Zertifikat seit 2008
audit familiengerechte hochschule

TECHNISCHE UNIVERSITÄT BERLIN

Kennziffer: IV-428/26 **Veröffentlicht:** 12.11.2026 **Beginn:** frühestmöglich, befristet für 3 Jahre
Vergütung: Entgeltgruppe 13 TV-L Berliner Hochschulen
Umfang: 100% Arbeitszeit; Teilzeitbeschäftigung ggf. möglich **Bewerbungsfrist:** 26.11.2026

Wiss. Mitarbeiter*in (d/m/w)

Fakultät IV - Elektrotechnik und Informatik, BIFOLD / FG Maschinelles Lernen und IT-Sicherheit

Über uns

Das Fachgebiet „Maschinelles Lernen und IT-Sicherheit“ am Institut BIFOLD sucht eine*n wissenschaftliche*n Mitarbeiter*in für das Forschungsprojekt „AMIRA“. Das Projekt untersucht die Sicherheit und Vertrauenswürdigkeit von Mechanistic Interpretability, einer Familie von Methoden zur Erklärung neuronaler Netze wie Circuit Discovery und Sparse Autoencoders. Ziel ist es, die Verwundbarkeiten dieser Methoden zu charakterisieren, neue Angriffe zu konzipieren und geeignete Verteidigungsmechanismen zu entwickeln.

Ihre Aufgaben

- Sicherheitsanalyse von Methoden der Mechanistic Interpretability
- Erforschung und Entwicklung von Angriffen gegen diese Methoden
- Erforschung und Entwicklung von Verteidigungsmechanismen für diese Methoden
- Wissenschaftliches Publizieren und die Möglichkeit zur Promotion

Ihr Profil

- Erfolgreich abgeschlossenes wissenschaftliches Hochschulstudium (Master, Diplom oder Äquivalent) in Informatik oder einem verwandten technischen Fach
- Sehr gute Kenntnisse im Bereich Maschinelles Lernen
- Sehr gute Kenntnisse im Bereich IT-Sicherheit
- Erfahrung im Bereich Adversarial Machine Learning
- Erfahrung im Bereich Mechanistic Interpretability
- Gute Deutsch- und/oder Englischkenntnisse sind erforderlich; Bereitschaft, die jeweils fehlenden Sprachkenntnisse zu erwerben
- Fähigkeit zu selbstständiger wissenschaftlicher Arbeit erwünscht

Zertifikat seit 2008
audit familiengerechte hochschule

TECHNISCHE UNIVERSITÄT BERLIN

Wir bieten

- Spannende und anspruchsvolle Forschung
- Kollegiales und kooperatives Arbeitsumfeld
- Renommiertes und engagiertes Team
- Internationales Forschungsnetzwerk

Hinweise zur Bewerbung

Ihre Bewerbung richten Sie bitte unter **Angabe der Kennziffer** mit den üblichen Unterlagen (in einem PDF-Dokument, max. 5 MB) ausschließlich per E-Mail an **hashmi@tu-berlin.de**.

Alle Unterlagen sind in einem einzigen PDF-Dokument einzureichen. Dazu zählen mindestens:

- Anschreiben
- Vollständiger Lebenslauf (CV)
- Zeugnisse und Notenübersichten (Master und Bachelor)
  Falls vorhanden, sind zusätzlich Empfehlungsschreiben und Publikationen beizufügen.

Mit der Abgabe einer Onlinebewerbung geben Sie als Bewerber*in Ihr Einverständnis, dass Ihre Daten elektronisch verarbeitet und gespeichert werden. Wir weisen darauf hin, dass bei ungeschützter Übersendung Ihrer Bewerbung auf elektronischem Wege keine Gewähr für die Sicherheit übermittelter persönlicher Daten übernommen werden kann. Datenschutzrechtliche Hinweise zur Verarbeitung Ihrer Daten gem. DSGVO finden Sie auf der Webseite der Personalabteilung: https://www.abt2-t.tu-berlin.de/menue/themen_a_z/datenschutzerklaerung/.

Zur Wahrung der Chancengleichheit zwischen Frauen und Männern sind Bewerbungen von Frauen mit der jeweiligen Qualifikation ausdrücklich erwünscht. Schwerbehinderte werden bei gleicher Eignung bevorzugt berücksichtigt. Die TU Berlin schätzt die Vielfalt ihrer Mitglieder und verfolgt die Ziele der Chancengleichheit. Bewerbungen von Menschen aller Nationalitäten und mit Migrationshintergrund sind herzlich willkommen.

Datenschutz:

Vollständige Stellenausschreibung: