

Tjay Burger

Junior Machine Learning Engineer

Potchefstroom, North West (relocating to the Western Cape) · tjayburger2004@gmail.com · +27 71 439 0187 · Driver's licence and own car

[LinkedIn](#) · [GitHub](#)

I build machine learning systems end to end: the statistics, the model, the API and the deployed product. Computer Science and Statistics student at North-West University, group leader on every team project, and third at the South African wrestling championships. Looking for junior ML engineer roles in Stellenbosch, Cape Town or remote.

PROJECTS

Causal Credit Uplift · Python, EconML · [live dashboard](#)

- Compared T-learner and CausalForestDML uplift models to find which customers a treatment actually changes, first on the real Hillstrom email RCT, then on synthetic credit data with known true effects.
- CausalForestDML recovered the true effects with 0.927 correlation and kept customers with negative effects out of its top decile (0.00%).

Credit Scorecard Engine · Python, FastAPI, React · [live demo](#) · [GitHub](#)

- Built two credit scorecards (bureau-based and non-bureau) on 307,511 real applications from the Home Credit Default Risk dataset. WOE/IV binning and points scaling are implemented from scratch, with no scorecard libraries; logistic regression tuned with 5-fold cross-validation.
- Holdout AUC 0.737 (Gini 0.473, KS 0.355) for the bureau model and 0.631 for the non-bureau model, benchmarked against XGBoost at 0.753 and explained with SHAP.
- Found an "inclusion gap" of 14.7% of applicants who the bureau model rejects or refers but the non-bureau model accepts. They default at 12.1%, against 3.6% for agreed accepts and 20.8% for agreed rejects.

AI Options Trading Agent (hackathon, paper trading) · Python, alpaca-py, DigitalOcean · [GitHub](#)

- Built an autonomous agent where a volatility signal (implied vs 20-day realised vol) limits what an LLM may propose, and an independent risk gate re-prices every spread from live quotes and can veto it (2% of equity per trade, 3 open positions).
- Two logged failures (a spread that was not a real credit, and a wrong-direction trade) led to a hard direction constraint enforced in code.
- Runs unattended on a DigitalOcean server under systemd. A proxy backtest over about 2.1 years on six tickers showed volatility mean-reverting in 67.9% of 53 sell-side signals; the README states this is not proof of profitable trading.

EDUCATION

BSc Computer Science and Statistics, North-West University, Potchefstroom · final exam 16 November 2026

- STTN327 group project (five people, statistics): predicted sulphur dioxide levels across 41 US cities. Group leader; built the mixed-effects and random forest models and presented the model comparison.
- CMPG321 group research project: wrote the SQL behind a ride-hailing study of trip cancellations, safety metrics and fleet quality.

EXPERIENCE AND LEADERSHIP

- **Payroll spreadsheets, family firm:** build the monthly UIF and admin payroll workbooks in Excel from the monthly payroll files.
- **Class representative:** for first-year modules and again for a module this year.
- **Group leader:** led the team on university projects, including the five-person STTN327 project.
- **Charity events:** take part in charity runs, including the Hollard Balls Run for cancer.

SKILLS

- **Languages:** Python, R and SAS (strongest); SQL, TypeScript
- **ML and statistics:** scikit-learn, XGBoost, SHAP, EconML, logistic regression and WOE scorecards, mixed-effects models, random forests, causal inference and uplift modelling
- **Engineering:** FastAPI, Streamlit, React, Next.js, Git, Vercel, Render, DigitalOcean with systemd; Cursor and Claude for AI-assisted development
- **Next to learn:** cloud platforms, Java, C++

BEYOND THE CODE

Third place at the South African wrestling championships (Olympic-style wrestling). Played rugby, train daily, and am preparing for a Hyrox race. Three Kaggle competitions entered.