jeremy@thaller.dev • 978-496-7990 github.com/jthaller • jeremythaller.com

JEREMY THALLER

U.S. Citizen • New York, NY

CAREER SUMMARY

Dual MS engineer with a strong background in data analysis and statistical modelling Continued history of presenting highly technical data and insights to guide leadership decisions. Specializes in deep learning and natural language processing.

EDUCATION

UNIVERSITY OF MUNICH (LMU), Germany [2019 – Aug. 2021]
MS Materials Science and Engineering
ADAM MICKIEWICZ UNIVERSITY, Poland [2019 – Aug. 2021]
MS Computational and Applied Physics
WILLIAMS COLLEGE, Massachusetts [2015 – 2019]
BA in Physics, Honors | Captain of Track & Field Team | Sigma Xi Scientific Research Honor Society

WORK EXPERIENCE

BROOKHAVEN NATIONAL LABORATORY (Upton, New York)
MS Thesis Researcher – Machine Learning / Nanomaterials *[Feb 2021 – Present]*

- Optimized the industry-standard simulation workflows by developing a new statistical-based methodology utilizing skew-norm distributions; efforts reduced the industry-standard computation time by 50x
- Utilized TensorFlow to predict disorder from the resulting simulation dataset, reducing the amount of required modeling data by 90% while maintaining accuracy parity
- Expanded the above convolutional neural network's predictive domain via few-shot transfer learning
- Created and actively managed the lab's GitHub organization; constructed example projects to demonstrate best development and data management practices
- Presented weekly research insights to non-technical stakeholders; advised and influenced the lab's strategic approaches

YALE UNIVERSITY (New Haven, Connecticut)
Investigative Consultant – Metallic Glasses / Metallurgy *[Summer 2019]*

- Wrote and deployed a GUI Python program to automate and expedite the material-candidate screening pipeline
- Designed rigorous experiment parameters to isolate causal variables and perform root cause analysis (RCA)

WILLIAMS COLLEGE (Williamstown, Massachusetts)
BA Thesis Researcher – Condensed Matter Physics *[May 2018 - May 2019]*

- Automated the complex data analysis process through MATLAB scripts into a click-to-run pipeline
- Determined relative feature importance of experimental parameters via entropy, CART, and random forest algorithms
- Presented project results at multiple international conferences

RECENT DATA SCIENCE PROJECTS

FACEBOOK MESSENGER ANALYSIS

- Scraped 10+ years of messaging data via selenium and BS4; processed/cleaned and analyzed messaging trends with Pandas, NLTK, SpaCy, and Gensim
- Created a Bayesian-based friend classifier capable of profiling users based on messaging styles, content, and sentiment.
- Built a from-scratch generative chatbot trained on personal messaging data using Keras and GloVe embeddings

SPOTIFY SONG RECOMMENDATION

- Utilized PySpark for big data, scalar-aggregate-reduction for 20x faster SQL queries, and REST APIs for data supplementation in order to investigate the evolution of song and user listening trends
- Created a song recommendation algorithm based on song similarity utilizing t-SNE distances and deep embeddings

CRYPTOCURRENCY EDA AND PREDICTIVE MODELING

- Analyzed key financial indicators over time and forecast DOGE Coin values in Python using an LSTM architecture

SKILLS AND TOOLS

Programming Languages (Years of Experience): Python (5), SQL (2), Java (7), R (1), Arduino (C++) (1), MATLAB (4)
Python Packages: Pandas, NumPy, Scikit-Learn, Numba, PyTorch, TensorFlow, Keras, PySpark, Regex, WandB, Optuna, Dash
Visualization Software: Tableau, Excel, Mathematica, Jupyter Notebooks, Weights and Biases

Data Cleaning and Feature Engineering, SSH + VIM, BASH, Version Control, Image Classification, Recommendation Systems, Distributed Computing, Basic Web Development, Signal Processing, PID Controllers and Global Optimization