



Technische Universität Berlin



Bei der Technischen Universität Berlin ist/sind folgende Stelle/n zu besetzen:

Wiss. Mitarbeiter*in (d/m/w) - Entgeltgruppe 13 TV-L Berliner Hochschulen

Teilzeitbeschäftigung ist ggf. möglich

Das Berlin Institute for the Foundations of Learning and Data (BIFOLD) ist eines von sechs nationalen KI-Zentren in Deutschland und wird vom Land Berlin und dem Bundesministerium für Bildung und Forschung gefördert. BIFOLD besteht aktuell aus 12 Forschungsgruppen mit über 150 Mitarbeitenden, einer Graduiertenschule und der BIFOLD-Geschäftsstelle. Beteiligt sind außerdem Fellows der großen Berliner Universitäten, der Charité – Universitätsmedizin Berlin sowie verschiedener weiterer nationaler und internationaler Universitäten und außeruniversitärer Forschungseinrichtungen.

Fakultät IV - The Berlin Institute for the Foundations of Learning and Data (BIFOLD)

Kennziffer: IV-177/25 (besetzbar ab sofort / befristet für 4 Jahre / Bewerbungsfristende 20.06.2025)

Aufgabenbeschreibung:

Das DEEM Lab (<https://deem.berlin>) sucht eine*n wissenschaftliche*n Mitarbeiter*in, zur Forschung in Responsible Data Engineering. Die Forschung wird in enger Zusammenarbeit mit Prof. Julia Stoyanovich von der New York University durchgeführt (<https://airesponsibly.net/people/julia/>).

Responsible Data Engineering entwickelt sich zu einer neuen Disziplin an der Schnittstelle von Data Engineering und KI, die Ethik, Rechtskonformität und Inklusivität als zentrale Designüberlegungen behandelt. Der ganzheitliche Charakter dieses Ansatzes basiert auf der Einsicht, dass die Entscheidungen, die wir während der Datenerfassung und -aufbereitung treffen, tiefgreifende Auswirkungen auf die KI-Systeme haben, die wir erstellen und einsetzen.

Das Ziel dieser Position besteht darin, ein neues System zu schaffen, das Dateningenieuren dabei hilft, Datenvorbereitungspipelines zu entwerfen, die die Modellleistung im Hinblick auf eine Vielzahl von Metriken optimieren, beispielsweise im Blick auf Genauigkeit, Robustheit, Fairness und Rechtskonformität. Zu diesem Zweck wird das System Dateningenieure proaktiv durch die Auswahl und Bewertung einer großen Menge an Datenvorverarbeitungs-, Datenaugmentierungs- und Feature-Encoding Methoden führen. Im Rahmen des Forschungsprojekts wird eine zuverlässige, effiziente und einfach zu bedienende Open-Source-Implementierung dieses Systems erstellt.

Dieses Unterfangen ist in mehrfacher Hinsicht technisch anspruchsvoll. Erstens müssen Datenaufbereitung und Modellauswahl für mehrere Zielmetriken optimiert werden, im Gegensatz zu bestehenden Ansätzen, die sich nur auf eine einzelne Metrik konzentrieren. Zweitens muss das System eine große Anzahl verschiedener Pipeline-Varianten erstellen, umschreiben und gleichzeitig ausführen, was eine effiziente Laufzeit und neuartige Techniken zur Anfrageoptimierung erfordert. Drittens muss die Forschung aktuelle dramatische Veränderungen in den Entwicklungspraktiken von KI-Anwendungen berücksichtigen, z. B. KI-unterstützte Programmierung, Tabular Foundation Models und KI-basierte Data-Science-Agents.

Erwartete Qualifikationen:

- Erfolgreich abgeschlossenes wissenschaftliches Hochschulstudium (Master, Diplom oder Äquivalent) in Informatik oder Künstlicher Intelligenz
- Sehr gute Programmierkenntnisse in Python und mindestens einer weiteren Sprache (Java/Rust/C++)
- Kenntnisse in der Datenverarbeitung mit Dataflow-Systemen, relationalen Datenbanken und/oder Dataframe-Bibliotheken (z. B. Apache Spark, DuckDB, Pandas usw.)
- Erfahrung mit der Steigerung der Effizienz, Skalierbarkeit und Korrektheit datenzentrierter Programme
- Grundkenntnisse in maschinellem Lernen und gängigen Bibliotheken (z. B. Pandas, Sklearn, Pytorch, SparkML usw.)
- Gute Deutsch- und/oder Englischkenntnisse erforderlich; Bereitschaft, die jeweils fehlenden Sprachkenntnisse zu erwerben

Wünschenswert:

- Erfahrungen mit realen Datenverarbeitungssystemen und/oder ML-Deployments (z. B. aus Praktika, Jobs oder unternehmerischer Erfahrung)
- Beiträge zu Open-Source-Projekten

Ihre Bewerbung senden Sie bitte unter **Angabe der Kennziffer** mit den üblichen Unterlagen (mindestens Lebenslauf, Motivationsschreiben und Zeugnisse) ausschließlich per E-Mail an Prof. Dr. Sebastian Schelter unter **schelter@tu-berlin.de** .

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-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.

Die Stellenausschreibung ist auch im Internet abrufbar unter:
<https://www.personalabteilung.tu-berlin.de/menuue/jobs/>

