ERIC ZACHARIA

DATA SCIENTIST

CONTACT

(774) 239-5342
ekzachar@uchicago.edu
Resume Website
GitHub
LinkedIn

EDUCATION

M.S. Computer Science
Specialized in Data Analytics
University of Chicago '21

B.S. Aerospace Engineering
Syracuse University '16

SKILLS

(Deep) Machine Learning, Advanced Algorithms, Big Data Mining, Databases, Python, NumPy, Pandas, nltk, PyTorch, Keras, Scikit-learn, hugging-face, XGBoost, Flask, PyMC3, TweePy, Matplotlib, Jupyter, Google Colab, Amazon EMR & SageMaker, SQL, Spark, Golang, QuantConnect, Alpaca API, Git, HTML, CSS, $L^{A}T_{E}X$, Excel

LICENSES

Secret Security Clearance '20
Private Pilot License '19
Scuba Diving License '16

Other CS Projects

Authorship Text Verification
Blob Video Game
Bloom Filter Streaming Algorithm
Caesar Cypher File De/Encryptor
Concurrent Twitter Feed Clone
Content Management System
Diet Planner iOS App
Flajolet-Martin Streaming Algorithm
Jack Prog. Language Compiler
k-Nearest Neighbors Classifier
Locality Sensitive Hashing
Linear Regression Model
Map Reduce Matrix Multiplication
Roman Numeral Converter
Slack & Pinterest Clones
Vanilla Soft Margin Primal/Dual SVM
Speaker Recognition System
Stock API Graph Generator
Vanilla AdaBoost Decision Tree
Vanilla Decision Tree Classifier
Vanilla Graph Algorithms
Vanilla Ordinal Logistic Regression
Vanilla Searching Algorithms
Vanilla Sorting Algorithms
Text Completion Software
Wedding Website Database

WORK EXPERIENCE

NLP Data Science Intern, University of Chicago Medical Center — Summer 2021

Preface: A significant amount of medical knowledge exists in unstructured medical notes, and many doctors waste valuable time carefully picking the proper ICD codes for their patients. Correct ICD codes are necessary for patients to receive proper follow-up procedures and are required by insurance companies for proper payment to the hospital.

- Developed language models to correct missing or wrong ICD codes with 96% accuracy
- Predicted diseases in cardiology patients using historical medical notes with 64% accuracy

Machine Learning Research Intern, Argonne National Laboratory — Summer 2021

- Worked with molecular engineers in researching machine learning techniques to compensate for the issue of drifting readings for water contamination sensors
- Applied Bayesian inference to predict the curve of voltage drift on experimental data
- Reduced run-time of the team's compensation software from 45 hours to 27 minutes
- Educated the team about the applications of ML in molecular research
- Created non-CS-friendly pipelines for the researchers to use in their projects

Level II Aerospace Engineer, Spirit AeroSystems Inc. — 2016-2020

- Designed, built, analyzed, and tested Boeing's 787 Dreamliner, and Boeing's 777X, and an aerospace structure for The Department of Defense

Host of Glacier's Bed and Breakfast — 2018-2020

- Hosted guests in my 3-bedroom home with over 80 bookings and a 5-star rating

Fluid Dynamics Researcher, Syracuse University — 2015-2016

- Analyzed the propulsive aspects of dolphin tails and experimented with 3D-printed dolphin caudal fins that mimicked swimming motion inside a water tunnel
- Constructed 3D visualizations of vortex flows to demonstrate swimming efficiencies

PROJECT EXPERIENCE

Wedding Website — present

- Designed and currently maintaining a website for my upcoming wedding in April

Predicting the Genre of Music Samples using a Convolutional Neural Network — Fall 2021

- Developed a Convolutional Neural Network to learn the patterns from spectrograms
- CNN 83% accuracy outperformed benchmark Support Vector Machines 63% accuracy

Reinforcement Learning Robot that Navigates Variable Environments — Fall 2021

- Wrote DP and temporal difference algorithms to navigate a stochastic environment
- Optimized Double Q-Learning parameters through cross-validation for a 77% success rate

Quantitative Momentum Trading Algorithm — Fall 2021

- Applied knowledge of fundamental valuation metrics and momentum-based trading ideas to create a paper trading algorithm that beat the S&P 500 by 10% over a 3-month period.

Predicting the Stock Market with Sentiment Analysis of Live Tweets — Summer 2021

- Developed an NLP pipeline that algorithmically trades stocks using opinions from Twitter
- Tripled the S&P 500 after one month of bullish market trading

Vanilla Convolutional Neural Network Image Classifier using CIFAR10 Dataset — Spring 2021

- Wrote a CNN image classifier from scratch, e.g., using only Python and NumPy

Vanilla Neural Network Digit Classifier using MNIST Dataset — Spring 2021

- Wrote a NN handwritten digit classifier from scratch, e.g., using only Python and NumPy

Last updated : 12/14/2021