Alfred Prah

Portfolio | LinkedIn | 815-585-6481 | alfredprah@gmail.com

SKILLS

Languages: Python, PySpark, SQL, R, React, Scala, C#, .NET
Databases: Azure, Snowflake, Redshift, PostgreSQL, Oracle
Data Visualization: Tableau, R Shiny, Streamlit
Big Data Processing & Analytics: Databricks, Azure Synapse Analytics, Snowpark
Specialized Expertise: Causal Inference, Natural Language Processing, Deep Learning, Time Series Forecasting, Probability and Statistical Inference, Machine Learning, Big Data Scaling, Exploratory Data Analysis, Data Science Algorithms, Programming & Simulation, Communication

EXPERIENCE

Senior Data Scientist, Launch Consulting Group August 2021 – Present

- Designed and executed a data-driven win-back campaign using Snowflake databases and Snowpark infrastructure, resulting in a significant 62.10% increase in campaign revenue.
- Reduced the cold start problem of a data product by 25% by creating & translating UML diagrams to an Amazon Redshift database schema & populating 3 key tables with simulated data to depict real-world use cases & scenarios.
- Reduced latency by 10% by Refactoring & optimizing 21 PySpark & Scala notebooks to adapt their codebase & Machine Learning algorithms from Databricks to Azure Synapse Analytics for Microsoft's Workplace Analytics platform.
- Improved the performance of 4 recommender systems in production by an average of 60% by shifting from an exclusively collaborative-filter approach through predetermined tier groups to a hybrid version that included a content-based approach.

Co-founder & Chief Technology Officer, CarWiz.ai September 2021 – December 2023

- Leading the creation of CarWiz, a pioneering platform in AI-powered, personalized car selection, I've masterfully blended advanced technologies like React, C#, .NET, and Azure with a deep understanding of user needs, crafting a unique, data-driven car-finding experience that combines technical excellence with intuitive, tailored vehicle recommendations to revolutionize the traditional car selection process.

Data Science Intern, GEODIS January 2021 – May 2021

- Implemented, optimized and deployed 200+ Time Series models to forecast volume and labor demands in the warehouses of 37 different clients/accounts, leading to $100,000+ in savings.
- Created diagnosis reports for 11 "faulty" Time Series models by performing root cause analysis & implementing actionable next steps, leading to a 32% reduction in Root Mean Squared Error (RMSE) across board from the 300+ models in production.
- Developed the Python style guide adopted internally within a Data Science team of experienced hires.

Data Science Research Asst., Vanderbilt Research on Conflict and Collective Action (ROCCA) Lab May 2020 – May 2021

- Predicted civilian sentiment with an 81% accuracy by using text data & PyTorch to design & build a deep learning classifier.
- Created client-facing dashboards using visualization tools like Tableau, R Shiny and Streamlit to explore & highlight dynamic links between civilian-led collective action and conflict trajectories over time, within the United States of America.
- Used text data to design & implement descriptive social networks to highlight 100s of key conflict actors in Sub-Saharan Africa, their respective affiliations, and how their network connectivity has changed over the last 2 decades.

Data Science Intern, OhanaHealth May 2020 – August 2020

- Designed, implemented & deployed a hybrid, deep-learning, recommendation engine (using Transfer Learning) to surface 100s of job openings & their respective descriptions to 1000s of end-users, resulting in a 5% conversion rate for paid subscriptions.
- Performed A/B testing to investigate & strategically advise on user-interactions with varying website layouts & call-to-action buttons.
- Improved click-through rate from targeted ads by 12% by creating & equipping the marketing team with user personas and market segment descriptions, leading to better-informed ad placements & purchases.

Data Science Research Asst., Vanderbilt Data Science Institute February 2020 – May 2020

- Predicted settlement & 54 other outcomes of class-action lawsuits by using Deep Learning and open-source Natural Language Processing techniques to build a hybrid model (RoBERTa + Fast.ai), resulting in an 86% accuracy in predictions.
- Data: ~3,000 legal cases with 18,000 PDFs overall, some of which were not text readable.

Consumer Behavior Researcher, Vanderbilt Owen School of Management May 2018 - December 2018

(Psychographics and Emotional Motivators – for Dono)

- Informed the marketing and go-forward strategy of Dono by designing research methodology, creating surveys, and producing actionable insights from survey findings, resulting in a $11,000+ positive cashflow for a brand that was in debt 3 months prior.

EDUCATION

Master of Science in Data Science, 2021 **Vanderbilt University, Nashville, TN**

Bachelor of Arts in Communication of Science and Technology, 2019 Vanderbilt University, Nashville, TN

PUBLICATIONS

- Understanding Confounding – A critical Insight for Data Scientists
- Navigating the Unknown: The Power of Uncertainty Quantification through the lens of companies like Netflix & Spotify