

Contact

Phone: 087 1211 867

Location: Cork, Ireland

Email:
oliverjnorton@gmail.com

Other:
[Website](#), [GitHub](#),
[LinkedIn](#)

Skills

Programming

- Python
- SQL
- R

Data Science

Libraries

- PyTorch
- scikit-learn
- XGBoost
- Pandas
- NumPy
- Prophet
- BeautifulSoup

Data Engineering & Cloud

- Docker
- Azure
- dbt
- Linux

Data Visualisation

- Power BI
- Tableau
- matplotlib
- Plotly
- Seaborn

Other

- Git
- Confluence
- Jira
- Jupyter
- RStudio
- Bash
- JavaScript

Oliver Norton

Experience

Feb 2023 - Jan 2024, Junior Data Scientist, *Meltwater*

- Developed time series models with Prophet to forecast demand for food ingredients in a €150,000 client project
- Automated manual tasks and improved workflows, saving over 300 billable hours for colleagues
- Managed Linux VMs, leveraged social media data platforms, and developed Python scripts for ETL processes
- Collaborated with cross-functional teams to deliver insights and effectively communicated technical concepts to non-technical colleagues

Projects

Aether Energy Analytics data process automation

Jul 2024 | **Data Engineering** | View project on: [GitHub](#)

- Designed and implemented robust data pipelines with ETL -like process, using dbt, Docker, Python, Bash, Git (GitHub) and Linux
- Transformed raw data and created data models using SQL queries
- Implemented rigorous testing and quality checks to ensure data integrity

Carpooling A/S Cloud architecture design

Jun 2024 | **Cloud Computing** | View project on: [Google Drive](#)

- In 72 hours, designed a data software architecture for client interested in migrating from an on-premise solution to the Cloud

Quantifying polarisation using networks and text data

Sep 2023 - Jan 2024 | **Network Science** | View on: [GitHub](#)

- Created a machine learning model (cross-validated Lasso regression) to explore a large dataset of over 80 GB
- Used cloud virtual machines (VM) for model execution and analysis
- Developed a novel multi-dimensional approach to quantify political polarisation in multi-party systems using Twitter data

Investigating polarisation in Reddit Communities

Apr - May 2023 | **Data Science**

- Collected and cleaned over 50 GB of semi-structured Reddit data using a REST API (& Web scraping) for in-depth analysis
- Trained a multi-label sentiment classification model to categorise users

Analysing climate discourse with NLP & network analysis

May - Jun 2022 | **Data Science** | Visualisation on: [Google Drive](#)

- Analysed Twitter API text data, focusing on sentiment and discourse
- Developed ML classification models and used LDA topic modelling to identify key themes and trends in the conversation
- Visualised a retweet network employing community detection algorithms

Education

2021-2024, MSc in Social Data Science, *University of Copenhagen*

- Data Science with specialisation in behaviour, networks and ideas
- Grade Average: 9.5/12 with grade 12/12 in master's thesis

2017-2021, BA in Business, Economic & Social Studies, *Trinity College Dublin*

- Grade Average: 70% (First Class Honours)