Tahiya Rahman
Data Analyst

Research Triangle Park, NC | 984.212.0889 | toc764@gmail.com | [LinkedIn](#) | [GitHub](#)

Experienced **Data Analyst** with 6 years in Fintech, Biotech, Edtech, and telecommunication industries, adept at leveraging data for strategic decision-making and driving innovation. Transitioning to a Junior Data Scientist role focused on advanced analytics and solution-driven approaches.

Technical Experience

Topic Modeling: LDA, NMF, BERTopic, LSA, TensorFlow | **Programming Languages:** Python, SQL, R, Java| **Databases:** MySQL, NoSQL, PostgreSQL, MongoDB, Oracle | **Machine Learning Algorithms:** Pandas, NumPy, Scikit-Learn, Gensim Generative AI, LLM, Pattern Recognition, Deep Learning | **Data Visualization:** Tableau, Power BI, Matplotlib, Seaborn | **Transferable Skills:** A/B Testing, SEO, Clustering, Statistics

Education

Master of Science in Data Analytics & Bachelor of Science in Data Analytics | Western Governors University | 2024
Bachelor of Business Administration in Digital Marketing | North South University | 2019

Professional Experience

Data Analyst (Contract) | BASF |RTP, North Carolina | *Jan 2024 - May 2024*
Tech Stack: Azure Data, KNIME, Profisee, Tableau, Power BI, Excel
- Resolved **anomalous behavior** and errors by 18% within the master data processing through detailed research data collection, data quality.
- Validating data **geolocation and GIS/GLM** to ensure data accuracy and relevance, enhancing B2B lead quality with an 80% acceptance rate.
- Consistently identified and **flagged fraudulent accounts**, validated to reduce erroneous entries by 15%, and improved data integrity.

Data Science Researcher (Contract) - Enterprise Risk Management | Bread Financial |Ohio | *May 2023 - Sep 2023*
Tech Stack: Dataiku, MicroStrategy, Archer, Email Automation
- Created **Predictive Models** for repeating issues of customer banking, risk, and threats analysis using clustering unlabeled data.
- Performed topic modeling using Non-Negative Matrix Factorization (NMF) and Latent Dirichlet Allocation (LDA) to cluster text data for advanced text mining.
- Utilized Python libraries (gensim, nltk, scikit-learn) to generate topics, optimize coherence scores, and classify key risk issues, supporting remediation in a data warehousing and migration project.
- Implemented daily **control charting**, scheduled alerts using conditioned matrices derived from datasets, and synchronized operations for efficient monitoring and management.

Data Scientist Apprenticeship | NSTEM Data Analytics Department |Washington| Jan *2022 - Dec 2022*
Tech Stack: Facebook Analytics, Google Analytics, Google Cloud, Google Looker Studio, Big Query, Miro-REST APIs
- Data engineering of project timelines, budget constraints, and business metrics using machine learning (ML) techniques in Jupyter Notebook, incorporating big data analytics, data mining, and scalability to provide deeper insights and streamline decision-making.
- Developed strategies and analyzed business requirements for web analytics and NLP-powered dashboards, enhancing decision-making with metrics on landing page tracking, customer acquisition, impressions, budget allocation, and conversions by 26%.
- Ensured **compliance**, and data protection with Google's standards and regulatory requirements for data security and data privacy in NLP.

Data Analyst | Anthology (Acquired Blackboard Inc.) | Washington | *Jan 2021 - Dec 2021*
Tech Stack: Salesforce Analytics, Microsoft Excel, SQL
- Employed Salesforce **CRM to conduct a comprehensive** analysis of student enrollment, ticketing, engagement, and retention metrics.
- Implemented **minimax algorithm** with alpha-beta pruning for game tree creation and analysis, enhancing educational games for improved learning outcomes through student behavior insights, and other statistical modeling in collaboration with a team of 5 software developers.
- Improved engagement metrics for the student body of 5.2 million by 12% through targeted outreach campaigns; optimized data management procedures, yielding a 25% decrease in error rates for process improvements.

Reporting Analyst (Contract) | Cox Communications |Atlanta, GA| *Jan 2019 - Jan 2021*
Tech Stack: AWS, Hadoop, Apache Spark, Python, Tableau, Databricks, Jira
- Utilized Python in **Atom** to create automated ETL pipelines, generating daily reports on Elite Gamers segment purchase ratios, and **extracting, error-pruning**, and cleaning data from millions of users.
- Collaborated with a sales team of 20 to increase sales leads by 36% and engaged stakeholders in data collection methods, documentation, coaching and analyzing business strategy for transforming data into system deliverables.

Technical Projects

[Gaussian Naïve Bayes Classifier](#) |Predict income levels exceeding $50,000 annually, achieving a commendable accuracy score. |*May 24*
[Latent Dirichlet Allocation](#) | Analyze data of 350 read links across more than 1,000 documents to identify popular topics. | *April 24*
[Langchain OpenAi](#) | Allows users to create a simplistic chatbot from use cases, simple math, computer vision, coding, tooling. | May 24
[LLM Keras' NLP](#) | Utilized confusion matrix for text-mining to validate AI-generated texts against human inputs, verifying outcomes. | *May 24*