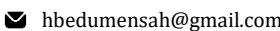
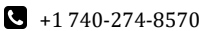


HENRY BEDU-MENSAH



PROFESSIONAL SUMMARY

AI Data Science professional with 7+ years of experience building scalable machine learning, deep learning, and generative AI solutions. Expertise in Databricks, PySpark, AWS, Azure, and LLM frameworks (LangChain, RAG). Proven record of deploying production-ready AI systems that deliver measurable business impact, including reducing churn by 15%, optimizing costs by 10%, and accelerating executive decision-making by 40%.

PROFESSIONAL EXPERIENCE

AI Data Scientist

Deloitte LLP | Jul 2022 – Present | Remote

- Designed, trained, and deployed ML, deep learning, and generative AI models in Databricks (PySpark, Python), automating decision systems for Fortune 500 clients and reducing manual decision-making time by **30%**.
- Built and optimized real-time AI pipelines with Snowflake and SQL to process structured/unstructured data, improving predictive accuracy and reducing latency by **35%**.
- Developed LLM-powered solutions (LangChain, RAG, embeddings) to extract insights from business documents and customer feedback, cutting analysis time by **50%**.
- Created executive-level Tableau dashboards with AI-driven insights, accelerating strategic decision-making by **40%**.
- Performed extensive data cleansing, applied business logic, and executed transformations using PySpark and SQL, improving model accuracy and ensuring high-quality analytics for production AI systems.

Key Tools: Python, PySpark, Databricks, Snowflake, SQL, Tableau, Scikit-learn, RAG, LangChain

Data Science Advisor

Cigna/ Express Scripts | May 2020 – Jun 2022 | Remote

- Led development of predictive models in Python to optimize customer retention, reducing churn risk by **15% across 1M+ members**.
- Applied NLP (sentiment analysis, topic modeling) on customer feedback, driving **10% improvement in personalization accuracy**.
- Embedded AI-driven workflows into business operations, generating insights from **100M+ healthcare records** to inform product design.
- Spearheaded logistics optimization with ML algorithms in Python and SQL, reducing operational costs by **10% annually (~\$5M savings)**.

Key Tools: Python, Scikit-learn, SQL, Tableau

Data Scientist

Blount Fine Foods | May 2017 – May 2020 | McKinney, TX

- Developed ML models to predict campaign response rates, increasing customer engagement by **20%**.
- Built regression models to forecast campaign profitability, guiding **multi-million-dollar marketing investments**.
- Designed and deployed ETL pipelines with SSIS, reducing data processing time by **50%**.
- Delivered real-time marketing intelligence dashboards in Tableau for executive decision-making.

Key Tools: Python, Scikit-learn, Pandas, SQL, SSIS, Tableau

Data Analyst (ML Focus)

Ventura Foods | Jun 2016 – May 2017 | Saginaw, TX

- Applied data wrangling and statistical modeling (Python, R) to uncover performance trends and improve forecasting accuracy.
- Optimized SQL reporting pipelines, improving query efficiency by **30%**.
- Automated reporting workflows and conducted UAT testing, improving adoption of analytics systems.
- Trained business teams on data-driven decision-making, boosting cross-team analytics adoption.

Key Tools: Python, R, SQL, Tableau

Research Data Scientist

Ohio University | Aug 2013 – May 2015 | Athens, OH

- Built regression models in R to analyze ecological data for environmental management projects.
- Partnered with Ohio Department of Natural Resources to deliver data-driven remediation strategies for acid mine drainage projects.
- Managed large-scale research databases, ensuring accuracy and integrity for multi-year studies.

Key Tools: Python, R, SQL, Tableau, Excel

TECHNICAL SKILLS

Programming & APIs: (Python | R | SQL | PySpark | Pandas | NumPy | Scikit-learn | REST APIs)

Machine Learning: (Predictive Modeling | Classification | Clustering | Model Optimization | Statistical Analysis)

Deep Learning: (Encoder-Decoder | Transformers | TensorFlow | Keras | Hugging Face)

Generative AI: (LLMs (GPT | OpenAI | Gemini | Claude) | LangChain | RAG | LangGraph | LLM Fine-Tuning | AI Agents)

Natural Language Processing (NLP): (Sentiment Analysis, Topic Modeling, Text Classification, Conversational AI, Embeddings, NLTK)

Big Data & Data Engineering: (Databricks | Apache Spark | Snowflake | Delta Lake | ETL Pipelines | Google BigQuery | AWS S3 | MS SQL Server | PostgreSQL)

Visualization & Reporting: (Tableau | Power BI | Streamlit | Jupyter Notebook | Matplotlib | Seaborn | Plotly | Excel)

Collaboration & Tools: (Git | Jira | Confluence | Agile/Scrum)

EDUCATION

Post Graduate Diploma in Data Science

Aug 2019 – Feb 2020

Emeritus / Columbia Engineering – New York, USA

- Relevant Coursework: Machine Learning, Data Engineering, Big Data Analytics, Statistical Modeling.

Master of Science in Environmental Science

Aug 2013 – May 2015

Ohio University – Athens, USA

- Relevant Coursework: Environmental Data Analysis, Research Methods, Climate Modeling.

Bachelor of Science in Biochemistry

Sep 2001 – Jun 2005

Kwame Nkrumah University of Science and Technology – Kumasi, Ghana

- Relevant Coursework: Biostatistics, Molecular Biology, Analytical Chemistry, Bioinformatics

CERTIFICATIONS

- Databricks Certified Data Engineer Associate