

(217) 255-3853
Boston, MA
zhang.allen@northeastern.edu

Allen Zhang

<https://github.com/AllenForReal>
[linkedin.com/in/allen-zhang-jobs/](https://www.linkedin.com/in/allen-zhang-jobs/)

EDUCATION

Master of Science *Northeastern University*

September/2022 — 05/2024

Master of Data Science

GPA 3.6

Bachelor of Arts in Physics *University of Colorado Boulder*

August/2017 — May/2022

Interdisciplinary Physics & Computer Science, Business Analytics Minor

Relevant Courses *Linear Algebra, Differential Equations, Electromagnetics, Nanophotonics, Classical Mechanics, Thermodynamics, Statistical Mechanics, Numerical Computing*

SKILLS

Languages and Tools Javascript, SQL, C++, Python, Shell Script, Matlab, Github, Excel, Powerpoint, Tableau

Environment+Framework VScode, Google Colab, Docker, Jupyter notebook, MySQL, AutoCAD, CST

PROFESSIONAL EXPERIENCES

Research Assistant (*Experiential AI*)

Jan/2024 — Present

- Fused 2 large datasets with resources from NOAA, FHWA and LSTW. Imputed then encoded selections of variables based on the study needs. Merged datasets by spatial proximity and longitudinal information, working under the Homeland Security funded SENTRY project.
- Engineered a custom data processing pipeline for a temporal graph neural network. Formulated feature & adjacency matrices for the entire US road network, using Pytorch, Pandas and Geopy. Expanded the input feature from 1 to 7 from earlier publications. Produced an accurate time & space aware model in traffic prediction, surpass the industry standard.

Statistical Programmer (*LLX Solutions*)

June/2023 — August/2023

- Automated SAS to extracted clinical trial data from different sources, impute missing information, normalized tables and derive treatment significance, according to the guideline from FDA's CDISC standard.
- Presented study summary, mythology and important findings with statistics and visualizations. Completed all training on time as well as 2 studies with more than 20 tables in the 10 week internship.

Graduate Research Assistant (*Northeastern Nanophotonics Lab*)

August/2022 — December/2023

- Optimized peer's work on meta-material photonics design with genetic algorithm, optimized neural net construction, reducing test error from 5% down to 2%, improving fabrication quality.
- Modeled meta-material using CST studio suite remotely. Tidied, augmented and parameterized the input data, making the deep learning design angle and polarization dependent. Implemented a custom VAE model, successfully designed an angle independent, frequency shift free laser proof glasses.

Undergrad Research Assistant (*CU Boulder Soft Condense Matter Lab*)

March/2020 — March/2021

- Automated synthesized signal generator using Labview, overcame IEEE, National Instrument to HP standardization and problem, increased experimental efficiency by more than 200%.
- Developed Shell scripts to control different scientific packages like Meep, design to simulate fluorescent materials in electric field.

PROJECTS

Sentiment based Summarization (*Northeastern NLP Project*)

March/2024 — April/2024

- Deployed a Google Chrome extension with Flask REST API, offering a selection of sentiment analysis and summarization models, aim to improve customer experience, simplifying finding the best product based on reviews.
- Developed sentiment analysis models using nltk, TextBlob and BERT. Then summarization model using BART, fine tuned GPT 3.5 and Gemini, all trained on high quality Amazon review dataset.