William Beckman

http://www.willbeckman.com

Email : will@willbeckman.com

Mobile : +1-410-236-7933

U.S. Citizen

Machine learning engineer with an emphasis on applied ML.

Education

Georgia Institute of Technology — Atlanta, GA

Master of Science in Computer Science - Specialization in Machine Learning; GPA: 3.9 — *Aug. 2019 – Dec. 2021*

Relevant coursework: Machine Learning, Reinforcement Learning, Deep Learning

University of Maryland, College Park — College Park, MD

Bachelor of Science in Computer Science; GPA: 3.76 — *Aug 2012 – May. 2016*

Experience

Appalachian Trail Thru-hike — April 2022 – August 2022

Successfully completed a thru-hike of the Appalachian Trail.

Capital One Financial Corporation — July 2016 – March 2022

Senior Machine Learning Engineer — *Feb 2019 - Mar 2022*

- **Risky Merchant Detection**: Worked to valuate and productionize a model to identify and flag risky merchants, estimated to save $5-10M/yr under various reasonable assumptions. Model valuation numbers were dependent on ranges of operational assumptions (population subsets, merchant restriction rules, best/worst case scenarios).
  Responsibilities included writing the valuation framework given input from business analysts and working alongside junior engineers to create the daily production ETL using pyspark on Kubernetes.
- **Ring/Graph-based Fraud Detection Algorithm**: Improved and productionized an algorithm to detect credit card ring fraud, resulting in an estimated incremental savings of $10M/yr. Improvements were mostly achieved by propagating updated first-party fraud risk scores through the graph of accounts and by intentionally targeting risky population subsets. Responsibilities included running experiments to target different population subsets and productionizing the updated algorithm.
- **Third-Party Fraud Model**: Assisted in the development of a new account takeover credit card fraud prevention model, resulting in over $15M/year in savings. Responsibilities included expanding the model's feature sets to include additional user actions, developing ETL reports to ensure data integrity, and validating the model/feature stability in production.
  Additional responsibilities included working with product and analytics teams to scope project, partnering with other fraud engineering and infrastructure teams to ensure accurate deployment, and mentoring junior engineers.
- **Fraudbook Data Pipeline**: Productionized batch data pulls to populate Fraudbook, an internal application to aid manual fraud investigation. Responsibilities included writing system tests to guarantee batch load functionality upon code changes and converting the existing data update model from complete database overwrites to delta batch updates, saving tens of GBs of database writes/updates per day and significant hours of compute costs.

Software Engineer — *Feb 2018 - Feb 2019*

- **Corporate Strategy Data Pipeline**: Worked on a data pipeline for credit bureau data for the Corporate Strategy division. Data pipelines processed 100s of GBs of raw data and converted them into formats consumable by analysts. Implemented a light domain-specific language in Scala for business analysts to use in order to simplify their code and reduce the burden of having them learn Spark.

Associate Software Engineer — *July 2016 - Feb 2018*

- **Athena**: Briefly assisted in the development of a React front-end of an in-house ontology-based graph database.
- **AWS Network Debugger**: Created a CLI utility using Boto3 that diagnosed network connections between two AWS resources from an AWS infrastructure standpoint (security groups, NACLs, VPC peering). Diagnostic issues were printed to the terminal in a human readable format to explain what was misconfigured between the instances.
  The tool was used by company-wide staff architects to quickly validate whether or not AWS resources were able to communicate on particular ports across VPCs. This automation greatly reduced the human effort and human error of a once-very-tedious process.

Programming Skills

Software/Infrastructure: Python, Scala, AWS, git, Airflow, Docker
ML/Data: SQL, pyspark, pandas, numpy, scikit-learn, Databricks, Snowflake