Mohsin Shah

mohsinposts.com | mohsinposts@gmail.com | 508-292-1312 | linkedin.com/in/mohsinposts | github.com/mohsinposts

Programming Languages: Python, Java, JavaScript, TypeScript, Julia, C, SQL, R, Splunk SPL, HTML, CSS, XML
Frameworks: React, PyTorch, Tensorflow, spaCy, Keras, Next.js, Node.js, Express.js, Sequelize, Scikit-learn, Pandas, NumPy, Scipy, Bootstrap
Tools: Git, AWS (S3, RDS, Lex), Hugging Face, Snowflake, Splunk, Jira, Jupyter, Linux, Excel, Agile, Scrum

Education

University of Massachusetts Amherst **GPA 3.97** **Expected Graduation: Dec 2024**
- **BS Computer Science and BS Mathematics -** Artificial Intelligence, Data Science, and Statistics concentrations

Experience

Fidelity Investments **June 2024 – Aug 2024**
Data Engineering Intern Smithfield, RI
- Created a chatbot for single sign-on service with Amazon Lex, tracked via Jira; aiding efficiency and projected to reduce inquiries by 40%.
- Designed Splunk dashboards for Password Resets & User Registration, identifying friction and abandonment points for millions of users.
- Applied Snowflake SQL tables & AWS S3 to migrate client-facing cybersecurity data, raising data security for 5000+ annual presentations.

Microsoft **Jan 2024 – Feb 2024**
Data Science Intern Cambridge, MA
- Extended Azure ML's Responsible AI Toolbox & Interpret Text for LLMs like GPT-4 & Llama, aiding 200,000+ users in model evaluation.
- Implemented LIME explainers, customizable benchmarking metrics, and error analysis modules in the comprehensive UI dashboard.
- Developed 5 tutorial notebooks showcasing model analysis with HuggingFace (GPT-Neo, RoBERTa) and OpenAI API (GPT-4, 3.5, 3).

University of Massachusetts Amherst **May 2023 – Sep 2023**
ML & NLP Research Intern | *Professor Jaime J. Dávila* | GitHub Code Amherst, MA
- Analyzed multimodal transformer models: BLIP, GIT, CLIP, and custom vision language model (VLM) with BERT (LLM) encodings, EfficientNet (CNN), and LSTMs with PyTorch (CUDA) to generate prompts of AI generated images, achieving a BLEU score of 63%.
- Created training and validation datasets for R&D using Python & Selenium, by web scraping 1000+ AI generated images and prompts.

Biologically Inspired Neural and Dynamical Systems Lab **Feb 2023 – Dec 2023**
AI & RNN Research | *Intern Professor Edward A. Rietman* Amherst, MA
- Built simulations in Julia to study the applications and dynamics of oscillatory neural networks; made computation 10x faster.
- Designed algorithms to solve the ongoing challenge of recurrent neural network oversaturation, with potential applications in robotics.
- Enhanced data visualization with 1200+ raster plots and video heatmaps, integrating clustering algorithms for data segmentation.

University of Massachusetts Amherst **Sep 2022 – Present**
Computer Science, Residential, & Academic Peer Mentor Amherst, MA
- Devised tailored academic success strategies for 600+ students in their transition to college through academic success mentoring.
- Boosted engagement by 75% through collaboration with campus organizations to plan and execute community-building events.

Projects

Travelers Insurance "Best Use of AI" Award Winner | signdecoder.com | Python, OpenCV, Tensorflow, Google Teachable Machine
- Developed an ASL translator using supervised learning and computer vision at Hack(H)er413, optimized to 83% classification accuracy.
- Led the creation of a custom dataset of 10,000+ images using computer vision algorithms to track and capture our hand gestures.

eBay: Machine Learning & Name Entity Recognition (NER) Competition | Python, spaCy, Pandas, NumPy, SMOTE, Scikit-learn
- Created a 94% accurate NER model using 10 million raw & imbalanced eBay listings in German; effectively classifying each word.
- Enhanced quality and searchability of listings using data processing, data analytics, and natural language processing with spaCy.
- Conducted data analysis & data cleaning on raw, non-English dataset; streamlining feature extraction & validity for a F1 score of 87%.

Roommate Finder Full-Stack Web App | GitHub Code | JavaScript, React, Node, Express, MySQL, AWS cloud (RDS, S3), MUI, Tailwind
- Developed & deployed full stack product architecture for users to get matched, go chat, and customize profiles; enhanced UI/UX via 30+ animations and implemented software testing to make 45+ routes & queries for frontend, backend, and database interaction.
- Incorporated Amazon web services to store 1000+ images & MySQL database; optimizing data management & security.

Customer Churn Prediction | GitHub Code | XGBoost, Scikit-learn, Python, SHAP, SMOTE, Pandas, NumPy, Matplotlib
- Applied XGBoost for fast & scalable model training, using SMOTE to balance a 7000 entry dataset; resulting in 84% accuracy & F1-score.
- Analyzed 20 features via XGBoost's feature importance & SHAP; providing actionable insights to reduce churn rate & boost profitability.

Automated Social Media Video Content Creation App (Windows & macOS) | Python, FastAPI, Vercel, MoviePy, FFmpeg, SoX
- Built & distributed app and custom API for licensed users to select a book segment & narrator voice; grew 7500+ social media followers.
- Retrieved background video clips from 4 APIs & synchronized on-screen text with narration, fully automating the video creation process.

Deep Learning & Reinforcement Learning Flappy Bird AI Game | GitHub Demo | Python, NEAT, PyGame
- Created Flappy Bird (60 FPS) via OOP, simulated physics and collisions, and trained AI birds to be unbeatable by the 11th generation.
- Implemented the NEAT genetic algorithm with an evolving artificial neural network design, resulting AI outperforms 99% of humans.

Leadership, Communication, & Courses

- Activities: Vice President of UMass Brazilian Jiu Jitsu club; Member of UMass Machine Learning Club and UMass Wrestling Club.
- Courses: Machine Learning, Artificial Intelligence, Software Engineering, Data Structures and Algorithms, Computer Systems, Statistics, Linear Algebra, Discrete Math, Multivariable Calculus, Differential Equations, Software Developer Project Management (Scrum Master)
- Coursera: Stanford University - Machine Learning Specialization by Andrew Ng