

Tjay Burger

Junior Quantitative Machine Learning Analyst

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 models for decisions under uncertainty: credit scorecards written from scratch, causal uplift models and a volatility-based trading agent, each tested against its own failure cases. Computer Science and Statistics student at North-West University and group leader on every team project. Looking for junior quantitative ML and credit risk roles in Stellenbosch, Cape Town or remote.

PROJECTS

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

- Built an options agent around the volatility risk premium: it sells premium when implied volatility is more than 2 points above 20-day realised volatility, buys when it is cheaper, and stands aside in between. An LLM only chooses strikes inside that signal.
- An independent risk gate prices spreads at worst-case fills (sell at bid, buy at ask), rejects non-positive credits and caps loss at 2% of equity per trade.
- Proxy backtest on six tickers over about 2.1 years: volatility mean-reverted in 67.9% of 53 sell-side signals and 58.0% of 69 buy-side signals. No historical options P&L is claimed, and the agent trades on a paper account only.

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

- Built bureau and non-bureau scorecards on 307,511 applications with WOE/IV binning, a 300-850 points scale (PDO 20, base 600) and accept/refer/reject cutoffs set from score percentiles, all implemented from scratch.
- Holdout AUC 0.737, Gini 0.473, KS 0.355 for the bureau model. An XGBoost benchmark reached 0.753 with SHAP, but the scorecard stays the decision model because regulators require explainable decisions.
- Found a 14.7% "inclusion gap" of applicants 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.

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

- Uplift models (T-learner, CausalForestDML) for deciding who to treat, such as who gets a credit-limit increase. Tested on a real 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%).

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
- **Statistics and ML:** logistic regression, WOE/IV scorecards, mixed-effects models, random forests, XGBoost with SHAP, causal inference and uplift modelling (EconML)
- **Quant tooling:** alpaca-py, NumPy, pandas, scikit-learn; volatility and backtesting code in Python
- **Engineering:** FastAPI, Streamlit, Git, 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.