

EMMANUEL OYEKANLU

+1-267-581-8688 | manuelbomi@yahoo.com | Big Flats, NY | linkedin.com/in/emmanuel-oyekanlu-6ba98616/

PRINCIPAL DATA SCIENCE ENGINEER

Exceptional and proven experience in big data engineering, full stack machine learning (ML) and data science projects. Successful professional with diverse experience in delivering several impactful data engineering, ML and software products for manufacturing, education, telecommunications and entertainment industries. Proven success in several data engineering projects using GCP BigQuery, Azure, Google or AWS Databricks (based on data specifications, customers' need and use cases). Data engineering projects have always successfully achieved the required data discovery, ETL, ELT, ACID, data security, data management and data privacy specifications. ML models are always developed with greater than 94% accuracies with an outlook towards end-to-end product delivery and reliability while factoring-in resilient Cyber-security methods in product development. Always open to collaborate with and learn from other data science, engineers and other industry professionals while also adding quantifiable values across different product value chains. Core competencies include:

Databricks Products | SQL/NoSQL Databases | Dbt | Machine Learning & AI Applications | GCP | BigQuery Products | AWS | Azure | Computer Vision & Generative AI Applications | OpenCV | Cross Functional Leadership | Cascading Style Sheets (CSS) | Incident Response | Threat and Vulnerability Management | Algorithms | Process Improvement | Software Development

EXPERIENCE

VerzX (Client: Zintro), NY, USA

October 2024 – Ongoing consulting relationship

Consultant - DevOps Architecture for Edge Computing & IoT. System Reliability & Risk Assessment

VisasQ/Coleman , NY, USA

August 2024 - Ongoing consulting relationship

Consultant - AI: Technology Trends and Industrial Applications

GIS , NY, USA (remote office at Tokyo, Japan)

May 2024 – June 2024

Consultant - Digital Transformation (DX) Specialist (Remote/Online from the US)

Architecture and design specifications of secure and data privacy template for successful digital transformation for large and medium scale enterprises.

Data Engineering, Data Management, Data Privacy, Data Science, and Cybersecurity specifications for successful integration of digital transformation for companies and government parastatals.

Corning Incorporated, Painted Post, NY

September 2019 - January 2024

Software and Systems Project Engineer

Project Engineer - Generative AI , August 2023 - January 2024

Development of a customized ChatGPT interface for use in industrial environment using Docker, PostgreSQL, Vue and Django framework.

To be able to use company's customized dataset, an instance of GPT4All was deployed with GCP's (Google Cloud Platform) Vertex AI using the Vertex AI Workbench instance on Ubuntu OS.

Software Engineer, March 2023 - August 2023

Developed anomaly detection systems to identify fault contours on glass surfaces for life sciences, automotive and manufacturing applications.

- Worked with other engineers using Databricks, SQL, Spark, Python, VSCode and Jupyter Notebook to develop solutions for online detection of faults on glass surfaces during manufacturing.
- Data engineering work using Python for ELT application for Bronze layer data. Evolved the datasets to Silver layer data sets for creating the anomaly detection toolbox.

- Developed a Windows desktop App for the anomaly detection system using Tkinter/PyInstaller. The App usage led to a 31% product waste reduction and better engagement with our customers.
- During app development, for CyberSecurity assurance, users' possible ranges of inputs were validated against a white list of acceptable input values to prevent injection attacks.

Machine Learning (NLP) Engineer, October 2022 - March 2023

Developed a NLP based model app using big time-series text data to predict faults types on ceramic products for automotive, manufacturing and life sciences applications.

- The NLP model was based on Spacy, Gensim, TF-IDF and Latent Semantic Indexing (LDF) libraries. NLP models were used along with LSTM time series training to develop reliable models. Model could successfully predict fault types and suggest likely reason(s) for faults.
- Data Engineering work to refine and curate the datasets using Scala, IntelliJ IDEA & SBT
- The model reduces time spent by engineers for fault identification from weeks to minutes, an approximately 97% reduction in average fault identification time.
- The app was deployed on GCP's Cloud Run. Time series data was funneled into GCP's BigQuery and Cloud Storage. BigQuery ML and GoogleSQL was used to derive additional insight from the time series data.
- During app development, for CyberSecurity assurance, users' possible ranges of inputs were validated against a white list of acceptable input values to prevent injection attacks. Generative Adversarial Networks (GANs) generated images were used to train the network so that the model will be able to identify spoofing attacks.

Machine Learning Engineer, April 2020 - December 2020

Designed of a Cordierite gasoline particulate filters image recommendation system that reduced fault inferencing time from approximately 2 weeks to 3 minutes. Extensive use of OpenCV based machine vision tools for life sciences, automotive and manufacturing applications.

- Developed an images recommendation system using VGG 16 and text input prompts trained with large language models (LLM) using Gensim and Spacy libraries. Perioperative Risk Assessment for ML models.
- Extensive big (images) data engineering task. Databricks was used for ELT applications and the model was designed using Databricks MLFlow. Flask was used to design a robust user interface for the system.
- To assist engineers in product manufacturing and development, a dictionary-based system was designed for big images and text data curation.
- During model development, for Cybersecurity assurance, comparable images from Generative Adversarial Networks (GANs) were used for model training and hardening.

Project Engineer, January 2020 - April 2020

Developed a deep learning/IoT based solution for Covid-19 social distancing application in industrial environment.

- YOLO algorithm, Nvidia Jetson Nano (IoT device), Nvidia Xavier and Raspberry Pi were used to design a Covid-19 social distancing enforcement tool in the industrial environment.
- Initial ELT using Python to curate training datasets. Perioperative Risk Assessment for ML models.
- For company-wide analytics and insight, data from the Nvidia's IoT device was streamed onto a GCP pub/sub instance on GCP. Cloud Dataflow and Apache Beam was used for data transformation. BigQuery Analytics was using to understand areas of concentrated social-distancing violations in the company. IoT data was stored on GCPs Cloud Storage. Dashboard was provided by using GCP Cloud Run.
- The solution designed was 93% more affordable than a comparable solution from an external vendor.
- Developed a vaccine vial tracking solution for healthcare application using Hough Transforms, image segmentation and OpenCV. Solution could successfully track vaccine vials on the design pipeline across several product development/manufacturing stages.

Project Engineer - Smart Manufacturing, September 2019 - April 2023

Led a cross-functional team comprising of engineers from Corning, Verizon, Amazon AWS, and Gestalt Robotics (Germany) to develop several data and AI-based use cases for Corning private 5G network applications.

- Automated Guided Vehicles (AGV) and Autonomous Mobile Robots (AMR) monitoring and management through IoT devices and sensors' data harvesting and transmission over private 5G network on the factory floor.
- IoT data analytics for remote AGV/AMR vehicles' diagnostics. IoT/sensors' data sets are harvested from devices, and systems on the AGVs. Analytics and insights are accomplished on AWS IoT Events platform.
- Real-time monitoring of AGV/AMR battery charge levels.
- Worked with academic researchers from Cornell University to discover methods of achieving near zero-waste for several factory floor processes.
- AGV & AMR-based product monitoring (cart and optical fibers cable) application using Docker and YOLO algorithm with over 97% product classification accuracy. Monitoring was accomplished across the factory floor. Private 5G was used as transmission network. AWS was used for data integration at system backend complete with alerting and dashboard system built in-house.
- Extensive Proof of Concept (PoC) and Proof of Values (PoVs) work regarding the utilization of AGV/AMR traffic to institute smart factory applications on the industrial floor using 5G communication networks.
- During the project, an IEEE journal paper jointly developed with Verizon Engineers is available here: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=S-jTMfkAAAAJ&citation_for_view=S-jTMfkAAAAJ:BqipwSGYUEgC
- During the project, an approved patent that I developed at Corning with another Engineer is available here: <https://patents.google.com/patent/US11878422B2/en>
- A combination of batch and streaming data emitted from industrial machines, Python, Spark, Airflow, VBA, Jupyter Notebook, Bokeh and Power BI were used to create a reactive and reliable dashboard that reduced product faults identification time from approximately 2 weeks to 7 seconds.

Comcast, Philadelphia, PA

March 2019 - September 2019

Data Science Engineer

Worked with a team that prototype ideas for new machine learning and artificial intelligence algorithms. We developed and supported data pipelines for system's optimization.

- Analyzed massive amounts of data both in real-time and batch processing utilizing Databricks, Presto, Spark, Kafka, and AWS technologies such as Kinesis, S3, Elasticsearch, and Lambda.
- Worked with a team that analyzed network health data of approx. 29 million devices in addition to customers' clickstream data to identify network issues and perform root cause analysis leading to improved Comcast customers' experience.
- Product led to a decrease in average time for resolving customers' issues. It also reduced customer churn on Comcast products.

Drexel University, Philadelphia, PA

September 2018- March 2019

Adjunct Professor (Applied Physics)

Taught college students on vectors and their applications to real life and industrial use cases. Other subjects taught include: statics, kinematics, and system dynamics, including Newton's laws, torque, projectile motion, circular motion, work, power and energy, impulse and momentum, rotation and computational physics.

- Applied and Computational Physics with emphasis on big data technologies
- Developed and evaluate students' projects and examinations.
- Consistently score high ratings in student evaluations. An example is indicated here: <https://www.ratemyprofessors.com/professor/2240559>

In Schools at Sweden and USA for M.Sc. & Ph.D. studies

August 2008 – August 2018

Ph.D. thesis was focused on implementing embedded wavelet-based statistical machine learning applications at edges of large Industrial Internet of Things (IIoT) and large networks such as the electrical

powerline, smart cities and the smart grid using a low-cost, fixed-point digital signal processor (DSP) called C28x and Embedded C programming language.

Waveforms implemented with C programming language on the C28x are comparable to waveforms available in standard Matlab toolboxes. Some of the waveforms were successfully used to characterize:

- (i) ECG (electrocardiogram) signals (healthcare applications) &
- (ii) Low-voltage electrical system powerline (below 400V).
- (iii) Grid voltage characterization for large non-linear loads such as EV vehicles. Stability effects of large non-linear loads with emphasis on EV vehicles.

Wavelets were also used to develop deep learning applications useful for improving the powerline system's performance at network edges near the consumer application layer. Suites of Embedded C based wavelets that I developed are the very first in the industry for the C28x DSP.

As part of my PhD studies, I worked briefly with the US Department of Energy to understand the effects of wide-scale usage of large non-linear constant power loads such as EV vehicles and other large direct current battery systems (e.g. large data centers) on the US electricity grid.

Smart grid and smart cities applications.

Effects of smart vehicle technologies such as Vehicle-to-Infrastructure (V2I) applications on the grid.

Examination of 3GPP and ITU standards for smart grid and smart cities applications.

ADDITIONAL RELEVANT EXPERIENCE

Rechez Communications, Awka, NG

Feb 2005 – August 2008

Network Data & Systems Engineer

During my time as Network Data Engineer with Rechez Communications, I oversaw network data analysis for real-time network stability monitoring and consumer traffic data assessment. I use Linux (Fedora) to create a system of network monitors using streaming data from iRG-S2 and TMI Transmodulator.

Diverse network data from the LAN, VSAT, consumers, WAN and the Wide Area Network of ISP (Ariave Satcom, Isreal) is locally monitored and evaluated for predictive and prescriptive network site reliability and stability analysis.

I helped Rechez Communications to secure a license from the NCC that enable Rechez Communications to expand their Internet coverage area to a wide area in the hinterland. I also helped Rechez to reduce monthly Internet bandwidth charges by about 42% by eliminating a middle Internet Service Provider (Kit Technologies) and linking directly with the primary ISP (Ariave Satcom) in Israel.

TECHNICAL SKILLS

Backend Applications:Python | R | C# | Golang (Go) | MATLAB | C++ | CUDA | OpenCL | Dask | VBA | dbt | jinja | Scala | SBT | Akka | Java | PySpark | SQL | Vitess | GoogleSQL | promQL | SparkSQL | KQL (Azure) | Hadoop | Kotlin | Flink | Apache Beam | Kafka | Hive | Pig | Storm | Zookeeper | MailChimp Programming (API)

Frontend Applications:JavaScript | RESTful APIs | Vue | Nuxt3 | Ruby | CSS | HTML | Flask | PyQt5 | Tkinter | PowerBI | Tableau | GraphQL

Cloud Applications:BigQuery | GCP | Vertex AI | GCP Cloud Run | GCP Pub /Sub | GCP Dataflow | Azure Databricks | GCE | GKE | AWS Databricks | Snowflake | Oracle ERP Cloud Platform

Databases:PostgresQL | SQLite | MongoDB | Kdb+ | Neo4J (Cypher) | Cassandra | ScyllaDB

DevOps:Docker | Kubernettes | Git | GitLab | Airflow | Terraform | Jenkins | Github Actions

Dev. Platforms:VSCode | .Net | Jupyter Notebook IntelliJ IDEA | Powershell | Linux

Machine Learning Libraries:TensorFlow | Keras | RAPIDS

Machine Vision Libraries:OpenCV | Halcon.

Site Reliability: Uptime Kuma | Prometheus | Datadog | Grafana

Cybersecurity Tools:Linux | Splunk (SIEM) | TCP/IP | Suricata | Wireshark | TCPDump

General Applications: Programmable Logic Controllers (PLC) applications, Embedded systems (C28x Fixed Point Real Time Digital Signal Processors and other Texas Instruments Suite of devices), Raspberry Pi for IoT and IIoT applications, Nvidia suite of IoT devices, etc.

WEBSITE

<http://emmanueloyekanlu.com>

OPEN SOURCE PROJECTS

<https://scholar.google.com/citations?user=S-jTMfkAAAAJ&hl=en>

<https://github.com/manuelbomi/>

EDUCATION

Ph. D. in Electrical Engineering, Drexel University, Philadelphia, USA

Master of Science (M. Sc.) in Electrical Engineering, Blekinge Institute of Technology, Karlskrona, Sweden

Master of Science (M. Sc.) in Signal Processing, Blekinge Institute of Technology, Karlskrona, Sweden

Master of Science (M. Sc.) in Telecommunications, Blekinge Institute of Technology, Karlskrona, Sweden

Bachelor of Technology (B. Tech) in Computer Engineering, Ladoke Akintola University of Technology, Nigeria

Data Science - Data to Insights (Massachusetts Institute of Technology - MIT) (Coursework), USA

CERTIFICATION

Data Engineering:

Databricks Certified Data Engineer Professional

Databricks Certified Data Engineer Associate

IBM: ETL and Data Pipelines with Shell, Airflow and Kafka

Machine Learning:

Databricks Certified Machine Learning Associate

Databricks Generative AI Fundamentals

Cybersecurity: Google Certified Cybersecurity Professional (Risk & Cybersecurity Assessment)

Project Management and Leadership: Yellow Belt - Project Excellence (Corning Incorporated, NY) | Green Belt - Project Excellence (Corning Incorporated, NY) | Black Belt Training Classes (Corning Incorporated) | Leadership Principles - Harvard University

VOLUNTARY ACTIVITIES

Board Member: IEEE Bridge HKN Magazine: Editing, sourcing for contents and reviewing

AWARDS

- **Corning Innovation Award:** 5G PoC, Use Case Development and Testing
- **Corning Quality Award:** IP Recognition (Corning Inc.)
- **Corning Performance Excellence Award:** Design of Recommendation System (AI) for cordierite filters (DESIGN)
- **Drexel University Teaching Excellence Award-** <https://drexel.edu/graduatecollege/news-events/graduate-student-day/2017-winners/>
- **Dean's Fellowship Award, Drexel University (2014)**
- **Freshmen Engineering Fellow Award**, Drexel University
- **Dean's Fellowship Award**, Drexel University (2016)

- **Schweitzer Engineering Laboratory (SEL)** - Equipment Award (for Ph.D. studies)
- **National Science Foundation (NSF)** - Sponsorship Award as Entrepreneurial Lead for the NSF project "Digital and Scientific Literacy and Training Solutions for Green Energy and Environmental Engineering"

PATENT AND SELECTED PUBLICATIONS

(<https://scholar.google.com/citations?user=S-jTMfkAAAAJ&hl=en>)

- E. A. Oyekanlu, W. P. Thomas III, "**Integrated Manufacturing Systems Architecture**," US Patent 11,878,422; Nov. 2022 (<https://patents.google.com/patent/US11878422B2/en>)
- Emmanuel Oyekanlu, David Kuhn, Grethel Mulroy, "**Data Engineering** for the Factory of the Future: From Factory Floor to the Cloud - Part 1: Performance Evaluation of State-of-the-Art Data Formats for Time Series Applications" IGI Global, April 2023 (DOI: 10.4018/978-1-7998-7852-0.ch001)
- Emmanuel Oyekanlu, David Kuhn, Grethel Mulroy, "**Data Engineering** for the Factory of the Future, Multimedia Applications and Cyber-Physical Systems: Part 2 - Algorithms and Python-Based Software Development for Time-Series Data Format Conversion," IGI Global, April 2023 (DOI: 10.4018/978-1-7998-7852-0.ch002)
- Book (Editor): Applied AI and **Multimedia Technologies for Smart Manufacturing** and CPS Applications, IGI Global, April 2023 (<https://www.igi-global.com/book/applied-multimedia-technologies-smart-manufacturing/266358>)
- E. A. Oyekanlu et al., "A Review of Recent Advances in **Automated Guided Vehicle Technologies**: Integration Challenges and Research Areas for **5G-Based Smart Manufacturing Applications**," in IEEE Access, vol. 8, pp. 202312-202353, 2020, doi: 10.1109/ACCESS.2020.3035729. (IEEE Access; January 2020)
- E. Oyekanlu, "**Predictive edge computing for time series of industrial IoT and large scale critical infrastructure based on open-source software analytic of big data**," 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 2017, pp. 1663-1669
- E. Oyekanlu, "Distributed Osmotic Computing Approach to Implementation of **Explainable Predictive Deep Learning at Industrial IoT** Network Edges with Real-Time Adaptive Wavelet Graphs," 2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Laguna Hills, CA, USA, 2018, pp. 179-188
- E. Oyekanlu, S. Onidare and P. Oladele, "Towards **statistical machine learning for edge analytics** in large scale networks: Real-time Gaussian function generation with generic DSP," 2018 First International Colloquium on Smart Grid Metrology (SmaGriMet), Split, Croatia, 2018, pp. 1-6
- E. Oyekanlu, "Osmotic Collaborative Computing for **Machine Learning and Cybersecurity Applications in Industrial IoT Networks** and Cyber Physical Systems with Gaussian Mixture Models," 2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC), Philadelphia, PA, USA, 2018, pp. 326-335.
- E. Oyekanlu, K. Scoles, P. Oladele, "Arbitrary Waveform Generation for **IoT and Cyber Physical Systems Communication Networks** Using C28x TMS320C2000 Digital Signal Processor", IEEE 10th Int'l Conf. on Advanced Infocomm Tech., Stockholm, Sweden, August 2018
- J. Uddin, E. Oyekanlu, C. Hong Kim, J. Myon Kim, "**High Performance Computing for Large Graphs of Internet Applications Using GPU**", Journal of Multimedia and Ubiquitous Engineering, Vol. 9, 2014
- Emmanuel Oyekanlu, "**Fuzzy Inference Based Stability Optimization for IoT Data Centers DC Microgrids: Impact of Constant Power Loads on Smart Grid Communication over the Powerline**", Journal of Energy, Vol. 68, Issue. 1, pgs 11-21