Submitted to Amentum for ARPA-H STATS

NAME	JOB TITLE	PWS Tasks
Veronica Boateng	ITDI - Intermediate Data Scientist	7.1.2; 7.1.10; 7.1.16; 7.4.5; 7.5.1; 7.5.5

Job Responsibility	Qualifying Skills
Data Analysis and Insights Development: Analyze internal data to identify trends, generate insights, and support evidence-based decision-making across the agency and programs.	· Experience conducting quantitative analyses to extract meaningful insights and drive data-driven decision-making. · Expertise in statistical methods, data visualization, and trend identification to support program initiatives. · Strong background in handling large datasets and performing exploratory data analysis for research and operational improvements.
AI/ML Model Development: Design and implement machine learning models and AI solutions to optimize agency operations and program effectiveness.	· Proficient in machine learning model development, training, and optimization, with a focus on deep learning techniques. · Experience working with AI/ML frameworks and libraries such as TensorFlow, PyTorch, and scikit-learn. · Hands-on experience in developing predictive models and AI-driven automation solutions for complex data challenges.
Policy and Strategic Planning Support: Provide data science expertise to support program development and strategic planning.	· Extensive knowledge of data governance, regulatory frameworks, and ethical AI principles for policy alignment. · Proven ability to collaborate with stakeholders to develop AI and data policies that support strategic objectives. · Experience in drafting and reviewing technical policies, ensuring compliance with federal guidelines and industry best practices.

Education

Ph. D., Computer Science, Bowie State University, Bowie, MD **2019 – December 2024**

MS, Computer Science, Wichita State University, Wichita, KS **Jan 2000 – May 2003**

BS, Computer Science, Kwame Nkrumah University of Science and Technology, Ghana **1990 - 1994**

Experience

Pinnacle Logic LLC, Laurel, MD, Data Scientist **07/2024 – 12/2024**

- Develop AI models leveraging techniques such as quantization, sparsity, and distillation for retail and real estate applications, achieving enhanced operational efficiency and increased revenue growth.
- Created NLP-based models using NLTK, Spacy, transformers, Deep Learning sentiment analysis to review customer satisfaction levels on retail products and services.
- Developed semantic search solutions using Vector Databases to store and retrieve embeddings for text and documents, significantly enhancing the search capabilities of the organization's information systems.
- Designed and implemented a question-answering model utilizing LLMs for response generation for storing contextual knowledge, providing users with relevant information in real time.
- Developed anomaly detection systems for fraud prevention by leveraging LLMs to generate normal and abnormal behavior embeddings for efficient anomaly identification.

- Utilized LLMs to create embeddings for documents for clustering and classification based on semantic similarity, enabling efficient document management and retrieval for the company.

Tools and Technologies: Natural Language Processing, LLM, Vector DB, Python, R, TensorFlow, PyTorch, sci-kit-learn, Git/GitHub/GitLab, AWS, GCP, Azure, Docker, Kubernetes, Tableau, Power BI, Matplotlib, Seaborn, Latex

Bowie State University, Bowie MD, AI/ML Researcher **08/2022 – 05/2024**

Industrial Robotic:

- Developed and trained robotic arms to execute pick-and-place tasks using policy gradient methods combined with deep reinforcement learning, optimizing processes for sorting, packing, and material handling in logistics and warehousing operations. Improved learning efficiency with transfer learning and environment simulation.

Transportation:

- Created NLP-based models using NLTK, Spacy, Transformers, and deep learning for sentiment analysis to review customer satisfaction levels on passenger train transportation.
- Created generative AI models for medical image analysis and tampering detection and classification of 3D CT scan benign and malignant cancerous real and fake images for medical diagnosis using AI algorithms, Keras, PyTorch, TensorFlow, Google Collab, Vertex AI, Python, and Eigen for efficient matrix operations, and Halide for optimizing image processing pipelines which outperformed existing baseline models by 10% in performance.
- Designed and created data pipelines for processing and analyzing Fitbit wearable device data to remotely monitor patients' health and activity levels.
- Created medical image analysis and classification models of 2D skin cancer images for medical diagnosis using deep learning ensembles and vision transformer models

Banking and Finance:

- Created algorithms with generalized linear model's stats models API for predictive analysis of daily closing stock prices.
- Created generative AI models for deep-fake mobile cheque deposit detection using GANs and Variational Autoencoders.

Forensics and Security:

- Created crime scene reconstruction for anomaly detection with Fitbit data on Android phone devices using ADB, FTK tools, SQLite browser, and Java on Linux
- Created generative AI models for Digital image forensic analyzer detecting AI-generated fake images using IoT devices.

Virtual Reality:

- Develop a real-time multiplayer virtual reality game using Unity 3D that incorporates realistic physics, immersive 3D models, and players interacting with virtual objects and environments.

Tools & Technologies: LaTeX, HTML, SQL, PyTorch, TensorFlow, Keras, PySpark, GCP, Azure, ROS, OpenAI Gym, IBM AI Fairness 360, NLTK, Spacy, Transformers, Deep learning, neural networks, Regression, classification, clustering, time-series forecasting, hypothesis-testing, optimization, simulation, clinical decision support, predictive analysis, exploratory data analysis, supervised, unsupervised learning and Generative AI and LLM frameworks.

Allwyn Inc, Herndon VA, Data Science/Data Analyst **06/2022 – 09/2022**

Transportation Domain:

- Built ML algorithms and data pipelines, and developed Spark ML models for statistical, qualitative, and quantitative analysis of passenger railroad data and preprocessed raw passenger train operations data for modeling using Python, pandas, PySpark, and SQL improving data accuracy by 10%.
- Created sentiment analysis models using Azure Synapse analytics and big data pipelines, improving customer satisfaction analysis by 15%.
- Built deep learning models to predict Amtrak passenger train delays, outperforming existing baseline models by 35%.
- Created data visualizations using Python, tableau, Power BI, and Qlik stage, improving KPI tracking and data analysis by 10%. Tools & Technologies: SQL, Python, Spark, Azure Synapse Analytics, AWS, Tableau, Power BI, Qlik Stage, regression, classification, clustering, time-series forecasting

Bowie State University Bowie, MD, Adjunct Faculty Data Science Instructor **08/2019 - 12/2024**

Education domain:

- I taught undergraduate courses in Data Science, Python, Java, and C++ programming courses.
- Prepared instructional materials and managed curriculum development.
- Applied technological options for online and course-related software.

Courses Taught: Data Science, Python, Java, C++, Data Concepts

Pinnacle Logic Technologies Data Scientist **01/2018 - 07/2019**

- Led data analysis initiatives and provided insights into key business trends.
- Developed and implemented machine learning models for predictive analytics for birth registry data, resulting in a 15% increase in operational efficiency.
- Led a team in conducting exploratory data analysis, feature engineering, and model validation.
- Built Spark ML models and performed statistical analyses to improve data insights.
- Created sentiment analysis models using Azure Synapse Analytics, improving customer satisfaction analysis by 15%.

Tools & Technologies: Python, Spark, Azure Synapse Analytics

Information Technology Agency, Application Software Management Consultant **01/2008 - 12/2017**

- Worked with a team of developers and business analysts engaged in a $72 million project for the creation of a new Company Registration, Tax Revenue, Correspondence Management web-based systems, Document management systems, e-commerce applications, and e-health management system achieving a 10% increase in revenue collections.

Business Connexion, Durban, South Africa Principal Software Developer **01/2005 - 12/2007**

- Led a team of 5 developers engaged in a $3-million project for the creation and maintenance of a new apparel wholesale manufacturing system to manage merchandising, sales, and product development divisions.
- Developed Customer Inquiries and Sales Order Tracking Systems with modules for customer inquiries management, progress points, and inventory monitoring and generation of reports using Oracle Forms 6i, Reports 6i, Oracle Project Builder, PL/SQL, and Oracle database (10g) running on HP-UX.

PUBLICATIONS Authored the following publications published in IEEE Xplore and Access Journal (2023)

A Global Modeling Pruning Ensemble Stacking with Deep Learning and Neural Network Meta-Learner for Passenger Train Delay Prediction

Ensemble Stacking with the Multi-Layer Perceptron Neural Network Meta-Learner for Passenger Train Delay Prediction

Technical Skills

- Predictive/Diagnostic Analysis
- Computer Vision/ LLM
- Natural Language Processing
- C++ /Python/ Java/Ruby Unity 3d
- Virtualization and Programming
- AI/ML/ Spark ML
- Intelligence gathering, Latex
- Data Modeling Design
- Statistical analysis (SAS, R)
- MySQL/SQLite3/ MongoDB
- Spark, Hadoop, Kafka, Google Cloud Platform, snowflake, Excel, MapReduce, Hive
- SQL (Redshift, MySQL, PostgreSQL)