Dae Hyun Lee

206-886-7537 | dlee23@uw.edu | linkedin | GitHub

Education

University of Washington — Seattle, WA
M.S. in Data Science — *(Expected) Sep 2025 – Mar 2027*

University of Washington — Seattle, WA
B.A. in Mathematics — *September 2019 – March 2025*

Technical Skills

Languages: Python, Java, R, SQL
ML/DL Frameworks: Scikit-learn, PyTorch, TensorFlow
Data Analysis: pandas, NumPy, Matplotlib, Seaborn, Tableau
Web Development: Dash, Plotly

Work Experience

Researcher — Oct 2025 – Current
Roy Lab, University of Washington

- Designing deep learning frameworks to operationalize cognitive reserve from multimodal neuroimaging data, addressing the fundamental challenge of quantifying latent cognitive resilience from observable clinical markers
- Conducting systematic literature review of cognitive reserve measurement methodologies, critically evaluating statistical approaches and implementing baseline models (linear regression, XGBoost) for benchmarking
- Training 3D convolutional neural networks on ADNI dataset (4,508 subjects with structural MRI and 127 clinical features) for three-way Alzheimer's disease classification (CN/MCI/AD), optimizing model architecture for diagnostic accuracy

Research Intern — Jun 2024 – Sep 2024
SNU VLDB Lab, Seoul National University

- Proposed and validated a z-score-based dimension reduction technique for embedding vectors, supported by mathematical formulation and statistical analysis
- Achieved 96% dimensionality reduction (1536→64 dimensions) while preserving 95%+ similarity accuracy, improving K-NN query efficiency
- Integrated algorithm into HNSW indexing system and conducted performance benchmarking using OpenAI embedding datasets (100K+ vectors)
- Authored internal technical reports including statistical analysis and algorithmic benchmarking

Projects

NFL Player Trajectory Prediction - Framework Comparison | *Kaggle* — Oct 2025 - Dec 2025

- Implemented structurally equivalent GRU-based encoder-decoder models in PyTorch and TensorFlow for NFL player trajectory prediction, controlling for architecture, hyperparameters, and random seeds
- Conducted rigorous comparative analysis revealing significant autoregressive inference divergence despite structural equivalence (PyTorch RMSE: 1.55 vs TensorFlow: 19.47 on validation set)
- Authored technical report contextualizing empirical findings with established literature on exposure bias and error accumulation, documenting implications for deep learning framework reproducibility
- Github Link

WatchDawg - Seattle Crime Analytics Dashboard | *University of Washington* — Sep 2025 - Dec 2025

- Developed interactive web dashboard analyzing 17+ years of Seattle Police Department crime data (500K+ incidents) with real-time filtering and geospatial visualization capabilities
- Engineered 11-stage data cleaning pipeline and memory-optimized architecture enabling stable deployment on resource-constrained environment (512MB RAM), reducing dataset noise by 35%
- Deployed production application on Render with 99%+ uptime, implementing user-centered design principles (Cooper's About Face, Bertin's visual encoding) for coordinated multi-dimensional filtering
- Github Link