

NATHANIEL FORDE | RESUME

Senior Data Scientist

📍 55 Bannow Road, Cabra D7, Dublin, Ireland
🏠 nathaniel.github.io

☎ +353 86 052 5426
✉ nathaniel.forde@gmail.com

Short Profile

- Experienced Data Scientist specialising in probabilistic modelling of risk and causal inference.
- Built, deployed and monitored ML solutions on AWS cloud and bare-metal stacks.
- Open Source Contributor @ PyMC
- Academic research background in mathematical logic and philosophy

Experience

Staff Data Scientist - Personio- Dublin

April-2024-Present

- Growing the data science function from 0 to 7 data scientists. Providing cross functional statistical consultation across the business guiding product development and revenue strategy.

Sample Projects: - Lead DS on Personio's Proactive Insights and Engagement Survey customer facing analytical products. Developed in-house experimentation programme to demonstrate the value of A/B testing and justify the purchase of Statsig for scaling experimentation in Personio. Forecasting revenue and costs to aid planning and resource allocation. Mentoring Junior colleagues and establishing best practices and ways of working within the team

Senior Data Scientist - Personio- Dublin

Nov 2021-April 2024

- Working within the Product and Engineering team to measure risk and optimise costs (broadly construed) in the software development life-cycle. Leading projects from design, data collection to modelling and model maintenance.

Sample projects: - Root cause analysis of lagging latency response times, Estimation of UX pain through metrics trees and factor analysis. Survival Modelling (Time to Failure) studies to prioritize remediation work and improve customer experience. Developing and experimenting with NLP and LLM solutions for customer feedback aggregation and diagnostics. Designed and built Personio's first ML deployment pipeline using a Python AWS tech stack

Data Scientist (R & D) - Cartrawler- Dublin

Nov 2019-2021

- Building and maintaining models to move markets. Focused on cost reduction by accurate demand forecasting and insurance cost prediction with explainable ML and probabilistic models.

Sample Projects: - Built customer segmentation based on price appetite to help optimise dynamic pricing models. Developed an interlocking claim propensity model and loss-ratio development forecast to help manage the uncertainty in the insurance product. Tech Stack: Python, PYMC, scikit-learn and in-house ML-pipeline packages

Senior Insights Analyst - Paddy Power/Betfair- Dublin

Feb 2018-2019

- Working within the Customer Operations Insights team using data science techniques to better understand the needs of the customer, in areas of Fraud, Responsible Gambling, CS, CX, Payments.

Sample Projects: - Developed NLP tool for a sentiment prediction on gaming industry customer service chat-data. Worked with industry think tank to refine and augment external and internal model for predicting problem gambling. Deployed propensity modelling projects to evaluate causal effects of interventions with customers e.g. the ROI for VIP relationship management.

Experience

Software Developer - Marsh & McLennan- Dublin

Mar 2017-2018

- Worked in Guy Carpenter's Catastrophic Risk research team developing Web-applications to serve up and automate the reporting of the team's Risk prediction models to aid the pricing of client portfolios against the risk of earthquakes, fires and floods. Technical Stack: Go, Postgres for the back-end. Javascript (React and D3) for the front end

Data Analyst - Marsh & McLennan- Dublin

Mar 2015-2017

- Worked within a team of four to design and implement the Data Quality function for Marsh globally.
Sample Projects: Developed statistical metrics and classification methods to assess the health of numerous data-sets including placement, sales, billing, and reference data. Built various reporting tools and analytic dashboards (Qlikview, Tableau) to motivate increased data-quality (GIGO). Co-ordinated with international teams to monitor and assess the process and progress of Data-Quality initiatives. Conducted bespoke analyses at request from senior stakeholders. Built a CRUD app with Shiny for the entry and management of data on customer calls for quality assurance

Selected Publications and Open Source Work

Bayesian Vector AutoRegressive models of Irish GDP - PyMC- Open Source

Dec 2022

- VAR and Hierarchical VARs for GDP models using PyMC. Publication https://www.pymc.io/projects/examples/en/latest/time_series/bayesian_var_model.html

Multilevel Regression and Post-Stratification in Forecasting - BamBi- Open Source

Sep 2023

- Demonstrated the usage of weighting techniques in the analysis of survey data for election forecasting. Publication: https://bambinos.github.io/bambi/notebooks/mister_p.html

Justifying Instruments in IV Designs - CausalPy- Open Source

Jun 2024

- Highlighted the manner of justification and role of comparative assessment in credible causal inference. Publication: https://causalpy.readthedocs.io/en/latest/notebooks/iv_weak_instruments.html

Education

Msc Logic - University of Amsterdam: Institute of Logic, Language and Computation- Amsterdam, Netherlands

2011-13

- Research thesis written on models of causal inference and other dependency relations. Focus in Linguistics and Logic with a focus on NLP and Formal Semantics.

MPhil (Research) Philosophy - Trinity College Dublin- Dublin, Ireland

2009-11

- Research on Modal Logics and identity. Awarded an IRCHSS scholarship for this work.

BA International - UCD/Sorbonne Université- Dublin Ireland & Paris, France

2005-09

- Various courses in philosophy and literature. Spent a year in Paris through the Erasmus Programme