

LUCIA PALOVA

DATA SCIENTIST

Summary

Research physicist turned Data Scientist and Software Engineer, leveraging advanced math, statistics, and data manipulation methods to draw practical insights from data.

Experience

Metis Data Scientist

New York, NY
Apr 2016 to Jun 2016

Worked on data science projects, with focus on computational machine learning and statistical techniques, predictive modeling. Used data acquisition, data wrangling, visualization, cloud computing. Worked in a team environment, presented and discussed results with peers.

New Jersey Institute of Technology Adjunct Faculty, Department of Mathematical Sciences

Newark, NJ
Jul 2012 to Current

Responsible for managing and running undergraduate classes, including creating syllabus, teaching materials, exams, grading and working with classroom presentation equipment. Classes taught: Statistics and Probability, Mathematics of Finance I (Theory of Interest), Calculus.

Columbia University Postdoctoral Research Scientist, Energy Frontier Research Center

New York, NY
Aug 2010 to Feb 2013

Conducted research in computational and theoretical condensed matter physics. Described electronic structure of single-layered nitrogen- and boron-doped graphene. Published results in top peer reviewed scientific journals, presented results. Collaborated closely on research projects with other scientists.

Rutgers University Teaching and Research Assistant, Department of Physics and Astronomy

New Brunswick, NJ
Sep 2004 to Jun 2010

Researched condensed matter physics. Characterized ferroelectric materials in bulk, thin-film and nanostructure geometries using a combination of numerical and analytical techniques. Published results, presented at conferences. Taught general physics classes, worked with students.

Projects

Publications Recommendation System

Implemented a system to index scientific publications within an academic subfield, such as condensed matter-materials science. Used natural language processing tools, including nltk, sklearn and gensim libraries to process raw text documents. Applied vector space models, in particular, latent semantic indexing, and extracted topics, clustered publications. Recommended similar publications based on a text query or another document. Implemented network algorithms to rank publications. Built a demo flask application that allows the user to search for publications and explore the publication network.

Storm Events Classification

Predicted the amount of property damage based on storm-related features, including the event's location, event type and season, among others. Separated low (zero) property damage and high (non-zero) property damage by implementing a random forest classifier. Categorized different storm events into eight event types. Showed trends in the property damage for the past 20 years using exploratory analysis. Performed calculations using NOAA storm events database. Implemented a simple visualization application that predicts the median damage based on the years 1996-2012.

MTA Turnstile Data Analysis

Analyzed MTA turnstile data and identified traffic patterns of the most frequent subway stations. As expected, "commuter hub" stations, such as 34St - Penn Station or 42St - Grand Central, showed most turnstile activity. The goal here was to identify other frequent commuter stations. My analysis compared subway traffic of two "regular" stations, 72ndSt - 123 and 86thSt - 456, to the subway traffic of the "commuter hub" stations, and found similar patterns, both in time and in traffic volume.

An Exploratory Study: Yelp Reviews vs. Check-ins

Studied relation between the number of reviews and the number of Check-ins for Yelp businesses. Built a linear regression model, with coupling between the number of Check-ins and the number of reviews. Explained the increasing variation in the number of Check-ins by a model coupling with features, including WiFi, wheelchair-access, business weekend opening hours, or star rating. The model's prediction algorithm explained about 75% of the data.

Contact

✉ lucia.palova@gmail.com

🌐 <http://lpalova.github.io/>

☎ 848-229-1881

in [linkedin.com/in/luciapalova](https://www.linkedin.com/in/luciapalova)

🔗 <https://github.com/lpalova>

Skills

Data Analysis, Statistics, Python, R, SQL, GitHub, AWS, D3.js, Flask, HTML, Java, Programming, Algorithms, Computational Materials Science, Quantum Espresso, Research, Scientific Writing, LaTeX, Linux, Physics, Mathematics, College Teaching

Education

Johns Hopkins University on Coursera Certification Data Science 2015

R Programming, Getting and Cleaning Data, Exploratory Data Analysis, Statistical Inference, Regression Models, Practical Machine Learning, Developing Data Products

Rutgers, The State University of New Jersey-New Brunswick

PhD Physics 2010

Thesis: Three Theoretical Studies of Ferroelectric Material in Different Geometries

Comenius University, Slovakia

BSc Physics 2004

Comenius University President Award for Excellent Diploma Thesis, Award for Excellent Study Performance and Academic Achievement

Publications

Atomistic Interrogation of B-N Co-dopant Structures and Their Electronic Effects in Graphene, ACS Nano (2016)

Segregation of sublattice domains in nitrogen-doped graphene, JACS 136 (4), 1391 (2014).

Local atomic and electronic structure of boron chemical doping in monolayer graphene, Nano Letters 13, 4659 (2013).

Connecting dopant bond type with electronic structure in N-doped graphene, Nano Letters, 12 (8), 4025 (2012).

Visualizing individual nitrogen dopants in monolayer graphene, Science 333, 999 (2011).

Universal behavior and electric field-induced structural transition in rare-earth substituted BiFeO₃. Advanced Functional Materials 20, 1108 (2010).

Quantum critical paraelectrics and the Casimir effect in time, PRB 79, 075101 (2009).