POSITION DESCRIPTION

Data Scientist

Position Level | 7 or 8
Faculty/Division | Medicine & Health
Position Number | 00098902
Original document creation | 20/10/ 2021

Position Summary

The Data Scientist will be responsible for strengthening and expanding the study of novel online data sources to identify emerging trends in illicit drug use, markets and associated behaviours for the Drug Trends program at the National Drug and Alcohol Research Centre (NDARC). The role will include expanding existing monitoring and reporting on darknet cryptomarket sale of illicit drugs, establishing new reporting on other online sources (e.g., internet discussion forums, social media platforms such as Reddit and Twitter, search engines) and building online interactive visualisations to interrogate and share findings. Drawing upon an array of computational and statistical skills, the Data Scientist will collaborate with leading figures in this field to address critical public health questions in relation to substance use.

The Data Scientist reports to the Senior Research Fellow, and has no direct reports.

Accountabilities

Specific accountabilities for this role include:
Level 7:

- Use Python for web scraping (through use of Selenium and Scrapy or BeautifulSoup), data cleansing, and analysis of online data.
- Manage and develop machine learning models using programming languages like Python (including maintenance of current LSTM neural network, and development of new classification tools utilizing natural language processing for online data).
- Maintain and expand online interactive data visualisation platforms (through use of Tableau, R Shiny or other alternatives), with a focus on automated reporting.
- Undertake high-quality data management, including updating and maintaining data processing pipeline and documentation, and ensuring data integrity.
• Proactively assist with research reports, peer-reviewed publications, presentations, ethics applications and progress reports. Report on study progress to the study investigator team to assist with timely completion of project activities within available resources.

• Contribute to planning for the program of research, including planning new projects and seeking new dissemination and communication strategies.

• Independently plan workload and manage timelines to meet project outcomes within specified timelines.

• Perform other duties as requested by your supervisor.

• Align with and actively demonstrate the UNSW Values in Action: Our Behaviours and the UNSW Code of Conduct.

• Cooperate with all health and safety policies and procedures of the university and take all reasonable care to ensure that your actions or omissions do not impact on the health and safety of yourself or others.

Level 8 (in addition to the above):

• Independently design, plan, and execute new projects in accordance with the overarching aims of the program of work.

• Provide substantial leadership in the preparation of peer-review papers.

Skills and Experience

Level 7:

• Relevant degree specialising in computer science, mathematics, biostatistics, data science, statistics, or other quantitative discipline, with relevant experience or equivalent competence gained through any combination of education, training and experience.

• Demonstrated experience in using Python for web scraping, data cleansing, and machine learning. Knowledge of other programming languages such as R would be advantageous.

• Demonstrated experience in building, testing and deploying machine learning models (GLM, Random Forest, GBM, Neural networks etc).

• Demonstrated experience in cleaning, merging, and managing large data sets, including implementing various strategies to maximise data integrity.

• Demonstrated experience in building online data visualisation platforms (e.g., Tableau, R Shiny).

• Excellent written and verbal communication skills, with a high level of attention to detail.

• An understanding of and commitment to UNSW’s aims, objectives and values in action, together with relevant policies and guidelines.

• Knowledge of health and safety responsibilities and commitment to attending relevant health and safety training.

Level 8 (in addition to the above):
• Doctoral degree in computer science, statistics, biostatistics, data science, mathematics, or other quantitative discipline, with substantial extension of the theories and principles normally requiring extensive relevant experience.

• Demonstrated capacity to conceptualise novel research questions and carry out independent research as indicated by a proven track record of leading reports or manuscripts for publication in peer-reviewed journals.

• Demonstrated evidence of bringing innovation and improvements in achieving core objectives and outputs.