Mark Dunne

Email: markdunne02@gmail.com   Website: markdunne.github.io

Experienced and adaptable data scientist with a history working in finance and genomics. Skilled in Python, Java, Machine Learning, and Data Engineering. Looking to solve interesting problems.

Experience

Data Scientist at Desktop Genetics
October 2016 - Present, London

Using data science and machine learning as part of the R&D team to gain insights into genomics and design of CRISPR experiments. Projects include building algorithms and models to predict CRISPR on-target and off-target activity. Developed novel CRISPR on-target activity scoring function that out performed the state of the art. Presented results to clients, wrote about [applying machine learning to CRISPR](#), and designed poster to present at Cold Spring Harbor Laboratory.

Analyst at Goldman Sachs
July 2015 - October 2016, London

Responsible for developing and maintaining software used daily by the Interest Rates sales team. Interacted daily with traders and sales people to support the software and understand the key requirements. Worked on project which used natural language processing to rapidly process trade inquiries, saving the company large amounts of otherwise lost revenue.

Summer Analyst at Goldman Sachs
Summer 2014, London

Developed key piece of in-house Interest Rates trading software used to process billions of dollars of transactions daily. Full stack development included everything from provisioning the bare-metal machines for the back end to working with end users to design the UI on the front end. Presented numerous smaller group projects to senior leaders. Gained in-depth knowledge about the financial industry, and in particular Interest Rate Products. Was asked to be campus ambassador at University College Cork for Goldman Sachs.

Software Engineer at HeyZap
Summer 2013, San Francisco

Developed innovative mobile adverts using HTML5 and CSS3. Quickly became the go-to guy for CSS related issues.

Freelance Technical Writer
2010 - 2012

Wrote tutorials for the TutsPlus network. Topics included maximizing performance in Adobe Flash, Threading in Android, and on Beginning Python. All articles were highly rated by the readers and received hugely positive feedback.

Education

University College Cork
2011 - 2015

B.Sc (Hons) Computer Science, First Class Honours

Key Subjects: Artificial Intelligence, Machine Learning, Algorithm Analysis, Theory of Computation, Linear & Abstract Algebra

Awards:

- *Title of College Scholar -* Awarded in 2013, 2014, and 2015 for achieve First Class Honours in end of year exams.
- *Most Innovative Final Year Project -* Awarded most Innovative Final Year Project by Fidelity Investments for work on the application of Machine Learning to the stock market. Analysed various stock market prediction techniques and robustly evaluated their predictive value.

Online Courses

- *edX 2016 -* Introduction to Biology - The Secret of Life
- *Coursera 2015 -* Machine Learning
- *Coursera 2014 -* Cryptography 1
- *Coursera 2013 -* Computational Investing

Ballincollig Community School   *2005 - 2011*

Key Subjects: Higher Level Mathematics (A2), Higher Level Applied Mathematics (A2), Higher Level Physics (B1)

Represented school in mathematics and linguistics competitions

Key Skills

- **Data Science:** Breadth of experience in machine learning and statistical modeling. Work with frameworks from the PyData Stack (**Pandas**, **Numpy**, **Scikit-Learn**, **SciPy**, **Matplotlib**, etc) on a daily basis. Also have extensive experience using **Tensorflow** to build high performance deep learning models, and **PyMC** to perform Bayesian statistical modeling. Big proponent of using anything but p-values and confidence intervals (they answer the wrong questions!). Understand how to translate statistical results for non-statisticians, and turn information into actionable information.
- **Data Engineering:** Have developed pipelines to process and model very large datasets, ranging from gigabytes into terabytes. Use **Dask** framework very regularly to perform large computations distributed over clusters running on **Google Cloud**. Also familiar with parts of the **Hadoop** stack such as **Apache Spark** and **Zookeeper**, and **Amazon AWS**.
- **Software Engineering:** Experience in working with enormous codebases (over 100M LoC), and working within teams. Understand how to write quality code that can be easily used and read by others later. Experience using **Git**, **SVN**, and **CVS** version control systems.
- **Programming:** Professional experience with **Python**, **Java**, **C#**, **Javascript**, and **HTML/CSS**. Also have non-professional experience with **Haskell**, my favorite programming language.

Personal Projects

- *[Deep Neural Net English Synthesizer](#)* - Developed end-to-end English language synthesizer using some of the latest techniques in deep learning research.
- *Statistical Arbitrage System -* Used to make certain profit by exploiting gambling sites that offered competing, and significantly different, odds on the same event.
- *[Chrome Extension with 20k+ (peak) users](#)* - WikiTube extension brings YouTube videos to Wikipedia. Featured on CNET and LifeHacker.
- *[Analysis of Food Standards in the UK](#)* - Analysed 515,748 GIS data points to map food hygiene standards across UK and London

References available on request