

|  |
| --- |
| THE UNIVERSITY OF NOTTINGHAM  Recruitment Role Profile Form |

**Job Title:** Research Associate / Fellow in Computational Data Analytics (Fixed term)

**School/Department:** School of Computer Science

**Salary:** £25,769 to £37,768 £25,769 to £37,768 per annum, depending on skills and experience (minimum £28,982 with relevant PhD). Salary progression beyond this scale is subject to performance.

**Job Family and Level:** Research and Teaching – Level 4

**Contract Status:** This post will be offered on a fixed-term contract for 3 years

**Hours of Work:** Full-time(36.25 hours per week)

**Location:** Jubilee Campus

**Reporting to:** Dr Daniele Soria, Prof. Jon Garibaldi

**Purpose of the Role:**

The purpose of the new role is to create computational algorithms to model and analyse wide-ranging data associated with the Nottingham Molecular Pathology Node project. Molecular pathology produces diverse, complex data and challenges lie in combining these to deliver clinically useful information. The data are multi-modal (numeric, textual, images), of high dimensionality and contain uncertainties (e.g. empty data fields). Work will include incorporating technologies allowing analytical outputs to be incorporated into clinical databases to generate new knowledge and research hypotheses with the ultimate goal of supporting clinical decision-making.

|  |  |  |
| --- | --- | --- |
|  | **Main Responsibilities** | **% time per year** |
| 1. | Develop new computational methodologies for the analysis of wide-ranging data | 60% |
| 2. | Produce conference and journal articles describing the developed methodologies | 20% |
| 3. | Disseminate results at conferences | 10% |
| 4. | Liaise with other stakeholders and members of the Node | 10% |
| 5. | Any other duties appropriate to the grade and role. |  |

Molecular pathology produces diverse, complex data and challenges lie in combining these to deliver clinically useful information. The data are multi-modal (numeric, textual, images), of high dimensionality and contain uncertainties (e.g. empty data fields). The successful candidate will create computational algorithms to model and analyse these wide-ranging data. In particular, the appointed PDRA will contribute to the following three areas:

(i) Data processing and feature extraction algorithms will process raw data using dimensionality reduction techniques to extract relevant features and key information, more amenable to processing.

(ii) Data aggregation algorithms will combine multi-modal data into meaningful sets of related data (e.g. pertaining to the same patient, same organ, same experiment etc.) whilst representing characteristics (e.g. uncertainty) limiting the overall size of the data and retaining essential information.

(iii) Knowledge modelling and decision support algorithms will include unsupervised/semi-supervised clustering techniques to elucidate new groups from unlabelled data, supervised classification techniques to distinguish between known classes or outcomes, forward/backward regression to test feature interaction and predictive techniques for time-series data. These algorithms will create new understanding such as novel models of patient stratification or novel interpretative systems supporting clinical decision-making. The successful candidate will work on incorporating technologies allowing analytical outputs to be incorporated into clinical databases to generate new knowledge and research hypotheses. Specific case studies will be utilised to populate the models with real-world data, both from within NMPN node and from other Nodes. The generic analysis pipeline the successful candidate will help creating will be useful to other similar Nodes in the country.

**Knowledge, Skills, Qualifications & Experience**

|  |  |  |
| --- | --- | --- |
|  | **Essential** | **Desirable** |
| **Qualifications/ Education** | * PhD awarded (or equivalent) or close to completion in computer science or relevant subjects. | * Data mining / AI experience. |
| **Skills/Training** | * Specialist research skills and techniques - intelligent analysis of multi-modal data and knowledge in the presence of uncertainty and noise. * Knowledge of statistical research methods and statistical analysis skills. * Writing scientific code in Matlab, and/or Python and/or C++ and/or Java and/or R. * Strong organisational, collaborative, and communication skills for dissemination of results. | * Computational intelligence techniques for medical decision support. * Analysis of medical-related data. |
| **Experience** | * Experience in data gathering, data aggregation, and machine learning techniques. | * Experience with the analysis of immunohistochemistry or gene expression data. |

**Decision Making**

1. **taken independently by the role holder**

|  |
| --- |
| Day-to-day processes and procedures, including quality assurance measures, data analysis processes, storage and archiving protocol. |

1. **taken in collaboration with others**

|  |
| --- |
| Organisation of research.  Working methods.  Development of research instruments.  Data collection and analysis strategy.  Research direction and dissemination – form and content.  Identification of relevant peer-reviewed journals and conferences for publishing research findings.  Structure of research reports and papers. |

1. **referred to the appropriate line managers (Dr Daniele Soria and Professor Jon Garibaldi) by the role holder**

|  |
| --- |
| Dissemination strategy.  Research directions. |

**Please quote ref: SCI/353615.**

|  |  |
| --- | --- |
| S:\PS\Registrars\HR Policy\CAT Office\Web Site\Images\AS_Silver Award.tif | *The University of Nottingham strongly endorses Athena SWAN principles, with commitment from all levels of the organisation in furthering women’s careers. It is our mission to ensure equal opportunity, best working practices and fair policies for all.* |