Allen Daniel Sunny

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

Data Scientist specializing in machine learning, explainable AI, and scalable data systems. Skilled in model development, deployment, and translating complex data into clear insights for stakeholders.

Experience

University of Maryland College Park — College Park, USA
Senior Data Analyst (Graduate Research Assistant) – Career Services — *Sept. 2023 – Present*
- Built a unified data architecture for career outcome analytics.
- Automated SQL/Python ETL for 50K+ records, cutting prep time 90%.
- Used user-centered design and iterative testing to ensure system usability and trust.

Kantar Analytics — Bangalore, India
Data Scientist / Associate Data Scientist - Walmart Account — *Aug. 2018 – Jan. 2023*
- Deployed R/Spark demand models for personalized promotions in 4,000+ stores, lifting ROI 12%.
- Developed streaming anomaly detection, cutting disruptions 35%.
- Maintained ML pipelines with Hadoop, Druid and Airflow, processing 2TB+ daily data.
- Optimized GPT-NeoX to generate SEO-friendly product descriptions, boosting search relevance by 15%.
- Conducted A/B tests and backtesting to validate models, ensuring consistent performance across regions.

Education

University of Maryland, College Park — College Park, USA
Master of Information Management — *Aug. 2023 - Dec. 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*
- Developed a neuro-symbolic XAI prototype for legally defensible SNAP eligibility.
- Combined embedding retrieval with rule-based logic for statutory alignment.
- Demonstrated improved transparency and legal traceability in AI-driven benefit decisions.

Human-Centered Explainable AI Research — *Jan. 2024 - Aug. 2024*
- Conducted interviews and surveys to analyze how users interpret AI explanations [paper].
- Designed and ran quantitative evaluations of tailored XAI interfaces [paper].
- Derived design principles to enhance user trust and transparency.

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 – Present*
- **StructuralDecompose**: R package for robust seasonal and trend decomposition.
- **TangledFeatures**: R package for interpretable feature selection (accepted at MLSB @ NeurIPS 2025).

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.

Technical Skills

Languages: Python, R, Go, TypeScript, SQL
Skills: LLM Fine-tuning, Prompt Engineering, Retrieval-Augmented Generation (RAG), Embedding Models, Ranking Models, Model Deployment, Scalable ML, Explainable AI
Frameworks: TensorFlow, Keras, HuggingFace, Faiss, Node.js, React, PyTorch, Spark, Hadoop, Druid, Milvus
Developer Tools: AWS S3, Snowflake, Airflow, Power BI, Databricks, Tableau, Git, VS Code, Docker, CI/CD