

Hyeonsup Lim, Ph.D., Transportation Data Scientist

hslim8211@gmail.com | 865-236-4963 | Buford, GA | [LinkedIn](#) | [Google Scholar](#) | Permanent US Resident

SUMMARY OF QUALIFICATIONS

- Data scientist specializing in machine learning and statistical modeling for freight, supply chain, and logistics.
- Hands-on programming skills to develop practice-ready algorithms and data solutions.
- Strong academic background with Ph.D. in transportation engineering and M.S. in statistics.
- Proficient technical writing and presentation skills as demonstrated by journal papers and technical reports.

CORE SKILLS

- **Python (most frequently used)**
 - Daily programming language with +5 years of experience.
 - Frequently used for checking data quality and generating summary tables.
 - Implemented various predictive models, including, OLS, Lasso, Decision Tree, Random Forest, Gradient Boosting, Support Vector, Gaussian Process, MLP, Gaussian Mixture, etc.
 - Used various libraries: pandas, NumPy, SciPy, scikit-learn, Geopandas, NetworkX, NLTK, OpenCV, BeautifulSoup, etc.
- **Web Development (JavaScript, ASP.NET, and MySQL)**
 - Full-stack development experience for a web-based data tabulation/visualization tool.
 - Utilized various open-source libraries (D3.js, PivotTable.js, Plotly) for the visualization/tabulation.
- **MATLAB (used to be go-to programming, but not recently used)**
 - Developed a self-learning license plate matching algorithm, error distribution estimation of multi-source data for connected vehicles, and other +30 data analysis tools.
- **Excel VBA (used to be go-to programming, but not recently used)**
 - Mainly used for re-organization of small-sized data and simple test simulations.
 - Capable of developing most of such functions within a couple of hours.
- **Other Software/Tools**
 - QGIS, ArcGIS, MS Access, SAS, R, JMP, and SPSS Modeler
 - Diverse experiences in transportation modeling and simulation tools (e.g., VISSIM and TransCAD)

PROFESSIONAL EXPERIENCES

- **Oak Ridge National Laboratory | R&D Staff (2019 - Present)**
 - Managing [Freight Analysis Framework 5](#) (FAF5, +\$3M project to provide public freight movement data among states and major metropolitan areas by transportation modes).
 - Developed key data products and [web-based data tabulation tool](#) for the FAF5 project.
 - Developed automated validation tools for FAF, reducing manual processing by 80%.
 - Designed and implemented a payload weight distribution algorithm for 100M+ truck records.
 - Built web-based visualization tools for [NextGen National Household Travel Survey \(NHTS\)](#).
 - Developed a sequential network imputation method that successfully imputes over 99% of missing roadway link data, leveraging less than 1% of observed truck volume data. ([the invention submitted to provisional patent application sponsored by the lab](#))
- **Bank of America | Assistant Vice President – Quantitative Finance Analyst (2017 - 2019)**
 - Validated financial risk models ensuring compliance with anti-money laundering regulations.
 - Applied machine learning for fraud detection, enhancing model efficiency.
- **University of Tennessee & Oak Ridge National Laboratory | Research Assistant (2013 - 2017)**
 - Developed a post-processing self-learning license plate algorithm, improving matching accuracy from 60~70% to over 90%.
 - Designed an integrated traffic simulation algorithm for real-time transport analysis.
 - Developed nationwide county-level energy commodity movement models.

- **Korea Smart Card Co., Ltd. | Business Analyst (2010 - 2013)**
 - Conducted global strategy assessments for Intelligent Transportation Systems (ITS).
 - Estimated passenger flow demand for transit systems.

EDUCATION

- **University of Tennessee, Knoxville, TN (2013 - 2017)**
 - Ph.D. in Transportation Engineering | M.S. in Statistics (GPA: 4.0/4.0)
- **Seoul National University, Seoul, S. Korea (2008 - 2010)**
 - M.S. in City Planning & Transportation (GPA: 3.8/4.3)
- **Ajou University, Suwon, S. Korea (2001 - 2008)**
 - B.S. in Transportation Systems Engineering (Ranked 1st in Major, GPA: 4.0/4.5)

DEVELOPED WEB-TOOLS

- Freight Analysis Framework 5 Data Tabulation Tool (public, sponsored by USDOT)
 - https://faf.ornl.gov/faf5/dtt_total.aspx
- National Household Survey NextGen OD Summary (public, sponsored by USDOT)
 - <https://nhts.ornl.gov/od/summary/>
- Truck Data and Modeling Visualization (pilot, presented in a conference)
 - https://lgagamel.github.io/TMAS-Data-Visualization/gmm_modeling.html
 - https://lgagamel.github.io/TMAS-Data-Visualization/gvw_distribution.html
 - https://lgagamel.github.io/TMAS-Data-Visualization/TMAS_WIM_coverage_2017.html

SELECTED PUBLICATIONS & PRESENTATIONS

- Lim et al., “A Comparative Study of Machine Learning Algorithms for Industry-Specific Freight Generation Model”, Sustainability, 2022, vol. 14, issue 22, 1-25
 - Implemented and fine-tuned seven machine learning models using Python scikit-learn library to optimize industry-specific freight generation predictions, achieving a 30% reduction in RMSE compared to traditional OLS regression.
- Lim et al., “A Data-Fusion Method using Bayesian Approach to Enhance Raw Data Accuracy of Position and Distance Measurements for Connected Vehicles”, 2021 IFIP/IEEE International Symposium on Integrated Network Management (IM), pp. 1018-1023.
 - Developed a Bayesian data-fusion model using Python SciPy library to enhance vehicle positioning accuracy, achieving over 98% R-squared in bias estimation of position and distance measurements.
- Lim et al., S. Chin, H. Hwang, and L.D. Han, “Incorporating Information Complexity into Regression-based Freight Generation Model Selection”, 2017 Transportation Research Board paper #17-06162 (selected as a “Practice-Ready” paper)
 - Developed an algorithm to evaluate and select optimal regression models for freight generation, enhancing predictive accuracy by accounting for information complexity.
- Liu et al. (Lim as second author), “Improving the accuracy of freight mode choice models: A case study using the 2017 CFS PUF data set and ensemble learning techniques”, 2024, Expert Systems with Applications 240, 12247
 - Provided conceptual framework and strategic guidance for machine learning-based freight mode choice modeling, mentoring a PhD student (first author) in implementing ensemble learning techniques to improve prediction accuracy by 19%.
- H. Lim, “Multi-Output Chaining for Time Series Forecasting to Estimate Impact of Supply Chain Disruption: Case of Russia’s Invasion of Ukraine on US Import and Export”, presented at 2023 Transportation Research Board Annual Meeting
 - Developed a multi-output chaining time series forecasting model using SARIMAX to estimate the impact of supply chain disruptions, integrating cross-correlations between industries to improve predictive accuracy for downstream sectors.