

# Esteban Navarro Garaiz

Los Angeles, California • ☎ (917) 412 - 9631 • ✉ eng272@nyu.edu • [in estebanng](https://www.linkedin.com/in/estebanng)

## Experience

### Los Angeles Dodgers – Quantitative Analyst

2020 – Present. Los Angeles, CA

- Collaborated closely with Player Development to improve analytic tools used in MiLB development
- Lead QA's collaboration with Pro Scouting to prioritize scouting time and improve player acquisition efforts
- Created probability models for relevant events to describe catcher defense (additional strikes, blocking, stolen base attempts, successful steal given attempt). Used these to create descriptive metrics (framing, blocking and arm) and built a projection system on top to forecast future performance given observed data

### MongoDB – Data Science Intern

2019. New York, NY

- Developed two quantitative segmentations to understand our Cloud customers, one at the product level and one at the company level. Created features from product usage and organizational characteristics
- Key insights are being used by Growth department to guide new offerings throughout customer funnel and Marketing department to personalize customer interaction, and boost customer retention

## Education

### New York University – Center for Data Science

New York, NY

Master of Science in Data Science – GPA 3.88

2018 - 2020

- Recipient of a CDS DeepMind Fellowship and a Fulbright-García Robles Scholarship.
- Course highlights: Machine Learning, Big Data, Deep Learning, Natural Language Processing, Data Engineering for Machine Learning, Optimization and Computational Linear Algebra, Forecasting Time Series Data

### Universidad Nacional Autónoma de México (UNAM)

Mexico City

Bachelor of Science in Mathematics - 9.03 GPA

2012 - 2018

- Course highlights: Probability (2), Statistics (4), Stochastic Processes (2), Linear Algebra (2), Operations Research (2), Graph/Game Theory (3), Calculus (4)
- Worked as Assistant Professor for Probability I, Stochastic Processes 1 and Select Topics in Operations Research
- For my graduation thesis, I created a soccer prediction model, based on the Poisson Distribution, and applied it successfully to the betting market using the Kelly Criterion <https://tinyurl.com/wtpssjt>

## Projects

- **NBA Shot prediction (Brooklyn Nets Capstone Project):** Using state-of-the-art, high-dimensional tracking data and a two-tier ensemble model, we are able to predict shot outcomes – whether a shot was made or missed, with the possibility of a defensive foul. Our model is used for analytical player evaluation and proves to be more stable than outcomes from season to season. <https://bit.ly/2wZjZtn>
- **Quora Insincere questions:** Built supervised learning models to predict Quora insincere questions (false, making a statement, pushing an agenda), using various NLP (Count Vectorization, TF-IDF, word embeddings) and modeling (Naïve Bayes, Logistic Regression, topic LDA) techniques. Correctly labeled 97% of sincere questions, while flagging 70% of insincere questions <https://tinyurl.com/ucocv4d>

## Skills

- **Python:** Data Science Packages (Pandas, scikit-learn, Matplotlib, NumPy, ScyPy, Pytorch, Tensorflow)
- **Machine Learning:** Neural Networks, NLP, Random Forests, Regression, kNN, Naïve Bayes, Boosting, PCA/SVD
- **Other Programming:** R – Tidyverse (stats, dplyr, ggplot2, tidyr), Unix Shell, Hadoop, Spark
- **SQL:** Industry level. Have worked with sqlite3, PostgreSQL / psycopg2, MongoDB query language (M001)
- **Other:** LaTeX, Spanish (native speaker, have worked as translator), GitHub in production environment