Gary L Clark Jr.
gclarkjr5@gmail.com || +17177986916

Bachelor's of Science, Mechanical Engineering
Howard University Class of 2013

Summary

- Data Scientist w/ 6+ years of experience from a fast-paced environment adding tangible value exploring data and implementing Statistical & Machine Learning algorithms
- Data Engineer w/ 4+ years of experience architecting, developing, and maintaining data ecosystems for large scale operations

References provided upon request

Technical Skills

- **Data Science**: R, Python, Spark, ML, Tensorflow, Apache Superset
- **Data Engineering/DevOps**: Python, SQL, Powershell, Azure DevOps, Apache Arrow YAML
- **Data Storage**: Data Warehosue, Data Lake, SQL Server, PostgreSQL, Apache Iceberg, Dremio
- **Environments**: Azure, AWS, Windows, Linux, WSL, Raspberry Pi, Docker, Kubernetes

Professional Experience

Wolters Kluwer ELM Solutions — Sep 2018 - Dec 2021

Data Engineer

- Lead the development effort for extracting client data into internal & external Data Lake solutions
- Automated data pipelines for data warehouse solution including performance improvement efforts leading to 300% data ingestion efficiency
- Lead the initiation of Version Control & CI/CD implementation around code and data management for multiple teams
- Developed an end-to-end pipeline that relied on client metadata to automate and build several layers of Data Processing and a client facing Dashboard

Data Scientist & Web Developer

- Developed & Presented a Machine Learning application that provided smart guidance to the Legal Space regarding attorney rate requests compared to the recommended rate of the ML model
  - Extracted ML models' coefficients to build view of how much each inputs contribution was to recommendation
- Lead a business initiative to improve revenue forecasting by leveraging Machine Learning
  - Improved Business Analyst's forecasting abilities by ~30% on average

Rolls-Royce Digital — Sep 2013 - Sep 2018

Data Scientist

- Developed time-series analytics rooted in Data Science to detect, predict, and prevent failures w/ heavy industrial equipment

    - Analytical accuracy of +75% resulted in cultural shift of customers to prioritize analytics
- Enhanced statistical techniques that aid in the production of predictive modeling and analytics
    - Resulted in 10% increase in lead time
- Used Machine Learning techniques to enhance analytics and discover insights in data
    - Supervised/Unsupervised, Linear/Logistic Regression, Random Forest, ARIMA, K-Means
- Utilized a wholistic approach to Data Science workflow that became a standard practice amongst team:
    - Exploratory Data Analysis, Data Manipulation, Data Visualization, Feature Selection/Engineering, Model Creation, Model Evaluation
    - Increase value of team to help communicate better understanding of data to customers
- Developed R & Python ETL scripts and translated to complex SQL stored procedures and dynamic queries

Technical Projects

- Data Delivery Optimization
    - Reduced an ad-hoc data intensive operation process from an estimated delivery time of 100+ days to 1 hour
- HomeBrewing
    - Achieved 80% accuracy classifying beer styles based on user submitted home recipes
    - Increased accuracy after using a nested model that classified styles per different brew processes
- Power Consumption Detection
    - Accurately detected periods of power consumption from noisy electrical outlet data