Aleks Hughes

Linkedin Direct message for contact details | flexibleframework.xyz | uk.linkedin.com/in/alekshughes

Education

MSC ECONOMICS | TRINITY COLLEGE DUBLIN AND UNVIERISTY OF NOTTINGHAM | AUGUST 2014
- Thesis title: 'How do Oil Price Shocks Affect Scottish Real Economic Activity?'
- Selected for the Trinity College's 'Econometric Game' team 2014, econometricgame.com.
- Writer for student run newsletter 'Nottingham Economic Review', neronline.co.uk.

A/AS LEVELS | ST. PETERS SCHOOL, YORK | JULY 2008
- Maths (A), Further Maths (A), Economics (A), History (B) and Politics (a).
- UK Maths Trust Olympics team. Silver in UKMT Senior Challenge. Vice-Captain of Rowing.

Experience

DATA SCIENTIST | PARKER FITZGERALD | APRIL 2017 TO PRESENT
- Natural Language Processing (NLP) engineer. Developed Python NLP applications for text classification, sentiment analysis, topic modelling and entity recognition. Used a range of techniques including explainable results, text processing pipelines, pre-trained language models and embedded word vectors. Applications have won and been shortlisted for two industry awards: risk.net and mca.org.uk.

- Topological Data Analysis (TDA) developer. Used Python and TDA to segment customers spending habits into 400 densely populated segments. Model is now used as an input for alternative credit scoring and suspicious account activity models. Backtesting demonstrated a 10% boost to suspicious activity model precision.
  Also built TDA model to triage mortgage fraud investigations, improving investigation effectiveness by 15%.

- Built Python Bokeh data visualisation applications: server hosted and standalone HTML dashboards. Coded JavaScript user interactions and custom enhancements to the vis.js 3D plotting library. Embedded own unsupervised segmentation algorithm for rogue employee behaviour into the dashboard.

- Architected and built R-shiny GUI data visualisation application. Original requirements of exploring large data column interdependence, generating summary stats and a representative sample. Application won an industry award for "best use of technology", auditandrisk.org.uk.

- Architected and built time series predictive ARIMA model in R to identify behaviours associated with poor risk management across bank business lines. Back testing over 2018 showed that application could have caught 4 high priority material events.

  Experience using range of different languages: Python, R, Java (Kotlin), JavaScript (TypeScript), UNIX and Solidity Ethereum blockchain. And tools: Docker, Travis CI, git, AWS ML (Sagemaker), Ayasdi TDA platform, Kortical automated ML.

- Leadership: recruited and line manage 2 consultant Data Scientists and 1 intern.

VOLUNTEER NLP ENGINEER | FULL FACT | OCTOBER 2018 TO FEBRUARY 2019
- Project to "Automate Fact Checking". Responsible for parsing parts of a claim into geography, direction, time frame, etc. Deployed model using Docker.
- Project was led and architected by a core maintainer of a leading Python NLP package "Genism".

TECH CONSULTANT | IBM CONSULTING | OCTOBER 2014 TO APRIL 2017
– Graduate Managers Choice Award, October 2016. IBM Europe Award 'Putting Client First', August 2015.
– Python Developer and Business Analyst on Data Remediation, large universal bank.
  Scripted lightweight and multi-threaded Python application. Git version control, test executed in Bluemix.
  Gathered requirements, wrote sections of solution design document.
– Business Analyst and Python Developer on Data Migration, large UK credit card platform.
  Built application to map 300,000+ end customers to 1,500+ target billing and payment configurations.
– Commercial Manager for IBM internal account, large universal bank.
  Modelled resourcing and prices for 40+ contracts with a value range of £100k to £10million. Scripted VBA macros to automate project financial trackers. Resolved 3 longstanding issues around invoicing, debt collection and purchase order issuance, ensuring inflow of funds.
– Organised 17 Lunch and Learns, topics include DevOps, Design Thinking, Agile and Blockchain. Also organised Excel training and a joint NHS-IBM healthcare hackathon.

Skills

TECHNICAL
– Machine Learning and statistical techniques
  NLP: text classification, topic modelling, entity recognition, text parsing, word vectors, and universal language models (e.g. BERT and ELMo).
  Neural Networks: CNNs, RNNs and LSTMs, building and parameterising different architectures.
  TDA: Various applications, lenses and metrics.
  General: regression, classification (multi-class, multi-label and ordinal), imbalanced datasets, time series.

– Python packages and techniques
  ML: SciPy, numpy, pandas, numba, keras and tensorflow, pyspark, scikit-learn processing model pipelines, AWS Sagemaker and controlled model release pipelines.
  NLP frameworks: spaCy, genism, pyText, allenNLP and NLTK.
  General: test driven development, building my own class hierarchy object model, modularising applications, decorators to enhance method functionality, SQL interfaces and data layer abstractions, API integrations, web scraping, regex matching and application logging.

– Familiar with diverse programming paradigms and patterns, e.g. distributed and multi-threaded execution, concurrency, building reactive GUIs and web frontends and lazy evaluation.

– Miscellaneous: Advanced Microsoft Office skills. Windows, Mac and POSIX operating systems. 85 words per minute typing speed.

LANUAGES
Native English, fluent Spanish, strong Dutch.

INTERESTS
Keen cyclist and swimmer. University Basketball 2nd team. 6 months volunteer English teacher in rural Mexican primary school. Combined 10 weeks Spanish immersion in Cuba and Spain. Biked 4,500Km through Alaska, hiked length of Pyrenees, through Patagonia and Atlas Mountains and spent 2 weeks dogsledding in Yukon.

References
Available on request