Candidate Information

Position:	Research Fellow in Artificial Intelligence Security
School/Department:	School of Electronics, Electrical Engineering and Computer Science
Reference:	24/112193
Closing Date:	Monday 7 October 2024
Salary:	£46,249 per annum
Anticipated Interview Date:	Wednesday 16 October 2024
Duration:	3 years

JOB PURPOSE:

To perform research in Trustworthy AI. The role involves gaining a deep understanding of AI systems' robustness, security, safety, and explainability for image classification problems relevant to a partner company. A key requirement is to design robust deep neural networks (DNNs) extending the latest state-of-the-art research and their formal verification exploiting frameworks such as auto_lirpa using both proprietary and public datasets.

MAJOR DUTIES:

1. Undertake high-quality and novel research in the area of adversarial training for robust image classification, adversarial attacks for model testing, general approach of formal verification of neural network image classifiers, and explainability of image classifier model decisions.
2. Design, develop, and refine experiments to evaluate CNNs performance and robustness computation.
3. Carry out analyses, critical evaluations, and interpretations using methodologies and other techniques appropriate to AI trustworthiness.
4. Present regular progress reports on research to members of the research group or to external audiences to disseminate and publicise research findings.
5. Prepare, often in consultation with the line manager, material for publication in national and international journals and presentations at international conferences such as IEEE Security and Privacy, ACM in Computers and Communication Security.
6. Produce high-quality research outputs consistent with project aims and commensurate with the career stage.
7. Undertake supplementary duties relevant to the success of the project including administrative duties and additional training and development activities as required.

ESSENTIAL CRITERIA:

1. Normally have or about to obtain a PhD in engineering or physical sciences area.
2. Recent high quality research experience in machine learning/AI or cybersecurity, or both, as evidenced by a strong track record of publications in leading journals and conferences in relevant areas.
3. Software programming skills.
4. Ability to assess and organise resources.
5. A consummate team player is prepared to work closely with other members of a large multidisciplinary research and development team, as well as with industrial collaborators.
6. Evidence of being a strong communicator with excellent oral and written communication skills.
7. Demonstrate resilience and the ability to work in a fast-changing environment with competing priorities.
8. Demonstrates a high degree of integrity, honesty and openness in professional conduct.
9. Able to visit collaborative partners and to attend meetings and conferences nationally and internationally as requested.
10. On-site presence required for minimum 3 days per week.

DESIRABLE CRITERIA:

1. Normally have or be about to obtain a PhD in the area of machine learning/AI, cybersecurity or cyberAI.
2. Experience in deep learning neural networks such as CNNs, auto-encoders, transformers, and self-attention networks.

3. Experience in trustworthy AI systems and applications such as:
   - Robust image classification
   - Adversarial attacks
   - Adversarial training and testing
   - Verification of models
   - Explainability of models.
4. Experience in initiating and developing research plans.
5. Experience in collaborating with industry.
6. Proficient in Python.
7. Experience using PyTorch, Keras, and/or TensorFlow. Jupyter Notebook.
8. Ability to design, train and test machine learning/AI systems using appropriate methodologies and datasets.
9. Familiarity with the theory of Machine Learning fundamentals (Statistics, Optimization, Linear Algebra, Partial Derivative Equations, etc.)
10. Demonstrable commitment to the purpose and objectives of CSIT and Cyber AI Hub.

ADDITIONAL INFORMATION:

Informal Enquiries to Sandeep Gupta: s.gupta@qub.ac.uk