THOMAS FELTON

Gainesville, VA | 571-398-4461 | Feltontd5@gmail.com | https://github.com/TDFelton | linkedin.com/in/tdfelton

EDUCATION

William & Mary, Williamsburg, VA — Expected Graduation: May 2027

B.S. in Data Science with a concentration in Artificial Intelligence | Minor: Chemistry | GPA: 3.79

Honors: Dean's List, 6 semesters

RELEVANT COURSEWORK

Applied Machine Learning | Linear Algebra | Applied Linear Algebra & Calc | Prob & Stats for Scientists | Intro to Data Science | Intro Multivariable Calculus | Generative AI | Supercomputing for Science

RELEVANT EXPERIENCE

Laboratory Assistant, Meldrum Physical Chemistry Lab, William & Mary — Jan 2024 - Present | 400 hours

- Developed an end-to-end MD simulation pipeline using OpenMM and MDAnalysis to compute and validate T_2 relaxation times, managing 1,000+ SLURM job submissions on W&M's HPC via MPI and Bash orchestration
- Engineered a validation suite comparing simulated vs. experimental T_2 values and automated publication-quality figures for diffusion and T_2 profiles using Matplotlib and SciPy

Student Manager/Offensive Analytics and Advanced Scouting, William & Mary Baseball — Mar 2024 – Present

- Built a semi-automated scouting pipeline in Python, scraping and aggregating opposing pitcher data from Synergy and Trackman to eliminate manual data entry
- Produced integrated scouting reports — combining Google Sheets analytics and video — for 10–20 opponents per season, delivered to both coaching staff and players ahead of each series

LEADERSHIP EXPERIENCE

Pitching Coach, Club Baseball, William & Mary — Aug 2024 – Present

- Provide individualized mechanical and pitch development coaching while directing in-game strategy based on batter-pitcher matchups and high-leverage situations
- Design structured practice programs driving team-wide development; recognized as NCBA 1st Team All-Region Pitcher, Mid-Atlantic North

PROJECTS

MLB Pitch Quality (Stuff+) Evaluation System | Python, PyTorch, LightGBM, SLURM

- Independently replicated and extended the FanGraphs Stuff+ methodology across two end-to-end pipelines trained on 700K+ Statcast pitches (2015–2025): a cascaded ensemble (Random Forest → contact classifier → LightGBM regressor) and a PyTorch MLP (256→128→64, GELU, BatchNorm) isolating pure stuff quality from 12 physical features with no location confounders
- Applied per-pitch-type normalization (avg = 100 per type), GroupKFold cross-validation grouped by pitcher to prevent leakage, and a gradient-based stability regularizer penalizing year-over-year inconsistency — validated via Spearman correlation and quartile persistence metrics
- Deployed on W&M's HPC cluster (Sciclone) via SLURM with automated multi-year data ingestion, model serialization, and leaderboard generation

SKILLS

Programming: Python, R, SQL, Bash, LaTeX

ML & Data Science: PyTorch, Scikit-learn, LightGBM, CatBoost, Statsmodels, SciPy, Pandas, NumPy, Matplotlib, Seaborn

Simulation & HPC: OpenMM, MDAnalysis, SLURM, MPI, TCSH

Data Engineering: JSON, CSV, NumPy binary arrays (.npz), molecular simulation formats (.dcd, .pdb, .mol2)

Tools & Platforms: Git, Jupyter, PyBaseball, Google Sheets, Excel Cinema4D