**M. Zachary Bierbaum (484) 888-6373 | zach.bierbaum@gmail.com | linkedin.com/in/zach-bierbaum**

***Security Clearance***

Active TS/SCI

***Software & Technology Skills***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Python | Apache Spark | Hadoop | Kafka | Zeppelin Notebook |
| R Studio | Tableau | ElasticSearch | Data Mining | Docker |
| SQL | Jupyter Notebook | AWS/Cloud | Data Analysis | Graph Databases |

***Experience***

***Data Scientist******–******Noblis,*** *Reston, VA Jan 2016 – Present*

* Conceived and built an end-to-end framework to rapidly search large amounts of log data with minimal latency. This framework utilized the Apache Hadoop/Spark infrastructure as well a light-weight Python API.
* Designed a distributed microservice architecture to integrate multiple separate data sources in order to perform schema-free data enrichment requests
* Produced a supervised machine learning analytic to detect potential malicious applications utilizing features in the secure connection between two entities.
* Developed analytic to perform time-series frequency analysis to detect attacks on DNS servers and determine if any had been compromised by identifying patterns in DNS queries.
* Created an unsupervised machine learning analytic to detect SQL injection, or other malicious behavior, inside strings in different internet traffic logs. This analytic employs a unique anomaly detection algorithm, which was modified to be applied with the Apache Spark infrastructure.
* Trained logistic regression classifier machine learning model with specifically engineered features to detect malicious subdomains with 97% accuracy.
* Developed similarity network graph and accompanying API to search over 200 million unique nodes. The data was processed in Apache Spark and stored in JanusGraph graph database, with the entire application made to be a Docker microservice.
* Reverse engineered an existing analytic and translated it to Python and SQL, which sped up analytic run time from hours to seconds.
* Produced interactive Tableau dashboards to aggregate and display analytic results, both on desktop client and Tableau Server.
* Created SQL-based analytics to run on customer databases to identify and classify internet traffic data into previously unknown categories.

***Consultant – Noblis,*** *Reston, VA**Jan. 2014 – Jan. 2016*

* Developed interactive timeline website tool to display pertinent information for the International Classification of Diseases -Revision 10 (ICD-10) for client to present to the healthcare industry.
* Prepared and facilitated weekly high-level meetings where top issues were discussed and action items are created for the project.
* Developed reports and executive dashboards regarding ICD-10 readiness for the client to present to higher levels within the government.
* Periodically conducted extensive research and evidence collection to assess the status of the stakeholders that the client supports.
* Created presentations for the client to use in webinars, national conference calls, and other industry stakeholder meetings.

***Education***

**M.S. Data Analytics Engineering** *May 2018*

George Mason University – Fairfax, VA

**B.S. Aerospace Engineering; Minor in Mathematics** *May 2013*

Virginia Polytechnic Institute and State University – Blacksburg, VA