

ANCIL CRAYTON

www.ancilcrayton.com ◊ www.linkedin.com/in/ancilcrayton/ ◊ ancil_crayton@hotmail.com

RELEVANT EXPERIENCE

Data Scientist, Booz Allen Hamilton

Sep 2019 - Present

Research and demo use cases of state-of-the-art deep learning, language models for text generation, reading comprehension, and QA. Building automated backend cloud pipelines for COVID-19 modeling, implementing data pull, data preprocessing, and modeling (SIR, IHME, etc.) using AWS Lambda, S3, CloudWatch Events, EC2, Systems Manager, and bash scripting. Researching causal ML methods to assess the impact of COVID-19. Developed novel BERT model trained from scratch to predict patient readmission for Centers for Medicare & Medicaid Services (CMS). Developed demo for generating realistic Booz Allen press releases and performing QA by fine-tuning an XL OpenAI GPT-2 model (1.5B parameters) with multiple GPUs. Co-presented text generation and reading comprehension demonstrations of transformer-based models at NVIDIA GTC 2019 (Washington, D.C.).

Project Manager, Team Lead; Data Science for Social Good Foundation

Jul. 2020 - Present

Lead team of four data scientists to research and develop a web tool, using state-of-the-art NLP methods and network analysis, to identify unmet needs and influential narratives after extreme weather events using Twitter discourse in partnership with the International Water Management Institute (IWMI). Leverage methods and tools such as zero shot text classification (with BART), topic modeling, POV extraction, text summarization, word/document embeddings, community detection, Google Translate API, Twitter API, and ElasticSearch. Serve as subject matter expert for applied natural language processing techniques. Facilitate communication between technical team and IWMI, hosting biweekly partner calls and weekly internal team meetings. Ensure efficient management of the project and timely delivery of deliverables.

Doctoral Research Fellow, UCD Geary Institute for Public Policy

Nov. 2016 - Aug. 2019

Developed semi-automatic methodology to assess the impact of central bank communication on U.S. Treasury yields using topic modeling. Collaborated in team to develop method to predict asset returns for U.S. equities (all stocks in SP 500 and Russell 1000) using predictive machine learning models. Built hybrid machine learning and econometric time series models for macroeconomic variables. Communicated research findings through journal articles, academic conferences, and non-academic conferences such as meetup groups.

Data Science Fellow, Data Science for Social Good Foundation

May 2018 - Aug. 2018

Collaborated in a team of data scientists to build a machine learning system for the Department of Labor in Chile to predict companies' likelihood of violating labor laws during labor inspections. Utilized AWS, PostgreSQL, and Python to compute, store data and build models with confidential government data spanning the entire country. Built a generalizable pipeline that can be used by other government institutions for similar tasks.

EDUCATION

University College Dublin

Sep. 2016 - Dec. 2020

Doctor of Philosophy in Economics

Thesis: Essays in Applied Machine Learning for Macroeconomic Policy Analysis

Relevant Coursework: Macroeconomics, Econometrics [Panel, Time Series, Cross Section, Causal Inference], Machine Learning, Advanced Machine Learning, Deep Learning, Multivariate Analysis, Bayesian Analysis, Mathematical Statistics

University of Glasgow

Sep. 2015 - Aug. 2016

Master of Science in Financial Economics

Relevant Coursework: Modeling and Forecasting Financial Markets, Econometrics I & II, Economic Fundamentals and Financial Markets

Christopher Newport University

Aug. 2011 - May 2015

Bachelor of Arts in Economics

Bachelor of Music in Clarinet Performance

SKILLS

Programming	Python (advanced), Matlab (intermediate), R (basic)
Data Tools	PostgreSQL, AWS, Google Cloud, Git/Github, Bash/Linux, Spark/PySpark
Statistics	Stata, EViews, OxMetrics
Typesetting	MS Office, L ^A T _E X
Languages	English (native), German (intermediate), Spanish (basic)

TRAINING & CERTIFICATIONS

AWS Certified Big Data - Specialty	<i>Oct. 2019</i>
AWS Certified Solutions Architect - Associate	<i>Sep. 2019</i>
Elastic Certified Engineer (Elastic)	<i>Nov. 2019</i>
Elasticsearch Engineer I (Elastic)	<i>Oct. 2019</i>
Elasticsearch Engineer II (Elastic)	<i>Oct. 2019</i>
Deep Learning Specialization (Deeplearning.ai; Coursera)	<i>Jun. 2018</i>
Social and Economic Networks and Models (Stanford University (by Matthew Jackson); Coursera)	<i>Jan. 2018</i>
Machine Learning (Stanford University (by Andrew Ng); Coursera)	<i>Aug. 2017</i>

TEACHING EXPERIENCE

Graduate Teaching Assistant, University College Dublin	<i>Sep. 2016 - Jul. 2019</i>
Graduate Mathematics for Economists (Lecturer)	<i>Fall 2018</i>
Graduate Macroeconomic Theory	<i>Fall 2017, 2018</i>
Intermediate Macroeconomics	<i>Spring 2017, 2018, 2019</i>
Principles of Microeconomics	<i>Fall 2016</i>

CONFERENCES & SEMINARS

Tackling Climate Change with Machine Learning Workshop, NeurIPS 2020	<i>Dec. 2020</i>
NVIDIA GPU Technology Conference (GTC), Washington, D.C.	<i>Nov. 2019</i>
Open Data Science Conference (ODSC), Dublin (invited speaker)	<i>Jul. 2019</i>
24th Spring Meeting of Young Economists, Universite Libre Bruxelles	<i>Apr. 2019</i>
(EC) ² Big Data Econometrics Conference, Bank of Italy	<i>Dec. 2018</i>
27th International Rome Conference in Money, Banking, and Finance	<i>Dec. 2018</i>
2018 PyCon Ireland	<i>Nov. 2018</i>
2018 DataFest Chicago	<i>Aug. 2018</i>
Department of Economics, Machine Learning Group, University of Oxford	<i>May 2018</i>
2015 Association of Private Enterprise Education Conference	<i>Apr. 2015</i>
2015 Virginia Association of Economists Conference, Randolph Macon College	<i>Mar. 2015</i>

RESEARCH

[1] Narratives and Needs: Analyzing Experiences of Cyclone Amphan Using Twitter Discourse (with João Fonseca, Kanav Mehra, Michelle Ng, Jared Ross, Marcelo Sandoval-Castañeda, and Rachel von Gnechten), *ArXiv Preprint*, <https://arxiv.org/abs/2009.05560>. *Tackling Climate Change with Machine Learning Workshop at NeurIPS 2020. AI for Social Good Workshop at IJCAI 2020.*

[2] Central Bank Communication and the Yield Curve: A Semi-Automatic Approach using Non-Negative Matrix Factorization, *ArXiv Preprint*, <https://arxiv.org/abs/1809.08718>

[3] Antinoise (with Clemens Struck and Enoch Cheng), *Working Paper*

[4] Macroeconomic Forecasting: Comparing Linear, Nonlinear, and Hybrid Approaches, *Work in Progress*

[5] Reg2Vec: A Spatial Mapping of Economic Regulation (with Aniket Kesari), *Work in Progress*

PROFESSIONAL ACTIVITIES

Journal Referee: Journal of Econometrics, Econometric Reviews

Program Committee Member: AI for Social Good Workshop (IJCAI 2021)