Dan Salo

dancsalo [at] gmail.com • (919) 698-2742
Github • LinkedIn • Website

EXPERIENCE

Proofpoint — Durham, NC

Email cybersecurity leader for the Fortune 100. Tech Lead for 7 years on various teams within central AI/ML organization.

Staff Machine Learning Scientist — *May 2024 - Present*

- **Agentic Webpage Sandbox**: Technical architect for new AI/ML team in Cork, Ireland, leading a team of 4 engineers to integrate specialized agents into the dynamic webpage analysis. v0 was designed, implemented, and deployed all in 4 months, thanks to Windsurf, Claude Code and Opus 4.6.
- **URL Defense Oracle**: Developing a multi-call LLM workflow with high precision for a growing number of URL-based threats. Currently offers a second-opinion to SOC analysts; knowledge distillation in process to scale detection in prod.
- **Benign Email Data Generation with Differential Privacy**: Finetuned foundation models using Huggingface, Opacus on email corpus from beta customers. Evaluated generations on traditional metrics and classifier training.

Machine Learning Manager — *April 2022 - May 2024*

- **ML Platform Team**: Designed a new cross-platform model development framework and CLI using Flyte and MLFlow on k8s to replace SageMaker-centric tooling. 8 different projects still leveraging the framework today.
- **DLP ML Team**: Released two features in production: Sensitive Form classification using deep metric learning for 1000's of DLP customers and FP Detection feature using XGBoost and approximate NN for internal triage analysts. Deployed on AWS using Kinesis, ECS, and Lambda.

Senior Data Scientist & Data Scientist — *May 2019 - April 2022*

- **Promissory Images Classifier**: designed and delivered an image classification system for financially promissory image detection. Leveraging object detection, classical ML, and approximate NN. Key differentiator for all finserv customers. Precision/recall maintaining over 24+ months with retrainings and annotation. Created autoscaled ECS deployment.
- **NLP Classifier System**: implemented and maintained and scaled a spaCy-centric framework to serve 54 NLP classifiers over 100Ks social media posts daily across 1000's of customers. Addressed customers FP/FNs weekly.
- **Physical Locations Classifier**: constructed custom dataset of 1000's of spans for an NER model to detect mentions of a location with an address. Selected DistilBERT over BiLSTM-CRF. Enabled for hundreds of Patrol Security accounts.

Automated Insights — Durham, NC

Natural language generation startup before NLP had its Imagenet moment. Lead Data Scientist on two projects.

Senior Data Scientist & Data Scientist — *July 2017 - May 2019*

- **Augmented NLG**: Designed and prototyped a natural language generation system using Node.js, React as an add-on feature for BI dashboards. Published a patent.
- **Contextual Synonym Suggester**: Productionized custom word vectors trained using gensim and Common Crawl for synonym suggestion as a Flask microservice. Core ML feature of Wordsmith product.

LEADERSHIP & RESEARCH IMPACT

- **Organizational Leadership & Architecture**: Co-founded new AI/ML team in Cork, Ireland; hired founding manager and initial team members. Co-authored successful IDA (Irish Development Authority) grant proposal on pre-training cybersecurity models, funded with near-perfect evaluation marks.
- **Hiring & Talent Development**: Directed 2 full hiring processes (junior and staff-level engineering); participated in 12+ additional interview panels across engineering and ML roles. Managed full performance lifecycle for direct reports, including structured coaching and formal improvement plans.
- **Mentorship & Academic Collaboration**: Mentored 4 summer interns; supervised 2 Duke MIDS teams and 1 UC Berkeley team on applied ML research. Contracted with Washington State University on PU Learning research and translated novel methods into production classification training loops.
- **Innovation & External Impact**: Named inventor on 3 patents in applied machine learning for cybersecurity. Placed top 3 in 3 company hackathons. Authored 1 technical blog, delivered multiple external presentations and panel discussions.

EDUCATION

- **M.S., Electrical Engineering** — Duke University — *2014 - 2017*
- **B.S., Biomedical Engineering** — North Carolina State University — *2010 - 2014*