## Godfred Somua – Gyimah, PhD

[godfred.s.gyimah@gmail.com](mailto:godfred.s.gyimah@gmail.com) ● Saint Louis, MO ● (573) 201-7316 ● github: [ggyimah1031](https://github.com/GGYIMAH1031) ● [linkedin.com/in/godfredsomuagyimah](https://www.linkedin.com/in/godfredsomuagyimah)

**Legal Status: US Permanent Resident**

---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### **Work Experience**

#### **MONSANTO COMPANY (acquired by BAYER AG)** Saint Louis, MO

***Senior Data Scientist (Monsanto Emerging Leaders in Science Program),*** *January 2018 to present (1 year)*

* Translating business problems into research questions for exploratory, descriptive, predictive and prescriptive analytics.
* Combining UAV field imagery with environmental data to build plot-level yield prediction models for cotton, corn and soybean.
* Developed deep learning models for seed vigor prediction with ~99.6% sensitivity and ~90.1% specificity using Python, Keras and CNN. Developed seed clustering insights using K-Means and Jupyter Notebook.
* Developed a dryer recommender model for optimizing seed quality using DataRobot, R and the Random Forest algorithm.
* Provided technical support for pilot phase model deployments. Collaborated with business leaders constantly to update / agree on project goals and expectations. Created visualizations to communicate model results to both technical and non-technical audiences. Documented projects on Github and prepared technical reports.
* As an ELS scholar, my major leadership task for 2018/19 is to co-ordinate integration efforts within Data Science to promote cross-team collaboration and similar common best practices across all Data Science teams in the R&D organization.

#### **INSIGHT DATA SCIENCE** Manhattan, NY

***Artificial Intelligence Fellow,*** *July 2017 – October 2017* *(4 months)*

* Performed topic modeling and sentiment analyses on the 2016 POTUS election-related tweets using the Twitter Streaming API, NLTK and SAS E-Miner.
* Implemented the 3D CNN action recognition model by [Schindler et al. (2008)](https://www.vision.ee.ethz.ch/publications/papers/proceedings/eth_biwi_00532.pdf) using Python, Keras and the KTH video dataset

**MISSOURI UNIVERSITY OF SCIENCE & TECHNOLOGY**  Rolla, MO

***Doctoral Researcher (Geomechanics, Numerical Modeling & Machine Learning)****, August 2014 – June 2018* *(4.5 years)*

* Developed and validated a novel method for calibrating the geomechanical parameters of DEM using XGBoost and RStudio. The method combines simulation of tri-axial rock testing with XGBoost to achieve prediction accuracies of up to 95.54%. Developed and validated a 3D numerical model for studying the failure patterns of geomaterials during rock excavations.

#### **HUAWEI TECHNOLOGIES** Accra, Ghana

***Costumer Analytics Manager,***  *October 2011 – September 2013 (2 years)*

* Provided formal mentoring and leadership to a team of 5 analysts. Led exploratory, descriptive and predictive costumer analytics studies to discover insights and opportunities. Led end-to-end analytic projects for revenue forecasting, customer segmentation, customer churn and sentiment analyses. Made presentations to technical and non-technical audiences.

***Engineering Data Analyst,*** *January 2010 – November 2011 (2 years)*

* Produced exploratory, descriptive and predictive analyses for material usage patterns, geotechnical investigations and construction project cost forecasting. Interfaced with clients and created visualizations to communicate analyses / recommendations. Presented and explained technical recommendations to multi-disciplinary audiences.

### **Education**

#### **MISSOURI UNIVERSITY OF SCIENCE & TECHNOLOGY** Rolla, MO

#### . ***PhD Mining Engineering*** *(GPA: 4.0 / 4.0)* Aug. 2014 – Dec. 2018

#### ***Graduate Certificate, Business Analytics & Data Science*** *(GPA: 4.0 / 4.0)* Aug. 2016 - May 2017

#### *(Courses: Data Mining & Machine Learning, Data Visualization, Text Mining, Business Analytics & Data Science)*

#### ***MS Mining Engineering*** *(GPA: 4.0 / 4.0)* Aug. 2014 - July 2016

#### **UNIVERSITY OF LEEDS** Leeds, England

#### ***MS Engineering Geology*** Sep. 2013 – Aug. 2014

#### **KWAME NKRUMAH UNIVERSITY OF SCIENCE & TECHNOLOGY** Kumasi, Ghana

#### ***BS Civil Engineering*** Aug. 2006 - June 2010

#### 

### **Technical Skills**

* ***Languages*​:** Python, R, SQL, Matlab
* ***Statistics, Machine Learning & Deep Learning*:** Tensorflow, Keras, Scikit-Learn, Caret, Weka, Azure ML, Amazon ML, DataRobot, NLTK, NumPy, SciPy, SAS Text Miner, RStudio, H2O
* ***Data Wrangling & Storage*:** MySQL, Apache Hive, Pandas, R Dataframe, Trifacta
* ***Data Visualization*:** Tableau, Power BI, ggplot2, matplotlib
* ***Cloud Computing:*** AWS, MS Azure, Domino, Paperspace
* ***Other skills:*** Jupyter Notebook, Git, Flask, Docker, Hadoop, HDFS, Apache Spark

### **Profile Summary** (based on advertised role requirements)

* 9+ years combined experience in analytics research and industry applications.
* 5+ years industry experience in predictive analytics.
* Prior experience leading end-to-end data science project implementation and managing a small team with direct reports.
* Prior experience identifying, developing, and advancing new products and businesses in an R&D environment.
* Demonstrated experience solving problems using machine Learning, predictive analytics, regression, classification and clustering techniques (decision trees, SVM, Neural Networks, k - NN, Naive Bayes, Decision Trees, Random Forests, etc.)
* Experience utilizing Deep Learning and Natural Language Processing (NLP) techniques.
* Strong ability to effectively communicate highly complex and technical results to a diverse audience. Expert ability to break down and clearly define problems. Strong interests in AI / machine learning innovation in the healthcare industry.

### **First Author Publications**

**Journal Articles**

* **Somua-Gyimah, G**., et al., A machine learning approach to Distinct Element Model calibration for earth material. International Journal of Constructive Research in Civil Engineering. 2019. In Press.
* **Somua-Gyimah, G**., et al., Formation fragmentation modeling and impact on dragline excavation performance in surface mining operations. International Journal of Mining Science, 2019. Volume 5, Issue 1: p. 11-21. [[PDF]](https://www.arcjournals.org/pdfs/ijms/v5-i1/2.pdf)
* **Somua-Gyimah, G**., et al., A Material Flow Model for Dragline Bucket-Formation Failure Analyses Using the Distinct Element Method. International Journal of Mining Engineering and Technology, 2018. 1(1): p. 1-15. [[PDF]](https://uniquepubinternational.com/wp-content/uploads/2018/11/UPI-IJMET-2018-1.pdf)

**Theses & Dissertations**

* **Somua-Gyimah, G.** Dragline Excavation Simulation, Real-Time Terrain Recognition and Object Detection. PhD Dissertation. Missouri University of Science & Technology. 2018. [[PDF]](https://www.researchgate.net/publication/330117889_DRAGLINE_EXCAVATION_SIMULATION_REAL-TIME_TERRAIN_RECOGNITION_AND_OBJECT_DETECTION)
* **Somua-Gyimah, G.** Finite Element Modeling of The Proposed Tunnel for the York Potash Mineral Transport System. MSc Thesis. University of Leeds. 2014. [[PDF]](https://www.researchgate.net/publication/325794854_Finite_Element_Modeling_of_The_Proposed_Tunnel_For_The_York_Potash_Mineral_Transport_System)

**Peer-reviewed conference papers**

* **Somua-Gyimah, G**., et al., A computer vision system for terrain recognition and object detection tasks in mining and construction environments. Proceedings of the 2019 Annual Conference of the Society for Mining, Metallurgy & Exploration (SME). 2019. In Press. [[Preprint - PDF]](https://www.researchgate.net/publication/330130008_A_computer_vision_system_for_terrain_recognition_and_object_detection_tasks_in_mining_and_construction_environments)

### Extra keyword summary**:** analytical, solr, reporting, experimental, fraud detection, big data, data science techniques, open source tools, data processing, data sets, physics, data piping, design, computer science, continuous improvement, distributed computing