ERIC ZACHARIA

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

CURRENT POSITION

Data Analytics Master's Student Sept, 2020 – Dec, 2021
Department of Computer Science, University of Chicago

EDUCATION

M.S. in Computer Science, University of Chicago Dec, 2021
B.S. in Aerospace Engineering, Syracuse University 2016

SKILLS

Python, NumPy, Pandas, PyTorch, Sklearn, XGBoost, PyMC3, Jupyter, Amazon EMR, Colab, SQL, Spark, Scala, Golang, HTML5, CSS3, MATLAB, L^AT_EX, Excel, Git

WORK EXPERIENCE

Data Science Intern, University of Chicago Medical Center 2021

- Developed (NLP) language model pipelines to parse through medical patient notes for the purpose of shifting a doctor's time from paperwork back to their patients
- Many doctors waste valuable time carefully picking the proper ICD codes for their patients
- Corrected notes missing ICD codes with 96% accuracy - necessary for patients to receive proper follow up procedures and required for insurance companies to pay the hospital
- Experimental project predicted diseases in cardiology patients with 64% accuracy using historical cardiologist notes with labeled disease outcomes

Machine Learning Research Intern, Junhong Chen's Research Group 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 and random sampling over distributions to predict the curve of voltage drift with respect to time based on experimental data
- Successfully compensated sensors at various water temperatures and pH levels
- Reduced the run-time of the research team's initial program from 45 hours to 27 minutes using Google Colab GPU's and PyTorch
- Educated the team about ML during weekly meetings and created user-friendly software for the team to apply to their research projects

Aerospace Stress Engineer, Spirit AeroSystems Inc. 2016-2020

- Worked on teams that designed, built, analyzed, and tested aerospace structures for The Department of Defense, Boeing's 787 Dreamliner, and Boeing's 777X; gained promotion to Level II Stress Engineer within 18 months
- Applied knowledge of MATLAB, Excel, linear algebra, ABD matrices, statics, dynamics, fatigue, damage tolerance, crack growth, Finite Element Analysis, Computer-Aided-Design, and properties of aluminum and composite materials

Host of Glacier's Bed and Breakfast 2018-2020

- Hosted guests in my 3-bedroom home with over 80 bookings
- Classified as a "Super Host" by Airbnb for sustaining a 5-star rating over two years

Fluid Dynamics Researcher, Syracuse University 2015-2016

- Studied the propulsive aspects of dolphin tails and experimented with 3D-printed dolphin caudal fins that mimicked swimming motion inside a water tunnel
- Synthesized knowledge of MATLAB, Excel, SolidWorks (CAD), 3D printing, Arduino, laser-induced fluorescence, HD videography, fluid dynamics, and circuits
- Used fluorescent dye and planar-shaped lasers to highlight trailing-edge vortices
- Constructed 3D vortex visualizations in MATLAB using pixel data from the HD recordings of the 2D vortex sheets by stacking and interpolating between points

PROJECT EXPERIENCE

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

- Developed a Natural Language Processing pipeline that trades stocks solely based on the trends of opinions about stocks on Twitter
- Trained a BERT classification model with tweets labeled bearish or bullish
- Communicated with Twitter API to classify sentiment of live tweets
- Devised an algorithm to determine the number of shares to trade based on the trend of a stock's sentiment, classification certainty, price per share, portfolio value, and buying power
- Conducted day-trades every few seconds using the Alpaca Trading API
- Tripled the S&P500 and DJIA after three days of bullish market trading

Predicting the Genre of Music from Audio Samples — 2021

- Engineering audio features from .mp3 files for supervised learning prediction
- Applied various classification models with a focus on Support Vector Machines
- Improved on this prediction by skipping the need to engineer features from the audio and developing a Convolutional Neural Network to learn the patterns from a music sample's spectrogram.
- Classified music genre's directly from .mp3 files with 60% accuracy.

Foundational Machine Learning Projects — 2021

Implemented the following models from scratch, with little to no toolkits, e.g., sklearn

- AdaBoost Decision Tree Classifier ensemble model
- Pegasos algorithm for solving the soft margin Support Vector Machine primal problem, skipping the dual problem, with regularization & hinge loss
- Soft margin SVM dual problem using a quadratic programming solver
- Neural Network classifier with computation graph, Xavier weight initialization, and forward and backward propagation of matrix multiplication and addition, cross product, sigmoid, ReLU, and SoftMax
- Convolutional Neural Network built upon NN classifier with forward and backward propagations for convolution, max pooling, and flattening filters
- k-Nearest Neighbors with cross validation to optimize k
- Linear Regression using gradient descent and cross-validated learning rates
- Logistic Regression with regularization, normalized inputs, & gradient descent
- Quadratic Weighted Kappa scoring function for ordinal logistic regression
- Decision Tree Classifier with splitting on class and continuous attributes, information entropy, generalization error, and pruning

Other Projects

Comparative Market Analysis using Linear Regression		Diet Planner iOS App
Mood Journaling App	Pinterest Clone	Slack Clone
Content Management System	Speaker Recognition System	Twitter Feed Clone
Jack Programming Compiler	Blob Video Game	VM to Assembly Translator
Text Completion Software	Stock API Graph Generator	Authorship Text Verification
Roman Numeral Converter	Caesar Cypher File Encryptor	1st Place Model Aircraft Flight

---

CERTIFICATIONS

Secret Government Security Clearance	2020
Private Pilot	2019
Open Water Scuba Diver	2016

ARTICLES

Predicting the Stock Market using the BERT Model and Sentiment Analysis of Live Tweets, 2021 (available here)
Hydrodynamics of Dolphin Caudal Fins, 2016 (available here)