**Aboubacar Keita**

Parkville, MD| akeita20@proton.me | (301)-820-2401 | [github.com/akeita11](https://github.com/akeita11)|[linkedin.com/in/aboubacarkeita20/](https://www.linkedin.com/in/aboubacarkeita20/)

**EDUCATION**

University of Maryland, Baltimore County

**MPS in Data Science**, *GPA: 4.0*

University of Maryland, Baltimore County

**BS in Information Systems**, *GPA: 3.**43*

**SKILLS**

*Programming Languages:* Java, C++, Python, SQL, PL/SQL, Scala, R

*Software*: Apache Spark, Microsoft SQL Server, MySQL, AWS S3, Eclipse, Tableau, Jupyter Notebook, Anaconda, Google Collaborate

*Libraries*: TensorFlow, Pandas, NumPy, Seaborn, Scikit-Learn, Matplotlib, SciPy, NLTK, Re, Beautiful Soup, Apache Hadoop,

PyDP (Differential Privacy), PM4PY, EasyGA

*Modeling Techniques:* Exploratory Analysis, Statistical Analysis, Data Mining, Data Analysis, Data Visualization, Linear Regression,

Logistics Regression, Random Forest, Support Vector Machines, K-Means Clustering

*Other*: Deep Learning, Cloud Computing, NLP, Sentiment Analysis, Big Data, Genetic Algorithms, Process Mining, Critical Path Method

**RELATED EXPERIENCE**

**Data Scientist Intern**, Remote May 2022 – August 2022

*Startup Crafters*

* Modeled a Convolutional Neural Network on a large dataset with a 96% accuracy to classify different sets of images
* Visualized a learning curve to represent the classification accuracy and cross-entropy loss
* Customized the Talking Head feature using Generative Adversarial Network through GitHub

**RELEVANT PROJECTS**

**Designing and Evaluating Critical Path Method and Genetic Algorithm Methodologies to Assess Critical Activities Within an Organization using Process Mining**

*Capstone*

* Researched the effects of implementing Critical Path Method and Genetic Algorithm into the field of Process Mining
* Implemented the concept of both Critical Path Method and Genetic Algorithms on an event log
* Utilized the open-source library PM4PY for doing process discovery, conformance checking, and performance analysis on an event log
* Generated different event log from the two methods
* Generated visual representation of difference in time duration of events, resources involved, activities performed
* Compared the results generated by both methods and provide future work to be performed
* Assessed the usability and effectiveness of Critical Path Method and Genetic Algorithm in the Process Mining field
* Concluded that Genetic Algorithm performed the best but precise tuning in its fitness function to properly work.
* Concluded that the Critical Path Method has it own effectiveness but under certain condition and objectives.
* Collaborated and supervised by Zeanique L. Barber, who provided with valuable resources in completing this project

**Bitcoin Cryptocurrency Coin Price Prediction**

*Platforms for Big Data Processing*

* Created a data lake in the cloud AWS comprising of cryptocurrency data that can be easily accessed by data scientists, business analysts, and data engineers for decision-making purposes
* Utilized AWS CLI to export datasets from a local computer into the AWS S3 bucket
* Compressed the dataset from a CSV to Parquet file format for easier data manipulation and faster processing
* Performed a Linear Regression to predict and visualize the future coin price of multiple cryptocurrency coins
* Compiled statistical metrics to evaluate the model using PySpark’s Regression Evaluator
* Performed the Silhouette Method to find the optimal number of clusters of three
* Built a K-Mean Clustering of three and visualized it using PCA

[**Machine Learning in High School Graduation Success**](https://github.com/akeita11/Machine-Learning-in-High-School-Graduation-Success)

*Data Analysis and Machine Learning*

* Collaborated with six data scientists to determine variables affecting graduations rates in the nation and predict successful H.S. completion threshold using regression analysis
* Identified and classified specific schools with a threshold greater than 67% as successful per ESSA legislation marker standards
* Visualized the model’s findings using the Seaborn library for ease of interpretation and communication

[**Flood Risk Level Modeling**](https://github.com/akeita11/Flood-Risk-Level-Modeling)

*Introduction to Data Science*

* Assessed the flood risk level of Maryland’s National Harbor in conjunction with five teammates
* Built a predictive model to project the sea level over five years in regions surrounding the MD National Harbor
* Visualized the findings with heatmaps, scatterplots, and boxplot

[**Metro System Simulation**](https://github.com/akeita11/Metro-System-Simulation)

*Database Application Development*

* Simulated a railroad metro card system using PL/SQL procedure on Oracle SQL Developer to simulate its real-life counterpart for travelers entering a metro station and paying associated fees for travel
* Interpreted user input (location, time of departure, idle time, age) to determine estimated arrival of Metro Train giving travelers accurate travel expectations by collaborating with three teammates to create a superior simulation outcome

[**Balagan Financial Company** **Business Analysis**](https://github.com/akeita11/Balagan-Financial-Company-Business-Analysis)

*Structured Systems Analysis and Design*

* Developed three different solutions for Balagan Financial Company based on financial analysis and narrowed them through a collaborative selection process
  + The first solution was to buy new networking equipment.
  + The second solution was to buy more of the same networking already implemented
  + The third solution was to mix and match by upgrading the necessary equipment causing a bottleneck
* Implemented the third solution to reduce their overhead cost and financial waste by 10%, while staying in the organization’s budget

**ADDITIONAL EXPERIENCE**

**Driver Associate** November 2019 – November 2022

*George Cassel Logistics, Hanover, MD*

* Decreased route time by two hours through effective package arrangement when loading and unloading the van

**Supervisor** June 2014 – November 2019

*Smokey Glen Farm, Inc, Gaithersburg, MD*

* Boosted customer satisfaction by 15% by immediately resolving guest problems and complaints
* Ensured smooth kitchen operation while managing twenty-five crew members

**LEADERSHIP – Board of Officer**

*Historian of Tau Kappa Epsilon Fraternity Inc - Upsilon Theta Chapter* December 2019 - December 2020

* Produced at least ten newsletters for chapter members and alumni to inform them of current organizational activities
* Increased alumni engagement by thirty percent through updating them with photos or videos of chapter events for the TEKE magazine
* Coordinated the development of two composites and cooperated with local alumni associations to provide programming for the TEKE alumni association