**Bryce Anthony**

(336) 389-2365 ▪ bkanthon@syr.edu ▪ linkedin.com/in/bryce-anthony ▪ github.com/bryce-ka

**EDUCATION**

**Syracuse University School of Information Studies - Syracuse, NY December 2024**

***Master of Science in Applied Data Science*** *3.8/4.0 GPA*

**Davidson College - Davidson, NC** **May 2023**

***Bachelor of Science in Computer Science***

**EXPERIENCE**

***Data Science Co-op-*** *Keurig Dr Pepper Research and Development*  **Frisco, TX** (*July 2024 - December 2024*)

* Accelerated product development timelines for 30% of future innovations through the application of a machine learning model
* Developed strategic recommendations for data storage to improve workflows and ensure data readiness for AI solutions
* Designed an AI-driven strategy to streamline sensory studies, reducing costs and saving time through process optimization
* Developed proof-of-concept (POC) models using Python, collaborating with KDP R&D teams to address diverse use cases
* Leveraged statistical analysis of historical data to accelerate product development timelines and minimize production waste for future carbonated soft drinks (CSDs).

***Gig-Hub Consultant-*** The Hurt Hub @ Davidson **Davidson, NC** *(August 2020 - December 2022)*

* Collaborated with Davidson alumni, college departments and student organizations, to develop custom software solutions
* Met clients regularly with updating them on project progress and incorporated feedback to deliver high-quality final products
* Past projects include: a niche job board platform for a venture capital firm, a virtual business card platform used during covid, and updating a website to be WCAG & ADA compliant

***Student Researcher*-** Boston University Summer Institute in Biostatistics and Data Science **Remote** (*June 2021- July 2021)*

* Analyzed the NHANES dataset examining the association between cannabis consumption and physical activity in U.S. adults
* Co-authored a paper and presented a PowerPoint of the findings to nearly 40 Boston University students, faculty, and staff
* Conducted statistical analysis using SAS and R, encompassing tasks such as generating descriptive statistics, performing analysis of variance, chi-square tests, and executing multiple linear and logistic regression analyses
* Applied these analytical techniques to datasets sourced from the Framingham and Jackson Heart studies as a significant component of the program
* Recipient of a full scholarship to the NHLBI funded Boston University Summer Institute for Research Education in Biostatistics

***Community Research Fellow****-* Davidson College Center for Civic Engagement **Remote** (*June 2020- August 2020)*

* Authored a comprehensive needs assessment of the North Mecklenburg area using public information and data we obtained through surveys, focus groups, and structured interviews with community members and subject matter experts
* Developed a data dashboard using R and created a library guide to centralize data about the area and create a starting point for future iterations of the program and other future projects
* Utilized R and Canva to create compelling data visualizations that complemented our analyses and highlighted evolving community needs.
* Delivered our recommendations to both organizational leaders and influential community members, aiming to streamline efforts. Our goal was to enhance organizational support for residents and ensure the sustainability of the services being offered

**Academic Projects**

**Text Mining Syracuse, NY** *(August 2024)*

* Implemented logistic regression, support vector machines, and gradient boosting models to classify recipes as healthy or unhealthy with an accuracy of up to 95%, optimizing for precision and recall.
* Cleaned, vectorized, and standardized nutritional and ingredient data scraped from the Spoonacular API, enabling efficient feature extraction for classification models.
* Utilized text processing methods, including stemming, tokenization, and stopword removal, to analyze and categorize recipe text, enhancing model interpretability and performance.
* Conducted PCA and clustering analyses, such as k-means, to visualize feature distributions and identify patterns in dietary classifications, supporting model development.
* Used SHAP analysis to interpret feature importance, identifying macronutrients as the key drivers of model predictions for recipe healthiness.

**Deploying Artificial Intelligence Solutions** **Syracuse, NY** *(Spring 2024)*

* Designed and deployed three machine learning applications (exam performance predictor, NBA game outcome predictor, and NutriScan AI) leveraging AWS services such as EC2, S3, Elastic Beanstalk, and CloudWatch for scalability and real-time monitoring.
* Developed a machine learning model predicting student exam performance using Flask and Python, integrating custom exception handling and logging for enhanced robustness and debugging.
* Created and deployed an NBA match prediction API using TensorFlow, FastAPI, and AWS EC2, focusing on efficient resource utilization, secure access control, and real-time predictions for high-traffic scenarios.
* Engineered NutriScan AI, combining computer vision with natural language processing, deployed via AWS EC2 to provide real-time nutritional guidance powered by Google Gemini Pro and image recognition.
* Utilized AWS IAM, Elastic Beanstalk, and Lambda for seamless deployment and API management, automating resource scaling and provisioning to optimize cost and performance for machine learning applications.

**NBA Game outcome and Spread Predictions Syracuse, NY** *(Spring 2024)*

* Developed a two-output deep learning model to predict NBA game outcomes and point spreads, achieving a test accuracy of 57.47% for game outcomes and 29.59% for point spread predictions on unseen data.
* Designed and implemented advanced feature engineering techniques, including rolling averages for team statistics, player-level aggregation, and principal component analysis (PCA), enhancing model performance and addressing overfitting.
* Conducted extensive hyperparameter tuning and model experimentation, optimizing dropout layers, activation functions (e.g., GELU, ReLU, Tanh), and loss functions (e.g., Huber Loss) to improve model generalization.
* Automated the scraping and preprocessing of over eight seasons of NBA data, incorporating team and player-level statistics into comprehensive datasets for analysis.
* Converted raw data into structured inputs for machine learning models, ensuring high data quality and relevance for predictive tasks.

**Melody Aligned Lyrics Generation Syracuse, NY** *(Fall 2023)*

* Created a SeqGAN network for lyric generation that writes song lyrics to match a given melody using natural language processing techniques
* Developed GAN-based text generation models to produce high-quality song lyrics, leveraging Bidirectional LSTM networks for word and syllable alignment, improving linguistic and rhythmic consistency.
* Engineered preprocessing pipelines to clean and tokenize over 400+ song lyrics, utilizing syllable tokenization and sequence generation for feature extraction and model input preparation.
* Designed and optimized discriminator models for text quality and syllable alignment evaluation, integrating hybrid reward systems to improve model performance using adversarial training.
* Implemented Monte Carlo rollouts and policy gradients to enhance long-term sequence generation, achieving a 50%+ improvement in text coherence and alignment metrics.
* Deployed scalable training loops on GPU-accelerated environments using TensorFlow and Keras, automating data ingestion and model iteration for large-scale text datasets.

**Farm2Fork Syracuse, NY** *(Fall 2023)*

* Created a supply chain logistics application to allow users to track the supply chain of the food they purchase from grocers
* Wrote SQL queries to deliver user data and product data from the database to the back end of our application
* Designed and implemented a relational database in SQL along with its stored procedures, transactions, and triggers
* Enhanced the security of our application by sanitizing input, implementing custom error handling, and other best practices

**Reducing Energy Consumption in South Carolina Syracuse, NY** *(Fall 2023)*

* Formulated six data-driven and actionable recommendations to curtail future energy consumption in South Carolina by leveraging hourly energy usage data of the Energy Company of South Carolina’s customers
* Engineered a sophisticated machine learning model, utilizing two years of energy usage data, to forecast future energy needs
* Designed and implemented a Shiny dashboard in R to complement our analyses and facilitate a comprehensive understanding energy consumption trend

**E-AI** (Emerging Artist Recommendation) **Davidson College Hackathon** *(Spring 2023)*

* Co-developed a prototype for an artist recommendation platform that generate emerging artist recommendations for users,
* Created a flask server to interact with Spotify’s API using authentication info from our front-end Node app and generate emerging artist recommendations to be used by the Node app
* Delivered a 2-minute pitch with to a panel of judges and answered follow-up questions justifying the potential of the project, our project received an honorable mention in the Artificial Intelligence category of the Hackathon

**Predicting Popularity Machine Learning Course** *(Spring 2023)*

* Developed a model to predict the popularity of a reddit post based on its text content and achieved a test accuracy of 77.5%
* Authored a research paper and created a poster for the project which was presented to students, faculty, and staff with varying amounts of knowledge about machine learning
* Used existing literature to identify successful model architectures and identify features important to the success of the model
* Built various classification and regression models for a variety of tasks ranging from wine quality rating prediction to the classification of cancer types based on individual mRNA profiles.

***Machine Learning Course* Davidson, NC** *(Spring 2023)*

* Used Python and SciKit Learn to build machine learning models for a variety of classification and regression tasks ranging from wine quality prediction to cancer type classification.
* Engineered a Machine Learning model to predict the popularity level of a reddit post based on its text content and sentiment achieving an accuracy of 77.5%
* Authored research papers and presented a poster showcasing a the reddit research project at the Verna Miller case symposium
* Used existing literature to identify successful model architectures and features contributing to the success of the models

**Deep Learning Projects Deep Learning Course** *(Fall 2022)*

* Built Deep learning models for a variety of classification and regression tasks
* Created a poster and presented our research to students, faculty, and staff with varying amounts of domain specific knowledge
* Tailored project to address the limitations identified in existing literature and implementing strategies from the same sources

**Multi-Agent Reinforcement Learning in Atari Joust Machine Reasoning Course** *(Spring 2023)*

* Developed and trained a bot to cooperatively play Atari joust using the research methodology outlined by Klijn and Eiben
* Implemented a frame stacking algorithm and genetic programming algorithm for updating the weights of the models in training
* Wrote a research paper and delivered our findings to a diverse audience through an oral presentation and poster display

**Awards & Honors**

* **Ben Callinder Award:** Presented annually to a member of the senior class who selflessly gives their time to others, unites people from different walks of campus, pursues their passions and is respected by their peers, faculty, and the administration
* **Gladstone Memorial Award:** presented to 2 rising seniors exhibiting high potential for service to mankind as indicated by leadership, service and academic record.