Alain Kuiete   **alainkuiete@gmail.com**
Data Scientist   **856-540-8179**
New York, NY

Summary of Relevant Experience:
- Well versed with **R, Python, Pyspark, SQL, NoSQL**, programming languages.
- Proficient with **Exce**l and **PowerPoint.**
- Holding Master of Science in Data Science from CUNY School of Professional Studies.

Summary:
- Data Scientist , with Data mining, Big Data with Hadoop and Pyspark,, Data Analytics.
- An insatiable intellectual curiosity, and the ability to mine hidden gems located within large sets of structured, semi-structured, and unstructured data. Ability to leverage, mathematics, high-level programming, statistics with visualization, with a healthy sense of explanation. All Combine with:
- 2+ years' experience in design, implementation, architecture discussion, building of brand-new pipelines with technology stack including R, Python, Pyspark, SQL, Hadoop with NoSQL.
- Proficient with R, Python, Pyspark, SparkR, SQL, NoSQL
- Proficient in Data Science and Machine Learning, Supervised/unsupervised Learning.
- Proficient in Feature Engineering, Modeling Techniques (XGBM, LGBM, RF, Regression)
- Knowledge of AWS, AZURE ML, GOOGLE CLOUD ML
- Knowledge of Docker Containers and Kuberneties
- Proficient with Excel and PowerPoint.
- Excellent communication and problem-solving skills.

Skills:
- Experience in machine learning based predictive modeling using: Gradient Boosting, Random Forest, Linear Regression, Generalize Linear Modeling, Logistic Regression
- 3+ years experience in R, Python, SQL
- 2+ years experience in Pyspark, SparkR and Hadoop Ecosystem
- 2+ year experience in GCP, AWS and AZURE
- Strong General programming skills and knowledge in Python , R and java
- Strong experience and expertise in applied statistics (Regression, Classification, Segmentation, Clustering, Association Rule, over fitting, modeling, lasso and Ridge regression, Cross validation, Bootstrapping, Bagging, randomization, Boosting, Multinomial logic regression, Knn, Kmean).
- Strong communication skills.
- Ability to work with different Machine Learning libraries: NumPy, Pandas, scikilt-Learn , statsmodels ...etc. / frameworks: anaconda for integration.
- Programming Languages: R, SparkR, Python, PySpark, Tensor flow, SQL,
- and Java,
- Operating System: Window & Linux applications,

Professional Experience:
Client: **Advanced Talent Solution New York, NY**
November 2014 – June 2018
Position: **Teacher Assistant**
Responsibilities:
- Analyzed and predicted meaningful knowledge from educational data that help for flexible planning, execution and prediction of information

- Collected and analyzed data from divers source to provide information and improve the quality of teaching and learning. That resulted in understanding students' behavior, assisting instructors in lesson planning, improving curriculum and learning system.
- Used data analytic tools to support continuous assessment on students' progress. Classification was used to analyze students' behavior and performance. Estimation predicted the outcomes of persistence, retention and course success.

Education:
M.S Data science, CUNY School of Professional Studies: Manhattan, NY
B.S Mathematics & Statistics Science, Lehman College: Bronx, NY
A.S Teacher Education, Bronx Community College: Bronx, NY

Certifications:
Neural Network and Deep Learning Online with Coursera.
Data Camp Certificates on SQL, Machine Learning, Python, R, Recommended Systems
Network:
https://alainkuiete.shinyapps.io/UN_Asembly_votes_by_year_country/

https://rpubs.com/AlainKuiete

https://covid-19-mobility-reports.herokuapp.com/