Hammad Ahmad Usmani

(646) 363-6463 — hammadus@proton.me — hammad93.github.io — fluids.ai — US Citizen

Ranked in the top 1% of over 4 million global Python developers on the HackerRank Python Leaderboard

EXPERIENCE

- **Machine Learning Engineer** United Nations   *2024 - Current*
  - Engineer Generative AI solutions for the UN Office of Information and Communications Technology
  - Advise on chatbots, LLM's, AI & ML Ops, RAG, Python, Copliot, prompt engineering, and data science
  - Governs AI enablement across the UN Secretariat with Microsoft CoPilot for more than 37k colleagues
  - Implemented chunking for RAG on Azure Cosmos DB to add document databases into an AI assistant
  - Enhanced RAG architecture for AI assistant by incorporating web search API to address training cutoff
  - Programmed GPT-4 on Azure with Python for UN80 suggestions to identify key themes and efficiencies
  - Facilitate UN initiatives with AI and LLMs to support inclusive development and responsible tech use
  - Developed ChatUN app with Ollama, OpenWeb UI, Llama 4, and nginx on AWS without external API's
  - Provides DGACM AI security and Python development for the ODS AI Assistant with RAG on Azure

- **Artificial Intelligence Engineer** Space   *2021 - 2023*
  - Search product reduced an estimated $700k/year per spaceship based on Lockheed Martin supply chain
  - Created RAG method for search tool using Lucene search index and a BERT-based question answering AI
  - Deployed AI, machine learning, dashboards, and data analytics for Space and Corporate lines of business
  - Engineered original data lakes & pipelines with 99.99% uptime utilizing NASA's TRL operational levels
  - Fine-tuned deep neural networks and statistical models through customer feedback and cross-validation
  - Integrated end-to-end ML with Azure, AWS, CI/CD, Git, Docker, Elasticsearch, REST APIs, & web apps
  - Engineered, trained, and fine-tuned LLM question-answering for quality and manufacturing dispositions
  - Produced & maintained machine learning algorithms through Azure, Python, SQL, sklearn, & Tensorflow
  - Integrated IBM Watson for search, translation, speech recognition, and vision to cloud apps using Python

- **Machine Learning Engineer** Moody's   *2019 - 2021*
  - Deployed web scraping for machine learning and LLMs for NYSE sentiment analysis and financial data
  - Solved analytics and data engineering problems on AWS, Azure, & Google Cloud with Python and Spark
  - Innovated server-less and containerized solutions to deploy and scale ML models with continuous training
  - Developed natural language processing with Python, Spark, Pytorch, deep learning, and language models
  - Specialized in deep learning techniques, including RNNs, CNNs, transfer learning, and cluster analysis
  - Improved AUC scores by 14% on recommendation problems using deep learning, ETL, and Tensorflow
  - Engineered a data lake infrastructure on AWS Athena, ECS, ECR, EMR, and S3 using Scala & Python
  - Developed automated sentiment analysis of NYSE utilizing deep learning from investors social media

- **Software Engineer** MIT   *2018 - 2019*
  - Conducted MLOps, data processing, data science, and machine learning for advanced weather problems
  - Innovated deployment of machine learning on the Azure & AWS cloud realtime globally with Tensorflow
  - Developed software to continuously monitor machine learning performance based on quality requirements
  - Engineered solutions using Python, JavaScript, and SQL with cloud computing to operationalize AI
  - Implemented lossless compression technique to reduce model output size by 99.2% in near real-time
  - Collaborated on the 2018 Best Paper Award from innovations in machine learning and neural networks

- **Data Scientist** Simpluris   *2017 - 2018*
  - Completed 200+ big data cases as a lead data analyst utilizing with SQL, Python, PyTorch, & Tensorflow
  - Produced and calculated analysis with SSRS reports using SQL and Excel for class action lawsuits
  - Improved efficiency of address parsing API algorithm by 97% from linear to logarithmic to linear growth
  - Developed duplication detection algorithm incorporating Levenshtein Distance in Python and Scala

- **Machine Learning Engineer** SHAMAN   *2015 - 2016*
    - Achieved multiple National Science Foundation Innovation Corps grants for IoT and big data analytics
    - Invented algorithms with Python on deep neural networks consisting of chat capabilities for NFC tags
    - Engineered microcontroller prototyping boards with RFID and NFC IoT functionalities in Java, C/C++
    - Discovered novel algorithms in deep learning to forecast hurricanes and tropical storms with LSTM's

EDUCATION

- **Georgia Institute of Technology**   Atlanta, GA US
  Master of Science in Computer Science
- **University of Central Florida**   Orlando, FL US
  Bachelor of Science in Computer Science

CERTIFICATIONS

- **NVIDIA**
  NVIDIA-Certified Professional: Accelerated Data Science   *2025*
- **NVIDIA**
  NVIDIA Certified Professional: Gen AI LLM's   *2025*
- **Microsoft**
  Azure AI Engineer   *2025*
- **Harvard Business School**
  CORe Credential of Readiness, Certificate in Entrepreneurship Essentials   *2024*
- **Google**
  Generative Artificial Intelligence, Machine Learning, Vertex AI   *2023*

SKILLS

- **AI & ML**: RAG, Fine-Tuning, GPT-4, Llama, DeepSeek, Cursor, Claude, Gemini, Grok, Midjourney
- **Programming Languages**: Python, Javascript, Java, C#, SQL, Unity, Oculus VR, iOS, Android, React
- **Data Engineering**: ETL/ELT, NLP, Large Language Models, CI/CD, Tensorflow, PyTorch, Sagemaker
- **Cloud Computing**: AWS, Azure, DataBricks, GCP, Snowflake, Tableau, Docker, PowerBI, OpenAI

PUBLICATIONS

- Patel, A. B., **Usmani, H.**, & Brant, J. C. (2021). *Multivariate LSTM Approach to Hurricane Intensity and Tracking Predictions.* 101st American Meteorological Society Annual Meeting.

- **Usmani, H.**, Habibi, A., & Habibi, D. (2020). *A Deep Neural Network to Globally Forecast the Track and Intensity of Tropical Cyclones.* 100th American Meteorological Society Annual Meeting.

- Veillette, Mark S, Iskenderian, H., Lamey, P. M., Mattioli, C. J., Banerjee, A., Worris, M., Proschitsky, A. B., Ferris, R. F., Manwelyan, A., Rajagopalan, S., **Usmani, H.**, T. E. Coe, J. E. Luce, and B. A. Esgar. (2020). *Global Synthetic Weather Radar in AWS GovCloud for the US Air Force.* 100th American Meteorological Society Annual Meeting.

- Iskenderian, H., Veillette, M. S., Mattioli, C. J., Lamey, P. M., Hassey, E. P., Banerjee, A., Worris, M., Cancio, K., Rajagopalan, S., **Usmani, H.**, Dreher, J. P., Hock, N., & Radovan, J. (2019). *Global Synthetic Weather Radar Capability in Support of the U.S. Air Force.* 99th American Meteorological Society Annual Meeting.

- **Usmani, H.** (2019). *A deep recurrent neural network to forecast the intensity and trajectory of Atlantic tropical storms.* 99th American Meteorological Society Annual Meeting.

- Almalki, H. M., Rabelo, L., Davis, C., **Usmani, H.**, & Hollister, D. (2016). *Analyzing the Existing Undergraduate Engineering Leadership Skills.* SYSTEMICS, CYBERNETICS AND INFORMATICS.