Brenton Mallen

2707 Mayfield Drive, Graham, NC 27253
561-376-6580 | brentonmallen1@gmail.com | brentonmallen.com

Profile

Ocean engineer turned data scientist with an eclectic background and experience in applying data processing and machine learning techniques to various applications.

Recent Work Experience

Senior Machine Learning Engineer | Capital One   Oct 2022 - Present

As part of the relatively new software division at Capital One, I am currently working on Snowflake Warehouse configuration based query performance prediction. This includes: research and exploration of the data available as well as generating data for the purpose of training a machine learning model query performance predictor, planning each step along the way for the purposes of, transparency and setting appropriate expectations for stakeholders and knowledge transfer to the rest of the team, designing a product to provide customer value in the form of Warehouse configuration recommendations while articulating risk, and developing said product as a scalable, resilient, low-latency serverless RESTful microservice.

Senior Machine Learning Engineer | Ethyca   Sep 2021 - Oct 2022

As the first machine learning oriented hire, I have been researching, designing, and developing a database-centric personally identifiable information (PII) detection and classification framework to aid in the annotation of privacy data. The framework has both API and CLI wrappers and can be used as a python module. The system takes advantage of the content in each table in the database as well as the contextual information surrounding said table for a multi perspective classification approach. I work with key stakeholders to establish product requirements and priorities as well as demonstrate progress.

Senior Machine Learning Engineer | DTN   Nov 2020 - Sep 2021

Acting technical product owner and manager of a team of three engineers. Created and maintained data science dashboards using the streamlit framework. Created and maintained scheduled data processing pipelines using Apache Airflow and AWS with Slack integration. Built and maintained internal data processing tooling to facilitate data research and data cataloging. Participated in multiple hackathons exploring ideas from weather impact on transportation, satellite imagery for crop yield prediction, and the application of Super Resolution neural network to enhance weather forecast modeling.

Data Science Mentor | SharpestMinds   Nov 2020 - Present

Developed a mentorship curriculum to take a mentee through the entire data science life cycle from product definition to end-user consumable product. First half focus on ideation, data processing, model selection, training and deployment followed by API app development to serve the model output. Facilitated mock interviews and periodic presentations with other data science professionals. Continued support through job search and beyond.

Senior Data Scientist | Syngenta   Feb 2019 - Nov 2020

Architected and built an event-driven, microservice centric backend pipeline and api that supports both streaming and ad hoc acquisition and analysis of satellite imagery. Performed research into crop type detection using image processing and machine learning techniques, including convolutional and recurrent neural networks. Interface with the team's biggest external customer on a regular basis to discuss interfaces, future work and communicate expectations.

Data Scientist | Distil Networks   **May 2015 - Nov 2018**

Engineered both real time and batch machine learning systems to detect malicious internet traffic at a rate of 3,000 records per second with sub-minute classification latency. Performed ad hoc analysis and research on large datasets (hundreds of millions of records) to remain state of the art in bot detection and mitigation techniques. Developed tooling and microservices for team and internal customers to ease the interface with the data science team and their products. Develop reproducible, clean systems and research code complete with unit and integration testing and continuous integration and deployment. Visualize and present system efficacy and research output to both technical and non-technical audiences. Lead culture group that focused on building trust and promoting agile practices across engineering and leadership.

Technical Skills

Data Analysis and Machine Learning:
- Random Forest, Regression, SVM, KNN, K-Means Clustering, Neural Networks, Natural Language Processing

Software and Programming:
- **Languages:** Python, SQL, Matlab
- **Data Processing:** scikit-learn, numpy, scipy, pandas, sklearn, keras, nltk, spacy
- **AWS Infrastructure:** EMR, EC2, ECS, Lambda, DynamoDB, API Gateway, Kinesis, Athena, S3
- **Database Tools**: Hive, Impala, Presto, PostgreSQL, Snowflake

Soft Skills

- Comfortable presenting to leadership and public speaking
- Able to work well in team environments
- Able to design, build products, and lead tasks with minimal oversight

Education

Florida Atlantic University
Master of Science in Ocean Engineering   **Fall 2005 - Spring 2012**
Bachelor of Science in Ocean Engineering

Publications & Presentations

"Proud Elastic Target Discrimination Using Low-Frequency Sonar   *U.S. Navy Journal of Underwater*
Signatures"   *Acoustics*

"Global Botnet Detector"   *PyData NYC 2015*