

LUCIA PALOVA

DATA SCIENTIST, RESEARCHER

Summary

Data Scientist and Researcher with strong knowledge of advanced math, statistics, and data manipulation methods.

Experience

Urban Green Council

Research Fellow: Fall 2016

New York, NY
Oct 2016 to Current

Leading work to develop a Carbon Compass tool. The Compass will model carbon emissions in New York City and offer pathways to 80% reduction by 2050. Assisting with analysis of benchmarking and audit data to find a possible correlation between building operator training and energy usage. Cleansing and analysis of data associated with Local laws 84 and 87, to be included in the next NYC Energy and Water Usage Report.

Metis

Data Scientist

New York, NY
Apr 2016 to Jun 2016

Completed various data science projects, with focus on machine learning and predictive modeling. Used cloud computing, data acquisition, cleansing and visualization techniques.

New Jersey Institute of Technology

Adjunct Faculty, Department of Mathematical Sciences

Newark, NJ
Jul 2012 to Dec 2015

Taught multiple undergraduate classes, including Calculus, Mathematics of Finance (Theory of Interest), Statistics and Probability. Managed 40-60 students per class. Created syllabi, exams, homework, quizzes, assigned grades.

Columbia University

Postdoctoral Research Scientist, Energy Frontier Research Center

New York, NY
Aug 2010 to Feb 2013

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

Rutgers University

Teaching and Research Assistant, Department of Physics and Astronomy

New Brunswick, NJ
Sep 2004 to Jun 2010

Conducted research in 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.

Education

Johns Hopkins University on Coursera

Certification Data Science 2015

Courses: 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

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

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

Machine Learning

Python

R

SQL

GitHub

AWS

D3.js

Flask

HTML

Java

Programming

Algorithms

Linux

Teaching, Research

Physics

Mathematics

LaTeX

Scientific Writing

PowerPoint

Presentation Skills

Quantum Espresso

Computational Physics Research

College Teaching

LUCIA PALOVA

DATA SCIENTIST, RESEARCHER

Selected Publications

- Atomistic Interrogation of B-N Co-dopant Structures and Their Electronic Effects in Graphene, ACS Nano** Jun 2016
T. Schiros, D. Nordlund, L. Palova, L. Zhao, M. Levendorf, C. Jaye, D. R. Reichman, J. Park, M. S. Hybertsen, and A. N. Pasupathy
- Segregation of sublattice domains in nitrogen-doped graphene, JACS** Jan 2014
A. Zabet-Khosousi, L. Zhao, L. Palova, M. S. Hybertsen, D. R. Reichman, A. N. Pasupathy, and G. W. Flynn
- Local atomic and electronic structure of boron chemical doping in monolayer graphene, Nano Letters** Sep 2013
L. Zhao, M. Levendorf, S. Goncher, T. Schiros, L. Palova, A. Zabet-Khosousi, K. T. Rim, C. Gutierrez, D. Nordlund, C. Jaye, M. S. Hybertsen, D. R. Reichman, G. W. Flynn, J. Park, and A. N. Pasupathy
- Connecting dopant bond type with electronic structure in N-doped graphene, Nano Letters** Jun 2012
T. Schiros, D. Nordlund, L. Palova, D. Prezzi, L. Zhao, K. S. Kim, U. Wurstbauer, C. Gutierrez, D. Delongchamp, C. Jaye, D. Fischer, H. Ogasawara, L. G. M. Petersson, D. R. Reichman, P. Kim, M. S. Hybertsen, and A. N. Pasupathy
- Visualizing individual nitrogen dopants in monolayer graphene, Science** Aug 2011
L. Zhao, R. He, K. T. Rim, T. Schiros, K. S. Kim, H. Zhou, C. Gutierrez, S. P. Chockalingam, C. J. Arguello, L. Palova, D. Nordlund, M. S. Hybertsen, D. R. Reichman, T. F. Heinz, P. Kim, A. Pinczuk, G. W. Flynn, and A. N. Pasupathy
- Universal behavior and electric field-induced structural transition in rare-earth substituted BiFeO₃, Advanced Functional Materials** Mar 2010
D. Kan, L. Palova, V. Anbusathaiah, C. J. Cheng, S. Fujino, V. Nagarajan, K. M. Rabe, and I. Takeuchi
- Magnetostuctural Effect in the Multiferroic BiFeO₃-BiMnO₃ Checkerboard from First Principles, Physical Review Letters** Jan 2010
L. Palova, P. Chandra, and K. M. Rabe
- Quantum critical paraelectrics and the Casimir effect in time, Physical Review B** Feb 2009
L. Palova, P. Chandra, and P. Coleman

Sample Data Projects

Publication 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 a network algorithm to rank publications. Built a demo search engine application that allows the user to search for publications and explore the publication network.

Storm Event Classification

Predicted the amount of property damage based on storm-related features, including the event's location, event type and season, among others. Implemented a random forest classifier to separate low (zero) property damage and high (non-zero) property damage events. Categorized 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 frequented subway stations. As expected, "commuter hub" stations, such as 34St - Penn Station or 42St - Grand Central, showed most turnstile activity. I have compared these to two "regular" stations, 72ndSt -123 and 86thSt - 456, 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.