Michael Wieck-Sosa

github.com/mwiecksosa • linkedin.com/in/mwiecksosa • mwiecksosa@gmail.com • 503-757-4726 • Urbana, IL

EDUCATION

University of Illinois at Urbana-Champaign | MS Statistics *Expected May 2022*

- GPA: 3.91/4.00 | Awards: 2-Year Full Tuition Waiver & Stipend through Guaranteed Graduate Teaching Assistantships
- Fall 2021 Courses: Graduate Measure-Theoretic Theory of Probability II (Stochastic Differential Equations, Brownian Motion, Stochastic Calculus - Le Gall, Karatzas & Shreve), Graduate Real Analysis (Measure Theory - Folland), Graduate Statistical Learning (Friedman, Tibshirani, & Hastie), Graduate Advanced Regression Analysis (Generalized Linear Models)

Fordham University – New York, NY | BS Mathematics with Minors in Computer Science & Economics *Graduated May 2020*

- GPA: 3.77/4.00 | Honors & Awards: *magna cum laude*, Loyola Scholarship, Dean's List, Pi Mu Epsilon Math Honor Society

RELEVANT EXPERIENCE

MIT Lincoln Laboratory, Interceptor & Sensor Technology Group | Summer Research Program (12 weeks) *May 2021-Present*

- Signal Processing: Optimizing parameters for statistical signal processing methods used in a space surveillance system
- Recalibrating System: Algorithmically calculating angle random walk & drift in system and recalibrating to minimize errors

University of Illinois at Urbana-Champaign, Department of Computer Science | Graduate Research Assistant *Jan. 2021-Present*

- Data Mining: Analyzing the stochastic process & network structure of info diffusion across Twitter, Facebook, & Reddit
- Big Data Processing: Querying 1+ exabytes of data with APIs & using Python scripts on server to analyze multi-GB of data

National Center for Supercomputing Applications, Software & Data Analysis Group | Programmer (Part-time) *Sept. 2020-Present*

- Time-Series Analysis: Building weighted regressions & plots to analyze time-series of nitrogen discharge from river networks

University of Illinois at Urbana-Champaign, Department of English | Graduate Research Assistant *Aug. 2020-Present*

- Data Engineering: Wrote Python scripts to query biodiversity databases & set up PostgreSQL database with data on marine species habitats for website to visualize & quantify changes in marine species distributions & ocean conditions since the 1800s

Corteva Agriscience, Data Science Division | Summer Research Intern (10 weeks) *June 2020-Aug. 2020*

- Time-Series Clustering: Reduced dimension & clustered time-series to study how environment types in the US have changed

University of Michigan, Department of Biostatistics | Summer Research Program (8 weeks) | Poster *June 2019-July 2019*

- Medical Signal Processing: Built CNN model in PyTorch to classify irregularly sampled time-series & attended ML lectures

Fordham University, Department of Computer Science | Volunteer Research Assistant | Poster *March 2019-March 2020*

- Computer Vision: Advanced tracker written in C++ with object detection & accelerated with GPU computing cluster
- Audio Signal Processing: Built & optimized deep learning models in PyTorch to classify noisy audio as coughs or snores

Carnegie Mellon University, Department of Statistics | Summer Research Program (10 weeks) | Poster *June 2018-July 2018*

- Interpretable Models: Built GLMs & plots in R to analyze patterns in human trafficking & attended Statistics lectures

Fordham University, Department of Economics | Volunteer Research Assistant *May 2017-March 2020*

- Time-Series Forecasting: Built parallelized models to predict recessions in the US & interpreted models with LIME & SHAP
- Economics + Machine Learning: Used surveys on schools in India to cluster schools into performance levels in each income level & used convolutional neural networks to analyze photos of those schools to learn features related to a school's success
- Econometric Models: Wrote review of the statistical analyses used to study how immigration impacts wages in the USA

PROJECTS

- 2020 Election Prediction based on Sentiment Analysis of Twitter (Group Project for CS 598 at UIUC) – Website, Code
- Introduction to Extreme Value Theory (Course Project for MATH 561 at UIUC) – Report
- Introduction to Fractals with Applications in Probability Theory (Course Project for MATH 444 at UIUC) – Report
- Bayesian Vector Autoregressions with Example in R (Group Project for STAT 429 at UIUC) – Report, Code
- Reproduction of *Scalable Generalized Linear Models* by Erdogdu et al. (Project for MATH 4006 at Fordham) – Report, Code
- CS Course Projects: Data processing in C++ with object-oriented programming, Turing Machine, bowling in MIPS assembly

SKILLS

- Programming Languages: Extensive experience in C++, Python, R, MATLAB, and proficient in C, C#, VBA, Java, Stan
- Libraries: NumPy, Pandas, PyTorch, TensorFlow, Scikit-Learn, NLTK, Dash, Shiny, Folium, tidyverse, ROS, OpenCV
- Frameworks: MySQL, PostgreSQL, MongoDB, Hadoop, Spark, MapReduce, Flask, Slurm, EC2, Linux, Docker, Git, Bash

RELEVANT COURSEWORK

- **Math:** Graduate Theory of Probability I (Measure Theory, LLN, CLT, Martingales - Durrett 1-4), Elementary Real Analysis, Topology, Abstract Algebra, Differential Geometry, Numerical Analysis, Probability, Statistics, ODEs, Linear Algebra
- **Statistics:** Graduate Geometric Flows, Graduate Mathematical Statistics I, Time-Series Analysis, Econometrics, Biostatistics
- **Computer Science:** Theory of Computation, Artificial Intelligence, Computer Algorithms, Operating Systems, Computer Organization, Databases, Programming for Math in Python, Data Structures in C++, Intro Computer Science I-II in C++

TEACHING EXPERIENCE

University of Illinois at Urbana-Champaign | TA for STAT 385 Statistical Programming Methods in R *Fall 2020, Spring 2021*

- Duties: office hours, videos, grading. Topics: indexing, loops, tidyr, stringr, ggplot2, dplyr, parallel programming, Shiny, Git

Fordham University, Mathematics Club at Fordham College at Lincoln Center | Volunteer Teacher *Jan. 2019-April 2019*

- Taught 6-week introductory course to undergraduates about how to use Python and R for basic modeling & analysis

LEADERSHIP

- Founder & President of Math Club at Fordham: Hosted speakers, taught intro course about using Python & R for data science, started the American Statistical Association Student Chapter, and hosted the ASA DataFest 2019 competition
- Lead Strategist (paid position) at the Fordham Social Innovation Lab: Led 12 students in weekly meetings to study issues facing recent refugees at Fordham University. Result: reported to Dean of Students and 1 student received full funding

EXTRACURRICULARS

- **Competitions:** Winner of JPMorgan Data for Good Hackathon 2021
- **Languages:** Native English, Proficient in Spanish
- **Volunteering:** Assisted at the food pantry of the Church of St. Paul the Apostle in Manhattan, NY where I packed bags of food, digitalized database, & translated documents & announcements into Spanish (3 hours/month, Sept. 2018 – March 2020)