BUA MATTHEWS

DATA SCIENTIST

CONTACT

949-237-1391

buavanit@gmail.com

github.com/bvanitsthian/photocircle_nlp

linkedin.com/in/buamatthews/

Brentwood, TN

SKILLS

Python	Spark
SQL	Power BI
R	Data Mining
Web Scraping	NLP
ML	Git

EDUCATION

Doctor of Computer Science

Colorado Technical University

2022 - present

MS Applied Economics

University of San Francisco

2015 - 2017

BA Quantitative Economics

University of California, Irvine

2011 - 2015

LANGUAGES

English	Fluent
Mandarin	Native
Thai	Native

PROFILE

I am a Blockchain enthusiast and Data Scientist with a passion for forecasting events amid uncertainty. With a background in quantitative research, I'm driven to tackle complex problems and deliver novel insights to stakeholders. I pride myself on being a detail-oriented problem solver and a team player, with technical skills that include data wrangling, predictive modeling, and data visualization using advanced tools such as Python, R, and Tableau. When I'm not working on data projects, I enjoy bodybuilding, fueling my passion for discipline, focus, and hard work. I'm excited to bring my unique perspective and creativity to every project, combining intelligence and innovation to deliver exceptional results.

WORK EXPERIENCE

Data Science Apprentice

Nashville Software School — August 2022 - May 2023

Intensive part-time bootcamp focusing on data science fundamentals and problem solving. Used real-world datasets and included projects where findings were presented to stakeholders from the community.

- Conducted exploratory data analysis using Python's pandas library and R's tidyverse packages, uncovering insights that led to process improvements and cost savings
- Created impactful data visualizations using matplotlib, seaborn, and ggplot2, enabling stakeholders to quickly understand complex data and make informed decisions
- Performed geospatial analysis using geopandas and folium, enabling businesses to understand location-based patterns and trends
- Gathered data through APIs and webscraping, providing businesses with access to valuable data sources that were previously inaccessible
- Retrieved and analyzed data using PostgreSQL and sqlalchemy, building robust databases and conducting analyses that led to actionable insights and improved decision-making
- Built and evaluated statistical and machine learning models using the scikit-learn and statsmodels libraries, resulting in accurate predictions and improved performance
- Applied natural language processing using the nltk and spaCy libraries, enabling businesses to extract insights from text-based data sources such as social media and customer feedback
- Performed network analysis on graph data using Neo4j, enabling businesses to understand complex relationships and networks
- Built and deployed interactive data visualizations using the R Shiny library, providing stakeholders with easy-to-use dashboards that helped them make informed decisions
- Utilized source code version control with Git/GitHub and project management/tracking tools such as GitHub project boards and issue tracking, ensuring efficient collaboration and tracking of progress
- Interacted with AWS using the CLI and ssh, building scalable and efficient cloud-based data solutions

Virtual Restaurant Founder

Chicken as Cluck — March 2019 - December 2021
(A cloud kitchen based in San Francisco & Austin)

- Utilized time series analysis and clustering techniques to increase revenue by 250% in 2021 by optimizing menu offerings and delivery times based on customer behavior
- Developed regression and decision tree-based pricing models to increase brand awareness, customer acquisition, and profitability
- Launched new services such as catering and private events, resulting in increased revenue streams, using market basket analysis and collaborative filtering to market and promote these services
- Created predictive pricing models using linear regression and random forest algorithms to optimize prices based on demand and market trends, resulting in increased profitability
- Improved operational efficiency and reduced costs through the use of process automation and optimization algorithms
- Conducted A/B testing and multivariate analysis to optimize website design and user experience, resulting in increased website traffic and online sales.
- Leveraged natural language processing and sentiment analysis to monitor online reviews and social media mentions, enabling timely response to customer feedback and improving customer satisfaction

DATA SCIENTIST

MEDIA PUBLICATIONS AND ACHIEVEMENTS

2018 PNBA Olympian Gold Medalist

Eater Austin

Chicken as Cluck X Zira

Austin Food Magazine

SF Chronicle

PROJECT HIGHLIGHTS

Aggravated Burglary in Davidson County

Pulled data from the Census API and Nashville's REST API to predict Nashville burglaries. Presented with ipywidets, GeoPandas, and Folium maps.
(**Python, GeoPandas, Folium maps)**

NLP-based User Segmentation

This project segments users based on album titles using NLP, resulting in improved personalized recommendations and increased user engagement.
(**Python, NLP, spaCy)**

TN Mortgage Probability Calculator

Investigated if race is a stronger predictor of mortgage loan denial than income and associate indicators of creditworthiness in TN using logistic regression.
(R, ShinyApp, Tidyverse, ggplot2, stats)

TN Healthcare Referral Patterns

Using an unsupervised machine learning algorithm and community detection, generated recommendations for which professionals Vanderbilt Hospital should contact to increase patient volume.
(Python, sqlite3, neo4j)

WORK EXPERIENCE

Forensic Economist

HSNO — August 2017 - May 2018

- Utilized statistical methods such as regression analysis, time series analysis, factor analysis, Monte Carlo simulation, ARIMA, and VAR models to investigate and solve complex economic issues related to cybersecurity hacks and fraud-related claims and natural disasters
- Conducted data aggregation, outlier detection, and built regression models in Python to project past damages for businesses affected by natural disasters
- Performed economic analyses relating to market power in antitrust disputes and employment discrimination, utilizing techniques such as hypothesis testing, cluster analysis, factor analysis, game theory, and econometric modeling
- Developed and maintained a customized python script that pulls data from the Bureau of Labor Statistics API, enabling streamlined and efficient access to key data sources
- Provided expert witness testimony and consulting services in legal disputes, including deposition and trial testimony, report writing, and client communication, utilizing excellent verbal and written communication skills

Business Intelligence Analyst

Rodan + Fields — July 2016 - July 2017

- Maintained and optimized SQL (MS SQL) and R scripts to populate tables and generate reports in data warehouse for daily, weekly, and monthly reporting across departments
- Partnered with the internal strategic insights team to identify key business questions and translate them into structured analysis
- Synthesized complex datasets to provide analysis, insights, and recommendations for optimizing sales strategies and tactics
- Produced regularly scheduled reports on purchase behavior and sales performance for R+F LASH Boost product, using Power BI and Tableau to communicate findings to stakeholders in a clear and actionable manner
- Collaborated with cross-functional teams to develop and maintain data governance processes and data quality standards, ensuring accurate and reliable data for reporting and analysis

Research Analyst

Panoramic Interests — October 2015 - June 2016

- Conducted ongoing monitoring and analysis of commercial and residential real estate market conditions, utilizing tools such as CoStar, REIS, and Real Capital Analytics to track trends and identify opportunities for investment
- Prepared comprehensive feasibility studies and due diligence reports for all new real estate acquisitions, leveraging data and analysis to construct detailed investment pro-formas and identify key risks and opportunities
- Collaborated closely with cross-functional teams, including Development, Acquisitions, and Asset Management, to gather and analyze data on potential investment opportunities and ensure alignment with company goals and strategies
- Facilitated accurate, timely, and insightful monthly and periodic variance analysis for executive management, providing meaningful insights into key performance indicators and trends in the portfolio

Undergraduate Research Assistant

University of California, Irvine — August 2013 - May 2015

- Conducted research and analysis on adverse economic trends posing risks to the labor market, utilizing industry reports and data sources to inform insights and recommendations
- Analyzed studies on derivative markets and gathered up-to-date data from the International Swaps and Derivatives Association (ISDA), ensuring accuracy and relevance in research findings
- Imported and prepared wage data, consolidated code, and utilized benchmarking tools to reduce processing time in statistical software packages such as SAS and Stata, increasing efficiency and productivity
- Compiled, manipulated, and reviewed large datasets from various government and market sources, facilitating the production of tables and charts to support research findings and presentation materials