

# MATTHEW GOLBEN

## DATA SCIENTIST

**E-mail:** matthewgolben@gmail.com  
**Phone:** 8455459258  
**LinkedIn:** matthew.golben/  
**Github:** NeOresme222/

Data Scientist with experience in R&D and US Navy veteran.

## Skills

### BIG DATA

Amazon Web Services  
Docker  
Spark Resilient Distributed Dataset  
Structured Query Language  
Web Scraping and Application Programming Interface

### MACHINE LEARNING

Artificial Neural Networks  
Boosting Methods  
Clustering  
Linear & Logistic Regression  
Markov Chain Monte Carlo  
Natural Language Processing  
Non-negative Matrix Factorization  
Principal Component Analysis  
Random Forest Classification  
Recommender Systems  
Singular Value Decomposition

### PROGRAMMING

Pandas  
Python  
Matplotlib  
Numpy  
Object Oriented Programming

### STATISTICS

Bayesian & Frequentist Statistics  
Hypothesis Testing  
Probability

### MISCELLANEOUS

Aviation  
Electronics  
Excel  
Grant Writing  
Inventor Design Software  
Leadership  
Machining  
Origin Software  
PowerPoint

## Education

Galvanize Inc.	Mar. 2020 to June 2020
University of Massachusetts at Amherst BS Psychology with an emphasis in Neuroscience 2017	Sept. 2015 to Aug. 2017
Orange County Community College AS Science	2013 to 2015

## Employment

Xergy Inc. Junior Engineer	Harrington, Delaware Mar. 2018 to Oct. 2019
- Design and manufacture of parts using calculations of specified performance values. - Management of teams performing previously stated procedures. - research and acquisition of critically special materials. - report, grant writing and rendering of schematics for prototype design. - Extensive work with hydrogen and hydrogen storage. - Extensive work with heat engines and pumps. - Co-author of patent Electrochemical Expander With Hydride Alloy Sub-System (application #: 62/762.510)	
Ergenics Corp. Assistant to Chief Engineer	Ringwood, New Jersey Sept. 2012 to Sept. 2014
- Research and acquisition of hydride alloys. - Manufacture of testing cells, data collection and analysis of said cells. - Lead of manufacturing demonstration units critical to the acquisition of a contract with ARPA-E (Advanced Research Projects Agency - Energy) for a 1-kilowatt heat engine prototype. - Integral in the research for, and construction of, said prototype.	
United States Navy Aviation Electronics Technician, Petty Officer 3rd Class Secret Security Clearance.	Virginia Beach, Virginia Sept. 2007 to Jan. 2012
Dates of Deployment: April 2011 – November 2011: Aviation Electronics, Organizational Level – 13 hour shifts, 7 days per week, 4 to 6 weeks at a time. Aircraft Carrier CVN 77; George H. W. Bush – Persian Gulf. Successful deployment in support of Operation Enduring Freedom and the Maiden Voyage of U.S.S. George H. W. Bush. Operations consisted of troubleshooting and repairing the electronic and computational systems of F-18 aircraft. In charge of maintaining Material Safety Data Sheets (MSDS) and Hazardous Waste (HW) systems.	
April 2010 -- January 2011: Military Police – 13 hour shifts, 6 days per week, up to 6 months at a time. Detention Facility in Parwan, Bagram Air Base, Afghanistan. Proper conduction of detention procedures in observance of human dignity; successful training of Afghan National Army for self-sufficient operation; supervision of up to 25 people; received correctional officer's journeyman certificate.	
September 2008 – March 2009: General Aircraft Maintenance – 13 hour shifts, 7 days per week, 5 to 7 weeks at a time. Aircraft Carrier CVN 71; Theodore Roosevelt -- Persian Gulf. Performed general aircraft maintenance, inspection, troubleshooting and flight preparedness operations in support of Operation Enduring Freedom. A squadron sortie rate of 98% culminated in a Battle Effectiveness Award.	

## Projects

Patent Query	June 2020 to June 2020
The world is enriched when people are empowered to contribute their ingenious creations. The endeavor of this project was to encourage ease of patent filing as well as investigate prior art. Subsequent code juxtaposes a patent query (description) with the most recent download cache of patents from the United States Patent and Trademark Office to return the most likely section classification as well as the top five most similar patents. A Gradient Boosted Classifying model performed with 71% accuracy in classification and Latent Semantic Indexing was used to identify inter-document similarity.	
Personality Categorization	May 2020 to May 2020
Understanding the minds of an online culture or forum is valuable in determining motivation and methodology. This project sought to classify users according to the Myers-Briggs (MB) personality classification system. Natural Language Processing methods were performed on the last 50 posts of 8,675 users on PersonalityCafe.com as reference for the self-disclosed personality of each user. Multi-classification models were applied of which Random Forest excelled with 26.9% accuracy. Classification for each binary personality subclass of the MB system was also performed with a variety of models. A Binomial Naive Bayes model outperformed the others in all cases, with a tope accuracy of 88.1%.	
World Development Indicators: Power and Income	Apr. 2020 to Apr. 2020
Electrical power generation is a critical feature of modern society. Understanding how and to what degree energy is produced in a country may provide economic and development insights for countries and corporations. Data sourced through the World Bank was initially organized according to each country's Gross National Income (GNI) bracket. Timelines, boxplots, and correlational heat-maps were generated in order to observe the dynamics of the relationship between a country's wealth and energy production. While there is a definite positive correlation of wealth to energy, there is an inverse correlation of wealth to energy efficiency.	