Quentin-Gabriel Thurier

New Zealand | quentin.thurier@gmail.com |

I thrive in cross functional teams working at the interface between industry and academic research in order to craft and develop great Machine Learning based products.

Skills

Languages	Python, Bash, SQL
Technologies	TensorFlow, scikit-learn, spaCy, Snowflake, Spark, Hadoop, AWS, Google Cloud Platform

Work Experience

Senior ML Engineer **Decoded Health** *June 2023 – Present*

- Rolled-out a fully automated medical intake process by implementing few-shot learning strategy for conversation summarisation, saving human in the loop contractors budget.
- Improved clinical entity extraction F1-score by 25% by using promp engineering strategies with Azure OpenAI.
- Established standards for data versioning and machine learning services observability on a Google Cloud Platform.

Senior Applied Scientist **Xero** *November 2020 – May 2023*

- Reduced the time spent on manual data entry by bookeepers by rolling-out a machine learning service on an AWS stack. It is used by 3M+ users and responds to 4M+ requests per day.
- Led the deep learning research (50+ models experiments) on bank reconciliation recommendations using TensorFlow.
- Simplified the data pipeline (PySpark) by halving the codebase and driving its migration to Snowflake and Prefect.
- Streamlined and aligned models documentation across 4 cross-functional pods by integrating Jira & Confluence.
- Patented 3 inventions in collaboration with the intellectual property team.
- Supervised 1 applied scientist and 3 graduates among which 2 became ML engineers.

NLP Engineer **PredictHQ** *May 2019 – November 2020*

- Reduced the overall records duplication rate to 3% by rolling-out a random forest classifier for event entity resolution with scikit-learn, giving the company a competitive edge in the events data providers space.
- Increased by a factor of 10 the level of granularity of events categorisation by rolling-out an ontology-based classifier using spaCy, improving downstream machine learning models performance.
- Led a 3-member team to deliver a gradient boosting regression model for concert events attendance prediction that decreased the mean absolute error down to 1000 attendees.

Data Scientist **Orion Health** *January 2017 – April 2019*

- Rolled-out an ensemble classifier (scikit-Learn) to predict post stroke outcomes into Auckland North Shore Hospital, achieving an area under the ROC curve of 94% on inpatient death outcomes.
- Managed a $300,000+ research budget and led a cross-functional 8-member team to deliver 2 interpretable machine learning models and 2 peer-reviewed publications.
- Started off 3 product features by delivering deep learning prototypes with TensorFlow and Keras.
- Enabled deep learning research in the team by building a GPU workstation with a $8,000 budget.
- Presented and shared my work at 4 venues, including the Ministry of Health.
- Mentored 3 data scientists and supervised 4 interns among which one was hired.

Data Scientist **Qrious** *2016*

- Delivered geospatial insights using mobile phone activity and public Wi-Fi data with Python, Hadoop & Spark.

Data Manager **NetBooster** *2015*

- Implemented data-driven digital marketing solutions leveraging web server logs on a GCP stack with Python.

Statistician **Société Générale** *2010 – 2013*

- Rolled-out a network analysis software to prevent unauthorized trading in response to a significant financial loss.

Education

Post Master's Degree (ML)	Telecom ParisTech	*2014*
Master's Degree (Statistics)	ENSAI	*2005 – 2010*

Patents

Method, Program, and Apparatus for Processing Sensitive Data	May 2023
Methods and Systems for Training Attribute Prediction Models	March 2023
Transaction Data Processing Systems and Methods	June 2022

Research

Physician understanding, explainability, and trust in a hypothetical ML risk calculator	JAMIA 2019
Smart MedRec: Using ML for reading dose instructions and incorporating this in clinical software	HINZ 2019
Inspecting a ML-based clinical risk calculator: a practical perspective	IEEE CBMS 2019
Interpretable ML for healthcare	HINZ 2018
Improving clinical named entity recognition with transfer learning	HIC 2018
HOPE (Health Outcomes Prediction Engine) for stroke	HINZ 2017
New Zealand health data review	HINZ 2017

Interests

Judo (second degree black belt, New Zealand nationals medalist, coach & former club secretary), Surfing, French Cinema