

# Eduardo Blancas Reyes

[e.blancas@columbia.edu](mailto:e.blancas@columbia.edu)   <https://blancas.io>   <https://github.com/edublancas>

## Education

---

### Columbia University

Aug 2017 - Dec 2018 (Expected)

M.S. in Data Science. Fulbright Scholar.

Coursework: Computer Systems, Machine Learning, Probabilistic Programming, Robustness and Security in ML, Applied Deep Learning.

### Universidad Nacional Autónoma de México (UNAM)

Aug 2016 - Jun 2017

Graduate Certificate in Applied Statistics (one-year program of graduate coursework)

Coursework: Statistical Inference, Sampling, Multivariate Analysis, Regression Models and Bayesian Inference.

### Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM)

Jan 2011 - May 2015

B.S. in Mechatronics Engineering. Graduated with honors.

## Experience

---

### Columbia University (Department of Statistics)

New York, NY

Research Assistant (Advisor: Liam Paninski)

Sep 2017 - Dec 2018

- Maintainer of **YASS**: an open source library for Computational Neuroscience
- Develop high-performance algorithms for processing Larger-than-Memory datasets (NumPy, SciPy, scikit-learn)
- Improve existing code: optimization using benchmarking and profiling tools, parallelization using the multiprocessing library and NumPy parallel operations
- Use Deep Learning for neural activity detection and classification (TensorFlow, Keras), GPU training using Google Compute Engine
- Worked with the research team to turn YASS old codebase into a tested, organized and documented Python library

### Krieger

Mexico City, MX

Lead Backend Engineer

Jun 2016 - Aug 2017

- Oversaw backend development across projects to ensure code quality, performance and reliability
- Developed internal software tools that are used in several ongoing projects (input validation, software versioning, deployment, user permissions). These tools have reduced development time and increased team's productivity
- Developed web services using Python (Flask) and deployed them using Google App Engine, Cloud Datastore and Cloud Storage
- Engineered a data pipeline to perform entity resolution using Apache Airflow

### University of Chicago (Center for Data Science and Public Policy)

Chicago, IL

Data Science for Social Good Fellow (Advisor: Rayid Ghani)

May 2015 - Jun 2016

- Developed a Machine Learning model for the City of Cincinnati to identify properties at risk of building code violations. 50GB+ of data
- Designed a predictive model for Infonavit (largest mortgage provider in Mexico) to predict home abandonment. 100GB+ of data
- Both pipelines (ETL, feature extracion, model training and model selection) developed using Amazon EC2, S3 and RDS (PostgreSQL and PostGIS).
- Communicated data analysis results to non-technical audience

## Publications

---

**E. Blancas Reyes**, J. Helsby, K. Rasch, P. van der Boor, L. Haynes, R. Ghani, E. Cunningham. Early detection of properties at risk of blight using spatiotemporal data. Data for Policy 2016 at University of Cambridge.

K. Ackermann, **E. Blancas Reyes**, S. He, T. Anderson Keller, P. van der Boor. R. Khan, R. Ghani. Designing Policy Recommendations to Reduce Home Abandonment in Mexico. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'16.

## Skills

---

**Programming languages:** Python, R, Objective-C, SQL. **Tools and technologies:** scikit-learn, NumPy, flask, Google Cloud (App Engine, Datastore, Storage and Compute Engine), PostgreSQL, PostGIS, Docker, MongoDB, AWS (EC2, S3, RDS and Beanstalk)

CV with prior experience, awards, press and other projects available at <https://blancas.io/files/cv.pdf>.