Sokratis Papadopoulos

sokpapadop@gmail.com | +1 646.705.3295

SKILLS

MACHINE LEARNING
Supervised learning • Clustering • Dimensionality reduction • Feature engineering • Probabilistic programming

STATISTICAL METHODS
Regression analysis • Time series analysis • Hypothesis testing • Experimentsl design • Panel data analysis • Causal inference • Bayesian methods • Spatial analysis

PROGRAMMING
Python (scikit-learn, pandas, numpy, scipy, gensim, pymc, beautifulSoup) • Spark • SQL • R • MatLab • SPSS

VISUALIZATION
Python (seaborn, matplotlib) • Tableau • Carto

COURSEWORK

GRADUATE
Machine Learning
Advanced Techniques in Data Science
Big Data Management and Analytics
Urban Informatics
Applied Statistics
Time Series Analysis

EDUCATION

NEW YORK UNIVERSITY
PHD CIVIL & URBAN ENGINEERING
Conc. Data Science/Urban Informatics
Expected May 2019 | New York, NY
Honors: Magna Cum Laude

MASDAR INSTITUTE
MS ENGINEERING SYSTEMS & MANAGEMENT
May 2015 | Abu Dhabi, UAE
Honors: Magna Cum Laude

LINKS

Web:// sokratispapadopoulos.com
Github:// spapadopoulos
LinkedIn:// sokratis-papadopoulos
Google Scholar:// Sokratis Papadopoulos

EXPERIENCE

INSIGHT DATA SCIENCE | DATA SCIENCE FELLOW
Jan 2019 – Present | New York, NY
- Created "AirbnBoost", a web app that enables users to make faster and more informed Airbnb decisions.
- Built a machine learning-based pricing model, crawled open data to engineer location intelligence features, and characterized listings using natural language processing.
- Designed a Flask app, deployed on Heroku cloud.

NYU CENTER FOR URBAN SCIENCE + PROGRESS | GRADUATE RESEARCH ASSISTANT
Sep 2016 – Present | New York, NY
- Contextualized and analyzed spatio-temporal patterns in large-scale building energy data using Python and Apache Spark.
- Developed machine learning models to predict urban scale energy consumption (~1M buildings) and understand its drivers.
- Created a city-specific energy index for more than 7,500 NYC's buildings in collaboration with the Mayor's office for Sustainability using XGBoost and KMeans clustering.
- Applied classic and Bayesian statistics for policy impact evaluation.
- Contributed in grant proposals and presented policy recommendations to project stakeholders.

MASDAR INSTITUTE | RESEARCH ASSISTANT
Aug 2013 – Sep 2016 | Abu Dhabi, UAE
- Used machine learning and agent-based modeling to quantify occupants' impact on building energy performance.
- Created an occupant behavior optimization framework based on machine learning and genetic algorithms. The tool speeds up the optimization process by 70% compared to brute-force simulation.
- Evaluated the accuracy and computational efficiency of time series methods and tree-based ensemble learning algorithms, such as Random Forest and Gradient Boosting, for electricity load forecasting.

INTELEN INC. | R&D INTERN
Jun 2014 – Aug 2014 | Athens, Greece
- Wrote scripts to automate building occupant energy behavior simulation using their feedback from questionnaires.
- Created visualizations of household energy data using Python's seaborn and matplotlib to increase occupants' awareness on their consumption.

AWARDS

2017 — United Nations, Data for Climate Action - Best visualization award (among 450 participating teams)
2016 — Siemens PhD fellowship recipient
2013 — Masdar Institute - Full Scholarship for MS studies