MICHAEL W. SHERMAN

publicresume@michaelwsherman.com · www.michaelwsherman.com · New York Metro Area

SUMMARY

— Technical lead, Machine Learning (ML) engineer, and Data Scientist on multiple successful production ML/AI projects.
— Tech skills (years): Software Development (11), Python (10), SQL/NoSQL Databases (10), AI/ML (8), MLOps (6), Public Cloud (6), Natural Language Processing/NLP (5), Tech Leading (5), R (4), AI Ethics/AI Fairness (4).
— Published AI/ML researcher.
— Flexible communicator (written and spoken) comfortable with technical and non-technical audiences.
— Experience in finance, media, energy, retail, legal services, manufacturing, and healthcare industries.
— Significant international and cross-cultural experience.

RECENT PROFESSIONAL EXPERIENCE

Google – Machine Learning Engineer; *Sunnyvale, California and New York, New York*   December 2017 – Present
- Primary Google engineer supporting Spotify (major customer at $XXXM/year). Collaborate on upgrades to Spotify's company-wide ML infrastructure. Help Spotify execute their AI/ML strategy. Advocate for Spotify to Google Cloud leadership.
- Build custom production ML systems for Google Cloud customers; advise customers (up to C-level) on AI/ML strategy.
- Review AI systems for fairness/ethics concerns, making technical design changes as necessary to uphold Google's AI Principles.
- Built search tools to help US Patent and Trademark Office patent examiners ensure the validity of newly issued patents.
- Built an ML system that makes daily interest rate estimates for $4T of financial assets for a major bank. Led a team of six. After the project, customer spend grew +489% over six months, and the customer became a Google Cloud public reference.
- Architected an ML-based pricing system for an e-commerce platform with $3B+ of annual purchases.
- Designed a Kubeflow/Kubernetes-based machine learning training and serving platform for a Fortune 100 manufacturer.
- Built an NLP system automating pharmaceutical document processing for a healthcare company with $100B+ annual revenue.
- Wrote, with collaborators, three technical blog posts (1 2 3) on Google Cloud products. The posts have over 30k views.
- Managed global AI/ML feature requests from Google Cloud customers, aided Product Managers with feature prioritization.
- Trained over 1000 individuals (Googlers and public) in topics including NLP, Kubeflow, sequence models, and ML in production.
- Co-wrote and maintain Google-wide legal guidance on safe and compliant use of public data.

Bloomberg L.P. – Senior Data Scientist; *New York, New York*   September 2015 – December 2017
- Led data science/ML/NLP for Points of Law, a legal search product mining a 100GB+ corpus of legal text. Led a team of three. Won the 2018 American Association of Law Librarians (AALL) New Product Award.
- Built a (production) NLP pipeline that included structured prediction, unsupervised learning, and graph processing.
- Conducted literature searches and implemented ML models from research articles, making changes to fit our use cases.
- Designed and supervised a $250,000 data labeling project.
- Won an internal technical award for training the first deep learning model in my department.

SunGard Consulting Services – Senior Consultant (Data Science); *New York, New York*   July 2014 – September 2015
- Tech lead of a (patented) $500M stock market surveillance prototype, processing 170M transactions/minute into graphs.
- Wrote multiple sections of SunGard's SEC Rule 613 bid and a whitepaper on our transaction processing system.
- Developed an ML model using SEC EDGAR financial filing data to predict future corporate bankruptcies.
- Designed, built, and deployed a set of analytics and visualizations (Tableau/MySQL) for the CEO of an energy investment fund.
- Received 2015 SunGard Consulting STAR award (awarded to 4 out of 200 consultants).

United States Peace Corps – Resource Specialist; *KaNyamazane, Mpumalanga, South Africa*   July 2010 – September 2012
- Designed and built a Microsoft Access operations database for a rural rape crisis center with over 1,000 annual victim intakes.
- Organized a series of public events promoting science and scientific careers reaching over 7,000 people.
- Elected by fellow Peace Corps Volunteers to advise Peace Corps leadership on policy and strategy.

OLDER PROFESSIONAL EXPERIENCE

- Statistical programmer (MATLAB) for a Neuroscience research lab.
- Head of billing ($10M+ annually) and part-time electronic medical record customization (PHP/MySQL) at a surgical clinic.

EDUCATION

The University of Texas at Austin	Master of Science - Information, Risk, and Operations Management	July 2014
Northwestern University	Bachelor of Science - Communications	June 2003

RECENT PUBLICATIONS, OPEN SOURCE CONTRIBUTIONS, AND PUBLIC SPEAKING

- L. Barrett and M. Sherman, "*Improving ML Training Data with Gold-Standard Quality Metrics*," in KDD '19: 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Workshop on Data Collection, Curation, and Labeling (DCCL) for Mining and Learning, 5 August 2019, Anchorage, Alaska, USA [Online]. Available: https://ai.google/research/pubs/pub48575.pdf.
- Contributor to Gensim (a popular Python library for NLP on large text datasets).
- L. Barrett et al., "*Civil Asset Forfeiture: A Judicial Perspective,*" in Proceedings of the Data For Good Exchange 2017, 24 September 2017, New York, New York, USA [Online]. Available: https://arxiv.org/abs/1710.02041.
- Productionization of Machine Learning, Panelist at Texas Analytics Summit, September 2018
- Practical Machine Learning for Financial Services, Keynote at UiPath Together NY, July 2018