligning timestamps with linear interpolation

PROJECTS

- Message Board Sentiment and Retail Order Flow** Independent Two-Person Project
Python, PyTorch, Hugging Face, KLUE-RoBERTa, pandas, statsmodels June 2026 – Present
 - Wrote a resumable Python crawler that collected 1.02M Korean stock message board posts over 134 trading days, using 5 times fewer requests after finding an undocumented page-size limit
 - Fine-tuned KLUE-RoBERTa on 86K posts labeled by GPT-5-mini to classify sentiment as bullish, bearish, or neutral, reaching 0.739 macro-F1 versus GPT-5-mini's 0.745
 - Labeled 1.37M posts with the fine-tuned model at no API cost, doubling throughput with length-sorted batching
- Jeonse Safety Recommendation System** Snowflake Hackathon, Two-Person Team
Python, SQL, Snowflake (Snowpark), Streamlit, pandas April 2026
 - Built a Python loader that pulls two years of government apartment sale and jeonse (lump-sum deposit) records for all 25 Seoul districts, retrying failed API calls and loading rows into Snowflake without duplicates
 - Wrote the scoring engine that turns a 6-question survey into a risk profile and ranks 288 neighborhoods by deposit-loss exposure, price cushion, budget fit, and commute
- AI News Summarizer** Two-Person Project
Python, FastAPI, Next.js, OpenAI API, Vercel December 2025 – March 2026
 - Built the Next.js app and FastAPI job server for a tool that groups and summarizes Korean news articles

SKILLS

- ML & Data Science:** scikit-learn, PyTorch, Hugging Face Transformers, pandas, BERT fine-tuning, model distillation
- Cloud & Platforms:** Google Cloud (Cloud Run, Vertex AI, GCS), Snowflake, Vercel, Linux CLI
- AI Tooling:** Claude Code, Gemini, OpenAI API, Prompt Engineering, Embeddings
- Programming & Frameworks:** Python, SQL, JavaScript, C; Streamlit, PyQt, Next.js, FastAPI, Flask, ROS
- Business & Languages:** Excel, PowerPoint, Technical Writing; Korean (Native), English (Fluent)