Alex Biuckians

Rockville, MD | alexbiuckians@gmail.com | 443-852-6283 | linkedin.com/in/alex-biuckians | github.com/alexbiuckians

SUMMARY

M.S. Data Science student who builds machine-learning systems end to end on real, messy data — from leakage-safe modeling through deployed, CI-backed services — and defends every result honestly against a baseline. Experienced across spatial, causal, time-series, operations-research, and anomaly-detection problems, reporting honest skill scores rather than inflated headline numbers.

EDUCATION

M.S., Data Science (GPA 3.72) Expected May 2027

George Washington University, Washington, DC

- *Relevant Coursework*: Data Warehousing, Data Mining, Algorithm Design for Data Science, Machine Learning I: Algorithm Analysis, Visualization of Complex Data, NLP for Data Science, Deep Learning and the Fundamentals of AI, Time Series Analysis & Modeling

B.A., Sport Management: Minor, Data Analytics (GPA 3.55) Dec 2024

High Point University, High Point, NC

- *Relevant Coursework:* Applied Multivariate Statistics, Data Analysis & Predictive Analysis, Intermediate Statistical Analysis, Introduction to SAS Programming, Introduction to Statistical Analysis, Linear Algebra

TECHNICAL SKILLS

- *Languages*: Python, SQL, R, SAS, C++, JavaScript
- *ML & Data Science*: pandas, NumPy, SciPy, scikit-learn, LightGBM, XGBoost, PyTorch, tidymodels, SHAP, DALEX, causalml, econml, Optuna
- *MLOps & Serving:* FastAPI, Pydantic, Docker, MLflow, GitHub Actions (CI), Evidently, pytest, ruff, Render, REST APIs
- *Data & Geospatial:* PostgreSQL, SQLite, GeoPandas, sf, NetworkX, Kafka/Redpanda, FAISS
- *Visualization:* Streamlit, Shiny/Leaflet, Plotly, Dash, Power BI, Tableau, Matplotlib, seaborn, ggplot2
- *Modeling & Methods:* Regression, Classification, Causal Inference / Uplift, Spatial Statistics (Moran's I, LISA), Time-Series Forecasting & CV, Anomaly Detection, Sequential Testing (SPRT/CUSUM), Operations Research, A/B Testing, Model Evaluation

PROJECT EXPERIENCE

TurnoutLens: Spatial Analysis of U.S. Voter Turnout

R, tidymodels, PostgreSQL, DALEX, spdep, Shiny/Leaflet | 2026

- Modeled U.S. county voter turnout across seven presidential elections (21,789 county-cycle observations from MIT Election Lab and U.S. Census data); demographics alone explained 61% of out-of-sample variance, trained on 2000–2020 and tested on an unseen 2024.
- Found the signal in the residuals: turnout gaps cluster spatially at Moran's I = 0.391 (p = 2.2e-290), and LISA clusters identify the same-day-registration states — recovering the effect of election law without being given it. Deliberately dropped a lag feature that improved accuracy 19.5% because it masked the demographic signal.

RouteCast: Last-Mile Delivery Dispatch Intelligence

Python, LightGBM, SciPy (Hungarian), SimPy | 2026

- Built an operations-research and ML system on 445,295 real deliveries (Alibaba Cainiao LaDe); three LightGBM quantile models predict delivery time at P10/P50/P90 with calibrated coverage (P90 covers 88.9% vs a 90% target), leakage-free by construction.
- Fed predictions into a Hungarian assignment optimizer and a SimPy discrete-event staffing simulation; showed the optimizer's edge grows with dispatch radius, but staffing the peak hotspot cuts average order time from 87 to 55 minutes — identifying staffing, not assignment, as the larger lever.

DriftWatch: Production ML Monitoring

Python, scikit-learn, Evidently, FastAPI, Streamlit, GitHub Actions | 2026

- Built a production monitoring system for a deployed credit-default model that detects data drift (KS/PSI), performance decay against late-arriving labels, and latency against an SLO, with hysteresis-gated alerting, a runbook, 15 tests, and CI.
- Demonstrated the failure mode drift-only monitoring misses: under concept drift, F2 fell 0.640 → 0.565 and recall collapsed 0.88 → 0.62 while data drift read only 4%. Built on the real UCI credit dataset (30,000 cardholders, honest held-out AUC ≈ 0.78).

Parquet Capital: NBA Contract Valuation & Forecasting Engine
Python, scikit-learn, PuLP, Streamlit | 2026

- Forecast NBA player performance three seasons out with quantile uncertainty bands calibrated to ~80% empirical coverage, priced contracts against market comps, and optimized roster construction under the cap with linear programming; built on 4,860 real player-seasons with 45 tests.
- Ran a decision-level backtest on held-out seasons showing Overvalued flags cost ~8.6× more per delivered value point than Undervalued ones, with robustness confirmed against an independent target (VORP) and an independent model class.

PROFESSIONAL EXPERIENCE

Analytics Intern, Division I Men's Baseball Team Jan – Apr 2024
High Point University – High Point, NC

- Analyzed Trackman pitch and batted-ball data for 30+ players to support player evaluation and performance review
- Evaluated team and opponent tendencies using Synergy and scouting workflows to support game preparation

Division I Men's Basketball Team Manager Aug 2022 – Nov 2023
High Point University – High Point, NC

- Conducted film analysis using HD Intelligence to identify high-efficiency shot zones

Film & Analytics Cohort Aug – Sep 2023
Grow the Game, Hudl – High Point, NC

- Generated box scores and shot charts for 40+ basketball players and created five solo film segments using SportsCode (custom coding language)

Game Day Intern May – Jul 2023
Winston-Salem Dash – Winston-Salem, NC

- Tracked live gameplay using DakStats for 20+ games