jeremy@thaller.dev 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 topics to broader audiences. Specializes in deep learning and natural language processing. Interested in applying mathematical and computational expertise to the business domain.

EDUCATION

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

WORK EXPERIENCE

CELSIUS NETWORK (New York, New York)
Data Analyst – Cryptocurrency Platform Growth and Marketing *[Sept. 2021 – Present]*

- Wrote and deployed a Python package to reduce repeated SQL queries and Pandas DataFrame transformations
- Conducted user surveys to augment data and predict vectors of asset growth
- Presented weekly research insights to stakeholders, steering strategic approaches
- Automated dashboards via python, Sheets API, and BASH scripts.

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

- Created and managed the lab's GitHub organization; constructed example projects to demonstrate best dev practices
- Presented weekly research insights to non-ML-savvy stakeholders; influenced co-workers' strategic approaches
- Reduced simulation compute time by 50x by developing a new statistical-based methodology
- Utilized TensorFlow to predict disorder from absorption spectra, reducing the amount of required data by 90%

YALE UNIVERSITY (New Haven, Connecticut)
Research Intern – Metallurgy *[Summer 2019]*

- Wrote and deployed a GUI Python program to automate and expedite material candidate screening
- Formulated an experiment to isolate the causal variable behind thermo-mechanically molded wire orientations

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

- Automated the complex data analysis process through Python scripts into a click-to-run pipeline Utilized
- Applied MATLAB image processing and multinomial regressions to measure surface tension from 3D image stacks
- Presented project results at multiple international conferences

RECENT DATA SCIENCE PROJECTS

FACEBOOK MESSENGER ANALYSIS

- Scraped 10+ years of messaging data via selenium and BS4; analyzed messaging trends with Pandas, NLTK, SpaCy, and Gensim, showcasing the results with charts and word clouds
- Created a friend classifier through Bayesian statistics, capable of predicting which friend sent an unseen message.
- Built a from-scratch generative chatbot trained on personal messaging data using Keras and GloVe embeddings

SPOTIFY SONG RECOMMENDATION

- Utilized PySpark, scalar-aggregate reduction optimized SQL queries, the Spotify Web API to investigate song trends
- Created a song recommendation algorithm based on song similarity utilizing t-SNE distances and deep embeddings

CRYPTOCURRENCY EDA AND PREDICTIVE MODELING

- Analyzed DOGE Coin value and key financial indicators over time with interactive plots
- Forecast DOGE Coin values in Python using Keras and an LSTM architecture

SKILLS AND TOOLS

Programming Languages (Years of Experience): Python (5), SQL (2), Java (7), R (1), Arduino (C++) (1), MATLAB (4)

jeremy@thaller.dev github.com/jthaller • jeremythaller.com

Python Packages: Pandas, NumPy, Scikit-Learn, Numba, PyTorch, TensorFlow, Keras, PySpark, Regex, WandB, Dash
Visualization Software: Tableau, Excel, Mathematica, Jupyter Notebooks, Weights and Biases

Data Cleaning and Feature Engineering, SSH + VIM, Unix Command Line (BASH), Git and Version Control, Probability and Statistics (Bayesian), Deep Learning, NLP, Image Classification, Recommendation Systems, Transfer Learning, Distributed Computing, Basic Web Development, Signal Processing, PID Controllers and Global Optimization