

Sunggu Park

sunggu0824@gmail.com | 010-4932-7157 | [linkedin.com/in/sunggu-park](https://www.linkedin.com/in/sunggu-park) | github.com/sunggu-park

EDUCATION

HBSc in Applied Statistics and Geospatial Data Science

University of Toronto

Sep 2018 - Jun 2025

Ontario, Canada

Relevant Courses: Introduction to Machine Learning, Space Time Data Analysis, Regression Analysis, Software Design, Spatial Data Science, Spatial Database, Surveys, Sampling and Observational Data, Experimental Design, Environmental Modeling

EXPERIENCE

Data Scientist Internship

SPORTS2i Co., Ltd.

May 2024 – July 2024

Seongnam, South Korea

- Developed a machine learning model to predict pitch types using KBO datasets and XGBoost, deployed in live baseball broadcasts to enhance viewer engagement.
- Addressed class imbalance issues (e.g., fastball dominance) through comprehensive feature engineering, boosting F1-score by 23% and accuracy by 8%.
- Reduced model training time by over 50% without compromising predictive performance via feature selection and hyperparameter optimization, significantly accelerating the experimentation cycle.

PROJECTS

Spatiotemporal Crime Analysis

GGR376 Final Project

Jan 2025 – Apr 2025

University of Toronto

- Applied a Spatial Panel Model (Spatial Error, random effects) in R to analyze assault crime across 140 Toronto neighborhoods (2015–2024), accounting for spatial and temporal dependence.
- Identified unemployment and housing cost burden as statistically significant predictors ($p < 0.01$) and built an interactive ArcGIS Dashboard to visualize spatial patterns and crime trends over time.

Transit Accessibility Analysis

GGR442 Capstone Project

Sep 2024 – Dec 2024

University of Toronto

- Analyzed transportation and demographic data in Alberta to identify transit-vulnerable areas and proposed optimal locations for infrastructure expansion through accessibility and suitability modeling.
- Automated the analysis workflow using ArcGIS Pro ModelBuilder, reducing turnaround time by 50% while maintaining reproducibility and data consistency.

City Prediction from Survey Data

CSC311 Final Project

Jan 2024 – Apr 2024

University of Toronto

- Analyzed survey data capturing public perceptions of global cities and developed a multi-class classification model to predict the target city based on respondents' answers, achieving 96% test accuracy and F1 score.
- Benchmarked diverse model families (linear, tree-based, and neural networks) and optimized performance using GridSearchCV with Stratified K-Fold cross-validation.

SKILLS

Programming Languages & Visualization: Python (Proficient), R (Proficient), ArcGIS Pro (Proficient), SQL, Java, Tableau, Power BI, Word, PowerPoint, Excel, LaTeX

Spoken Languages: English (Full Professional Proficiency), Korean (Native)