Fast Track Data/AI Graduate Programme

Covering Letter

Roy Wilson

26 August 2021

My interest in data science was sparked when I attended a handful of intensive courses for the R programming language from Jumping Rivers funded by the Bangor Doctoral School. These courses provided an excellent introduction to the R statistical programming language aimed to train researchers to become competent users of R. The courses include an introduction to R, data manipulation, statistical modelling, functional programming, and advanced graphics in R. These courses provided me with the confidence to interrogate and extract meaning from data, as well as the confidence to explore packages on CRAN to discover what else I could do with the language. I thoroughly enjoy exploring these capabilities which are beyond what Excel alone could offer. I went on to build a website in R using the blogdown package which can be found at RoyWilson.UK. This has proved useful in storing useful code snippets that I can easily share with others. I have also assisted my brother with his research into high entropy alloys. I built an interactive dashboard in R shiny which calculates the thermophysical properties and the thermal expansion coefficients of a selected alloy composition quickly with an easy to use user-interface. This dashboard can be found at RoyWilson.UK/HEA/. I have also explored machine learning techniques found within the caret package and applied them to datasets from the TidyTuesday project and those found on Kaggle with success in building predictive models.

I have also picked up practical skills in Python, and SQL throughout various courses. I really enjoy learning new tools as well as seeing what I can build with them. An example of where my Python skills were useful was when I heard that my brother and his supervisors were downloading molecular data from a supercomputer to then randomise in Excel, with each file being saved separately and reuploaded to the supercomputer. I wrote a script in Python that takes in arguments from the user to automatically produce any number of randomised molecular data files that retain their unique structure for use with the VASP system. The script structured the directories appropriately to aid in running calculations on the data and produces these input files directly on the supercomputer without any need for downloading or uploading data. The ability to run this process directly from the command line interface has saved my brother and his research group countless hours by allowing them to focus on the science, rather than repetitive processes that should be automated.

I studied Accounting and Finance with an MSc (Distinction) and a BSc (First-class). The recent masters programme provided me with training and an appreciation for the philosophy of science, and a strong interest in working with data and applied statistical techniques. Data and statistical heavy modules include statistical methods, research methods, financial econometrics, among general mathematical applications within finance. The programme has developed my analytical capabilities and instilled a confidence and curiosity for tackling complex problems. I also undertook maths at A level and have been developing my mathematical ability since.

I am excited for the opportunity that this programme provides. I want to be able to develop my data skills and understanding in cutting-edge techniques within data science and be able to combine with my business acumen to enhance what value I could bring to any future roles. What draws me to data science is its potential for life-long learning, where the data scientist is on the frontier of development of cutting-edge techniques and software packages. I see myself finding a lot of success on the Data/AI Fast Track Graduate programme as it looks to offer a stimulating challenge in an area that I thoroughly enjoy.

Sincerely,

Roy Wilson