Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Post Specification

Post Title:	Research Fellow in Security and Privacy implications of unused parameters of Machine Learning Models
Post Status:	Specific Purpose Contract – Full Time
Research Group / Department / School:	School of Computer Science and Statistics, Trinity College Dublin, The University of Dublin
Location:	School of Computer Science and Statistics, Trinity College Dublin, the University of Dublin College Green, Dublin 2, Ireland
Reports to:	Prof. Alessio Benavoli
Salary:	Appointment will be made on the Research Ireland Salary Scale at a point in line with Government Pay Policy (up to ~€71,000 per annum), appointment will be made no higher than level 2B point 1.
Hours of Work:	40
Closing Date:	12 Noon (GMT), 30th November 2025

Please note:

The successful applicant will be expected to take up post starting from January 2026.

Post Summary

We are seeking to recruit a Postdoctoral researcher with a background in game theoretic modelling, learning theory and both the theoretical and practical aspects of deep learning. In particular, we privilege the theoretical aspect. The successful applicant will collaborate within a skilled and motivated team of scientists (from Trinity College Dublin, Queen's University

Belfast and University California Riverside) dedicated to understand the security and privacy implications of the unused parameters present in machine learning models. The proposed research explores the problem both from the theoretical and empirical perspectives.

Description: Machine Learning applications increasingly rely on Deep Neural Networks (DeepNNs) to achieve state-of-the-art performance across a wide range of real-world tasks, including those powered by Large Language Models. DeepNNs are typically highly over-parameterised, but this characteristic has been shown to facilitate both optimisation and generalisation. Nevertheless, such over-parameterisation also introduces many parameters that are not essential to the final learned model and whose specific values have little impact on its behaviour (don't-care states). In both software and hardware systems, undefined behaviour and don't-care states have been shown to be potential sources of vulnerabilities. The postdoc will conduct research that contributes to the following streams of work: 1) Investigate the possible strategies that colluding parties could use to maliciously exploit the unused parameters (learning capacity) of a machine learning model using techniques from information and learning theory; 2) Model the attacks within a game-theoretic framework under complete or incomplete knowledge (Bayesian games), which will enable to find optimal strategies for the players (Attacker, Nature and Defender); 3) Validate these theoretical models and strategies on machine learning models and real applications.

The position will be under the direction of Prof. Alessio Benavoli. For informal inquiries please contact alessio.benavoli@tcd.ie

Standard Duties and Responsibilities of the Post

The Research Fellow is expected to conduct independent and collaborative research related to the project, develop and evaluate models or algorithms, and publish findings in leading journals and conferences. The role also involves collaboration with partner institutions and researchers across the project, participation in project meetings and travel as required, and contribution to project reporting, grant writing, and dissemination activities (e.g., workshops organisation).

Funding Information

The post is funded by Research Ireland.

Person Specification

Qualifications

The successful candidate must have a PhD in Computer Science, Computer Engineering, Mathematics, or a related field. The post is applicable to both new and experienced PhD holders, and salary will be commensurate with experience and achievement.

Knowledge & Experience (Essential & Desirable)

Essential:

- Expertise in game-theoretic modelling and statistical learning theory;
- Theoretical knowledge and practical experience in deep learning;
- Proven record of publishing in top-tier journals and conferences on topics central to the post's research focus.

Desirable: Experience in one or more of the following areas is considered preferable:

- Robustness in Statistics or Machine Learning;
- Adversarial attacks and training of machine learning models;
- Bayesian machine learning;
- Bayesian game theory;
- Research experience on Trustworthy AI.

Skills & Competencies

- Strong programming skills;
- Excellent written and oral communication skills;

- The ability to work well in a group;
- Strong self-motivation and willingness to learn.

Application Procedure

Applicants should submit a cover letter and a full Curriculum Vitae, including the names and contact details of two referees (with email addresses), no later than the 30th November 2025, via email (subject "Postdoctoral Research Fellow in Security and Privacy implications of unused parameters of Machine Learning models") to: Alessio.Benavoli@tcd.ie

Further Information for Applicants

URL Link to Area	www.tcd.ie
URL Link to Human Resources	https://www.tcd.ie/hr/

Trinity College Dublin, the University of Dublin

Trinity is Ireland's leading university and is ranked 75th in the world (QS World University Rankings 2026). Founded in 1592, the University is steeped in history with a reputation for excellence in education, research and innovation.

Located on an iconic campus in the heart of Dublin's city centre, Trinity has 18,000 undergraduate and postgraduate students across our three faculties – Arts, Humanities, and Social Sciences; Engineering, Mathematics and Science; and Health Sciences.

Trinity is ranked as the 31st most international university in the world (Times Higher Education Rankings 2024) and has students and staff from over 120 countries.

The pursuit of excellence through research and scholarship is at the heart of a Trinity education, and our researchers have an outstanding publication record and strong record of grant success. Trinity has developed 19 broad-based multidisciplinary research themes that cut across disciplines and facilitate world-leading research and collaboration within the University and with colleagues around the world. Trinity is also home to 5 leading flagship research institutes:

- Trinity Biomedical Sciences Institute (TBSI)
- Trinity College Institute of Neuroscience (TCIN)
- Trinity Translational Medical Institute (TTMI)
- Trinity Long Room Hub Arts and Humanities Research Institute (TLRH)
- Centre for Research on Adaptive Nanostructures and Nanodevices (CRANN)

Trinity is 1st in Europe for Producing Entrepreneurs for the 7th year in a row and Europe's only representative in the world's top-50 universities
(Pitchbook 2021-2022).

Trinity is home to the famous Old Library and to the historic Book of Kells as well as other internationally significant holdings in manuscripts, maps and early printed material. The

Trinity Library is a legal deposit library, granting the University the right to claim a copy of every book published in Ireland and the UK. At present, the Library's holdings span approximately 6.5 million printed items, 400,000 e-books and 150,000 e-journals.
With over 120,000 alumni, Trinity's tradition of independent intellectual inquiry has produced some of the world's finest, most original minds including the writers Oscar Wilde and Samuel Beckett (Nobel laureates), the mathematician William Rowan Hamilton and the physicist Ernest Walton (Nobel laureate), the political thinker Edmund Burke, and the former President of Ireland Mary Robinson. This tradition finds expression today in a campus culture of scholarship, innovation, creativity, entrepreneurship and dedication to societal reform.

Rankings

Trinity College Dublin is the top ranked university in Ireland. Using the QS methodology we are ranked 75th in the world and using the Times Higher Education World University Ranking methodology we are 173th in the World.

Full details are available at: www.tcd.ie/research/about/rankings.

The Selection Process in Trinity

The Selection Committee (Interview Panel) may include members of the Academic and Administrative community together with External Assessor(s) who are expert in the area. Applications will be acknowledged by email. If you do not receive confirmation of receipt within 1 day of submitting your application online, please contact the named hiring lead on the job specification immediately and prior to the closing date/time.

Given the degree of co-ordination and planning to have a Selection Committee available on the specified date, the University regrets that it may not be in a position to offer alternate selection dates. Where candidates are unavailable, reserves may be drawn from a shortlist. Outcomes of interviews are notified in writing to candidates and are issued no later than 5 working days following the selection day.

In some instances the Selection Committee may avail of telephone or video conferencing. The University's selection methods may consist of any or all of the following: Interviews, Presentations, Psychometric Testing, References and Situational Exercises.

It is the policy of the University to conduct pre-employment medical screening/full pre-employment medicals. Information supplied by candidates in their application (Cover Letter and CV) will be used to shortlist for interview.

Applications from non-EEA citizens are welcomed. However, eligibility is determined by the Department of Business, Enterprise and Innovation and further information on the Highly Skills Eligible Occupations List is set out in Schedule 3 of the Regulations https://dbei.gov.ie/en/What-We-Do/Workplace-and-Skills/Employment-Permits/Employment-Permit-Eligibility/Highly-Skilled-Eligible-Occupations-List/ and the Ineligible Categories of Employment are set out in Schedule 4 of the Regulations https://dbei.gov.ie/en/What-We-Do/Workplace-and-Skills/Employment-Permits/Employment-Permit-Eligibility/Ineligible-Categories-of-Employment/ . Non-EEA candidates should note that the onus is on them to secure a visa to travel to Ireland prior to interview. Non-EEA candidates should also be aware that even if successful at interview, an appointment to the post is contingent on the securing of an employment permit.

Equal Opportunities Policy

Trinity is an equal opportunities employer and is committed to employment policies, procedures and practices which do not discriminate on grounds such as gender, civil status, family status, age, disability, race, religious belief, sexual orientation or membership of the travelling community. On that basis we encourage and welcome talented people from all backgrounds to join our staff community. Trinity's Diversity Statement can be viewed in full at https://www.tcd.ie/diversity-inclusion/diversity-statement.

Pension Entitlements

This is a pensionable position and the provisions of the Public Service Superannuation (Miscellaneous Provisions) Act 2004 will apply in relation to retirement age for pension purposes. Details of the relevant Pension Scheme will be provided to the successful applicant.

Applicants should note that they will be required to complete a Pre-Employment Declaration to confirm whether or not they have previously availed of an Irish Public Service Scheme of incentivised early retirement or enhanced redundancy payment. Applicants will also be required to declare any entitlements to a Public Service pension benefit (in payment or preserved) from any other Irish Public Service employment.

Applicants formerly employed by the Irish Public Service that may previously have availed of an Irish Public Service Scheme of Incentivised early retirement or enhanced redundancy payment should ensure that they are not precluded from re-engagement in the Irish Public Service under the terms of such Schemes. Such queries should be directed to an applicant's former Irish Public Service Employer in the first instance.

Application Procedure

Application Procedure

Applicants should submit a cover letter and a full Curriculum Vitae, including the names and contact details of two referees (with email addresses), no later than the 30th November 2025, via email (subject “Postdoctoral Research Fellow in Security and Privacy implications of unused parameters of Machine Learning models”) to: Alessio.Benavoli@tcd.ie