Will Radford / Data Scientist

http://wejradford.github.io / https://twitter.com/wejradford / wejradford@gmail.com

I'm a data scientist with over ten years experience in building systems to solve interesting problems using machine learning, particularly those involving natural language processing.

EXPERIENCE

Red Marker, Sydney - *Data Scientist*

April 2017 - now

- Working with finance and legal experts to model domain problems, training a range of predictive models (rule-based, linear-models, neural).
- Architecting model deployment on different AWS environments (EC2, Lambda), monitor and scale them to production loads.
- Designing data-ops workflows to make the most of our data, design evaluation programmes, and integrate active learning to maximise annotator ROI.
- In-the-weeds detailed information extraction and coreference resolution with long, complex financial documents.
- Wholesale re-engineering of an NLP/ML stack to improve scalability and deployability.

Hugo.ai, Sydney - *Lead Data Scientist / Engineer*

January 2016 - March 2017

- Used machine learning to build a collective intelligence CMS that accelerates a human web search task, cutting median job time from 60 to 6 minutes in 12 months.
- Built system to identify at-risk posts in online youth mental health forums in an international shared task for Reach Out Australia.
- Investigated how humans use emoji using a translation of *Moby Dick* into emoji.
- Co-authored research into generating one-sentence biographies from structured knowledge-base records.
- Actively participated in the research community, reviewing and publishing papers and co-organising the Humans, Language, Technology Meetup, with over 200 members.
- Led a small team of data scientists, co-supervising an Honours student.

Xerox Research Centre Europe, France - *Postdoctoral Researcher*

June 2014 - December 2015

- Built natural language understanding systems for chatbots for telecommunications customer care.
- Participated in client events promoting and explaining chatbot technologies and workshopping engagements in the recruitment and online retail space.
- Worked on techniques for using cast-list information to study gender representation on film and television, resulting in published papers and New York Times coverage.
- Researched methods for improving named entity recognition with human assistance and structured data from a knowledge base, writing paper and submitting a patent.

- Researched improved systems for distinguishing similar languages in Tweets (e.g. Spanish and Catalan) using social networks.

Capital Markets Cooperative Research Centre, Sydney - *Researcher*

June 2008 - June 2014

- Completed a PhD project on linking names in text to Wikipedia knowledge base, building the system and evaluating techniques.
- Collaborated with Fairfax Media engineers to deploy Zoom timeline exploration system based on PhD project to paywall customers.
- Taught programming to advanced-unit undergraduates, delivered guest software engineering lecture, and tutored at the National Computer Science Summer School.
- Built system for tracking information flow from ASX company announcements to the Reuters news stories that covered them, distinguishing different types of flow.

Appen, Sydney - *Software engineer*

August 2004 - June 2008

- Worked on projects supporting collection, analysis and delivery of language data including speech and tablet-based writing/drawing data.
- Built part-of-speech tagger for Thai to support a data collection project.
- Worked on project to predict an author's demographic and psychometric traits from their email messages.

TECHNOLOGIES

Python, Jupyter, scikit-learn, *some* tensorflow/keras, AWS, bash, Spark, HTML/CSS, Javascript, SQL, Linux, git, vim, *rusty* Java/C/C++.

SKILLS

Working in agile software development teams using tools like Jira, Bitbucket, GitHub. Writing for and presenting to non-academic and academic audiences.

EDUCATION

University of Sydney - *PhD in Computer Science (2010-2013)*

Thesis: Linking Named Entities to Wikipedia / APA and CMCRC scholarships / Tutoring award

University of Sydney - *B. Liberal Studies 1st class Honours (2009)*

Thesis: Tracking information flow in finance text / CMCRC scholarship

University of Sydney - *B. Liberal Studies in Computer Science, Linguistics (2000-2003)*

PUBLICATIONS

Evaluating entity linking with Wikipedia - Artificial Intelligence (2013)

B Hachey / **W Radford** / J Nothman / M Honnibal / JR Curran

Learning multilingual named entity recognition from Wikipedia - Artificial Intelligence (2013)
B Hachey / **W Radford** / J Nothman / M Honnibal / JR Curran

Tracking onscreen gender and role bias over time - Web Science (2016)
W Radford / M Gallé

A full list is available on http://wejradford.github.io

OTHER

Australian citizen / British citizen / Useable French