

## Titel

# Fairness-Guided Graph Augmentation: Nutzung sensitiver Attribute zur Verbesserung von Fairness in Graph Neural Networks (B.Sc./M.Sc.)

## Hintergrund

Graph Neural Networks (GNNs) werden zunehmend in sensiblen Anwendungen eingesetzt, in denen Verzerrungen gegenüber geschützten Gruppen problematisch sind. Bestehende Ansätze zur Fairnessverbesserung verfolgen unterschiedliche Strategien: Methoden wie FG-SMOTE versuchen, sensitive Attribute indirekt zu entfernen oder zu entkoppeln, während Modelle wie FairGNN diese explizit im Trainingsprozess nutzen.

Ein bislang wenig untersuchter Ansatz ist die gezielte Nutzung sensitiver Attribute auf Datenebene zur Steuerung von Graph-Augmentation. Insbesondere fehlt eine systematische Untersuchung, ob Fairness durch strukturierte, gruppenbasierte Erweiterung des Graphen verbessert werden kann.

## Aufgabenstellung

Ziel der Arbeit ist die Entwicklung und Evaluation einer neuen Methode zur Fairnessverbesserung in GNNs durch gezielte Graph-Augmentation.

Die Methode soll sensitive Attribute explizit nutzen, um benachteiligte Gruppen im Graphen zu identifizieren und durch strukturierte Generierung neuer Knoten und Kanten zu ergänzen.

Die Arbeit umfasst:

- Analyse bestehender Fairness-Methoden
- Entwicklung einer Fairness-gesteuerten Augmentationsstrategie
- Implementierung und experimentelle Evaluation
- Vergleich mit bestehenden Ansätzen hinsichtlich Genauigkeit und Fairness

## Theorieteil:

- Grundlagen von GNNs, Definitionen von Fairnessmetriken
- Überblick über Fairness-Methoden:
  - Pre-processing (z. B. Augmentation wie FG-SMOTE)
  - In-processing (z. B. FairGNN)
  - Post-processing

- Rolle sensitiver Attribute in Fairness-Ansätzen
- Diskussion des Spannungsfelds zwischen ignorieren und expliziter Nutzung sensitiver Attribute
- Herausforderungen bei Fairness in Graphdaten (z. B. strukturelle Verzerrungen, Homophilie)

### **Implementierung:**

- Nutzung bestehender Benchmark-Datensätze (z. B. Pokec, German Credit, Bail)
- (Nach)Implementierung und/oder Nutzung von:
  - einem Standard-GNN (Baseline)
  - FG-SMOTE
  - FairGNN
- Entwicklung einer eigenen Methode, z.B.: Fairness-Guided Graph Augmentation (FGGA)
- **Kernidee der Methode:**
  - Analyse:** Bestimmung von Fairness-Unterschieden zwischen sensitiven Gruppen (z. B. Vorhersagebias)
  - Identifikation:** Lokalisierung benachteiligter Gruppen und Regionen im Graphen (nicht nur global, sondern auch auf Nachbarschaftsebene)
  - Augmentation:**
    - Generierung neuer Knoten basierend auf ähnlichen Knoten derselben Gruppe
    - Ergänzung neuer Kanten zu strukturell passenden Nachbarn
    - Berücksichtigung von Graph-Eigenschaften (z. B. Nachbarschaftsstruktur)
  - Ziel:** Reduktion von Fairness-Unterschieden bei möglichst geringer Verschlechterung der Modellgüte
- Evaluation: Modellgüte (z. B. Accuracy / F1), Fairness-Metriken
- Vergleich und Analyse der Ergebnisse

### **Erforderliche Kenntnisse und Fähigkeiten**

Python, PyTorch Geometric/DGL, GNNs

### **Betreuerin**

Barbara Hoffmann

Fragen jederzeit gerne via Teams oder E-Mail ([barbara.hoffmann@uni-bayreuth.de](mailto:barbara.hoffmann@uni-bayreuth.de))