

# Matthew Maresca

## Data Scientist

## Summary

As a self-proclaimed professional tinkerer, I like taking on challenges that provide an opportunity to solve interesting problems while expanding my skill set. I began learning to program in order to build applications to better serve clients in the fitness space. In the years since, while juggling various business ventures, my insatiable curiosity for learning new things led me to develop a method for understanding complex technical concepts at a deeper level. This has enabled me to make use of my diverse background while transitioning into a data science position. I look forward to merging my quantitative and creative skills in my next role.

## Experience

### Metis Data Science Bootcamp New York, NY

- Metis is a full-time 12-week immersive data science bootcamp with a focus on python programming, machine learning, and statistical analysis.
- Built several real-world projects end-to-end from data acquisition to statistical modeling.
- See Projects section below for examples.

### Fit Ignition LLC NY/NJ Personal Fitness Trainer - Corrective Exercise 2011 to Current

- Began my own private training practice with a focus on correcting postural distortions and faulty movement patterns.
- Built a ReactJS/Python web app to measure client progress through personalized challenges.
- Built my own website and blog from scratch to promote the business.

### Digital Business Consultant 2011 to 2014

- Engaged in various consulting projects with multiple dietitians and other wellness professionals.
- Redesigned and developed client websites.
- Worked to increase web traffic via content marketing and performed A/B split tests to optimize websites for email list building and product sales.
- Assisted with the online promotion for a major book launch, helping to build the author's web presence from scratch. Also partnered in the creation of an online coaching program to complement the book.

### New York Sports Clubs Newark, NJ Personal Fitness Trainer 2006 to 2010

- Created fitness programs for individual clients, tracking progress with basic charts and spreadsheets.
- Sold personal training packages to gym members.
- Assisted in gym membership sales and customer service.

## Projects

### Image Classification with Bag of Visual Words

- Classified a subset of 10 insect-related categories from the ImageNet dataset without neural nets.
- Achieved 43% accuracy using a bag of visual words pipeline utilizing KMeans clustering and a Support Vector Classifier.
- Used OpenCV for image processing and creation of feature descriptor vectors.

### Data Science Resource Recommendation Tool (NLP)

- Used natural language processing to recommend resources based on full-text inputs related to data science.
- Scraped a large corpus of several thousand blog posts.
- Wrote a shell script to perform a daily check to add new posts.
- Fit a TF-IDF matrix to determine documents most closely related to the input.
- Performed topic modeling to partition the corpus into distinct clusters and learn popular topics.

### Baseball Batting Average Predictor

- Used linear regression to predict a baseball player's batting average using stats that are not heavily dependent on randomness.
- Developed proof of concept for deeper analysis, as the model was able to make predictions better than a baseline of the previous season's batting average.
- The model did not even use previous batting average at all, as it falls in my "randomness-dependent" category.
- Scraped 20+ years of statistics for all MLB players (though only used stats from post-2001 as that is when certain advanced metrics arose).

### Satellite Image Segmentation

- Performed image segmentation to map farmland in high-resolution satellite imagery.
- Used Open Street Maps data and QGIS to perform exploratory geospatial analysis and create ground truth masks for the images.
- Trained several Convolutional Neural Network architectures, building off of designs from state-of-the-art papers
- Performed a bit of "net surgery" to build a model that fused multiple architectures in order to solve a problem of not being able to utilize pre-trained weights in conjunction with all eight channels of spectral data.

## Contact

✉ [mattmaresca@gmail.com](mailto:mattmaresca@gmail.com)

🌐 [mattymar.github.io](https://mattymar.github.io)

☎ 551-265-1953

📍 North Arlington, NJ 07031

📄 [in matt-maresca-347b6b22/](https://github.com/matt-maresca-347b6b22/)

👤 Mattymar

## Education

Drew University  
BA Economics

## Skills

### PROGRAMMING LANGUAGES (AND SQL)

Python

Javascript (Reactjs, Node.js)

SQL

### MACHINE LEARNING

Classification (SVM, Logistic Regression, Random Forest, GBTs)

Convolutional Neural Networks

Unsupervised (KMeans, PCA)

Natural Language Processing

Linear Regression

Recurrent Neural Networks

### COMPUTER VISION

Image Classification

Semantic Image Segmentation

OpenCV

Pose Estimation

### LIBRARIES

Scikit-Learn

Tensorflow

Keras

Numpy

Pandas

OpenCV

Matplotlib

Scrapy

### MISCELLANEOUS

GIS - Geospatial Analysis

Web Scraping

Probabilistic Graphical Models

AWS & GCP

Bash Scripting