Allen Daniel Sunny

Ph: +1 (858) 241-6155 | Email: allens@umd.edu | LinkedIn | GitHub | Personal Website

ML engineer and researcher specializing in explainability and scalable AI systems. Experienced in building production-grade pipelines, fine-tuning LLMs, and developing tools for interpretable machine learning.

Experience

University of Maryland College Park — College Park, USA
Graduate Research Assistant — *Sept. 2023 – Present*
- Built centralized data system for Career Services to unify job metrics across departments.
- Automated Power BI dashboards, reducing manual reporting time by 60%.

Kantar Analytics (Walmart Account) — Bangalore, India
Data Scientist / Associate Data Scientist — *Aug. 2018 – Jan. 2023*
- Deployed price forecasting across 4,000+ Walmart stores, enabling scalable pricing.
- Modeled promotional uplift with R/Spark, accounting for secondary effects.
- Built anomaly detection pipelines to catch real-time pricing errors and reduce leakage.
- Deployed ML pipelines via Hadoop, Druid, and Airflow for production-grade reliability.
- Partnered with pricing teams to align models with business strategy.
- Fine-tuned GPT-NeoX for SEO-optimized product descriptions, boosting search visibility.

Education

University of Maryland, College Park — College Park, USA
Master of Information Management — *Aug. 2023 - July 2025*

Visvesvaraya Technological University — Bangalore, India
Bachelor of Engineering in Computer Science and Engineering (with Distinction) — *Aug. 2014 - July 2018*

Research Projects

Master's Thesis – Tech Policy Lab — *Jan. 2025 - Present*
- Researching procurement agility, explainability, and public trust in public sector AI.
- Developed and evaluated a legally-defensible XAI prototype for SNAP eligibility.

Oxford AI Policy Group – Systemic Risk Monitoring — *Feb. 2024 - Apr. 2024*
- Designed a scraping pipeline to track LLM adoption across global sectors.
- Supported risk governance research of foundational models at the University of Oxford.

R Packages – Tools for Interpretable Machine Learning — *Jan. 2022 - Jul. 2023*
- **StructuralDecompose**: R package for decomposing level-shifted time series data.
- **TangledFeatures**: R package for feature selection in highly correlated datasets.

Academic Projects

RADAR: Retrieval Augmented Data Analysis and Representation [Code] — *Feb. 2025 - May 2025*
- Built a RAG system using LLaMA 3.1 8B to turn text prompts into visualizations.
- Integrated LangChain and Milvus for schema-aware data retrieval.

Dark Pattern Analyzer [Code] — *Aug. 2024 - Dec. 2024*
- Built a browser extension to detect deceptive UX patterns using DOM analysis and LLMs.
- Included auto-highlighting and GPT-generated inline explanations.

Explainable AI Research — *2024*
- Studied how users interpret AI explanations through interviews and surveys [paper].
- Designed and evaluated tailored XAI interfaces for public-facing decision systems [paper].
- Findings informed design principles for building trust in government AI systems.

Technical Skills

Languages: R, Python, Go, TypeScript, TSQL, VBA
Skills: AI, Machine Learning, Large Language Models, Explainable AI, Prompt Engineering, Retrieval-Augmented Generation (RAG), Agile Development, Data Pipelines, Web Scraping, Data Retrieval, Demand Forecasting
Frameworks: Transformers, Pytorch, LangChain, React, Redux, Node.js (Fastify), Spark, Druid, Hadoop, Milvus
Developer Tools: Anaconda, RStudio, Git, Google Cloud Platform, AWS, Power BI, Tableau, VS Code, Excel