DKU

A Comprehensive and Innovative Study on Fairness in Graph Anomaly Detection

FairGAD

Shouju Wang

STATS403-Deep Learning

Overview

who-what-why-how

- Introduction
- Statement of the Problem
- Review of Related Literature
- Methodology
- Q&A

Back to Overview

Introduction

- Name : Shouju Wang （汪守菊）
- Come form Wuhan University
- Junior, major in computer science
- human factor & techonology

Greta Gerwig
Frances Ha (2012)

Statement of the Problem

How can we address fairness concerns to graph anomaly detection?

Back to Overview

Review of Related Literature

Fairness Anomaly Detection

- The goal of fair anomaly detection is to ensure that the detection outcome of an individual in the factual world is the same as that in the counterfactual world where the individual had belonged to a different group

Graph Anomaly Detection

- Graph Anomaly Detection (GAD) is a technique used to identify abnormal nodes within graphs, finding applications in network security, fraud detection, social media spam detection, and various other domains.

Anomaly Detection

- Anomaly detection (aka outlier analysis) is a step in data mining that identifies data points, events, and/or observations that deviate from a dataset’s normal behavior.

TOWARDS FAIR GRAPH ANOMALY DETECTION: PROBLEM, NEW DATASETS, AND EVALUATION

Anonymous authors
Paper under double-blind review

ABSTRACT

The Fair Graph Anomaly Detection (`FairGAD`) problem aims to accurately detect anomalous nodes in an input graph while ensuring fairness and avoiding biased predictions against individuals from sensitive subgroups such as gender or political leanings. Fairness in graphs is particularly crucial in anomaly detection areas such as misinformation detection, where decision outcomes can significantly affect individuals. Despite this need, existing works lack realistic datasets that encompass actual graph structures, anomaly labels, and sensitive attributes for research in `FairGAD`. To bridge this gap, we present two novel graph datasets constructed from the globally prominent social media platforms Reddit and Twitter. These datasets comprise 1.2 million and 400,000 edges associated with 9,000 and 47,000 nodes, respectively, and leverage political leanings as sensitive attributes and misinformation spreaders as anomaly labels. We demonstrate that our `FairGAD` datasets significantly differ from the synthetic datasets used currently by the research community. These new datasets offer significant values for `FairGAD` by providing realistic data that captures the intricacies of social networks. Using our datasets, we investigate the performance-fairness trade-off in nine existing GAD and non-graph AD methods on five fairness methods, which sheds light on their effectiveness and limitations in addressing the `FairGAD` problem.

Subproblems

Back to Overview

Dataset

How to identify an appropriate dataset for FairGAD and ensure its suitability for our research.

Deep Learning Base Method

What kind of DL methods I can use?

Fairness Evaluation

How can I incorporate suitable fairness evaluation to measure the methods?

A Comprehensive Survey on Graph Anomaly Detection with Deep Learning

Xiaoxiao Ma, Jia Wu, *Senior Member, IEEE,* Shan Xue, Jian Yang, Chuan Zhou
Quan Z. Sheng, and Hui Xiong, *Fellow, IEEE* and Leman Akoglu

Abstract—Anomalies are rare observations (*e.g.,* data records or events) that deviate significantly from the others in the sample. Over the past few decades, research on anomaly mining has received increasing interests due to the implications of these occurrences in a wide range of disciplines - for instance, security, finance, and medicine. For this reason, anomaly detection, which aims to identify these rare observations, has become one of the most vital tasks in the world and has shown its power in preventing detrimental events, such as financial fraud, network intrusions, and social spam. The detection task is typically solved by identifying outlying data points in the feature space, which, inherently, overlooks the relational information in real-world data. At the same time, graphs have been prevalently used to represent the structural/relational information, which raises the *graph anomaly detection problem* - identifying anomalous graph objects (*i.e.,* nodes, edges and sub-graphs) in a single graph, or anomalous graphs in a set/database of graphs. Conventional anomaly detection techniques cannot tackle this problem well because of the complexity of graph data (*e.g.,* irregular structures, relational dependencies, node/edge types/attributes/directions/multiplicities/weights, large scale, etc.). However, thanks to the advent of deep learning in breaking these limitations, graph anomaly detection with deep learning has received a growing attention recently. In this survey, we aim to provide a systematic and comprehensive review of the contemporary deep learning techniques for graph anomaly detection. Specifically, we provide a taxonomy that follows a task-driven strategy and categorizes existing work according to the anomalous graph objects that they can detect. We especially focus on the challenges in this research area and discuss the key intuitions, technical details as well as relative strengths and weaknesses of various techniques in each category. From the survey results, we highlight 12 future research directions spanning unsolved and emerging problems introduced by graph data, anomaly detection, deep learning and real-world applications. Additionally, to provide a wealth of useful resources for future studies, we have compiled a set of open-source implementations, public datasets, and commonly-used evaluation metrics. With this survey, our goal is to create a "one-stop-shop" that provides a unified understanding of the problem categories and existing approaches, publicly available hands-on resources, and high-impact open challenges for graph anomaly detection using deep learning.

Index Terms—Anomaly detection, outlier detection, fraud detection, rumor detection, fake news detection, spammer detection, misinformation, graph anomaly detection, deep learning, graph embedding, graph representation, graph neural networks.

Fairness in Deep Learning: A Computational Perspective

Mengnan Du, Fan Yang, Na Zou, Xia Hu

Abstract—Deep learning is increasingly being used in high-stake decision making applications that affect individual lives. However, deep learning models might exhibit algorithmic discrimination behaviors with respect to protected groups, potentially posing negative impacts on individuals and society. Therefore, fairness in deep learning has attracted tremendous attention recently. We provide a review covering recent progresses to tackle algorithmic fairness problems of deep learning from the computational perspective. Specifically, we show that interpretability can serve as a useful ingredient to diagnose the reasons that lead to algorithmic discrimination. We also discuss fairness mitigation approaches categorized according to three stages of deep learning life-cycle, aiming to push forward the area of fairness in deep learning and build genuinely fair and reliable deep learning systems.

Index Terms—Deep Learning, DNN, Fairness, Bias, Interpretability.

1 Introduction

MACHINE learning algorithms have achieved dramatic progress nowadays, and are increasingly being deployed in high-stake applications, including employment, criminal justice, personalized medicine, *etc* [1]. Nevertheless, *fairness in machine learning* remains a problem. Machine learning algorithms have the risk of amplifying societal stereotypes by over associating protected attributes, e.g., race and gender, with the prediction task [2]. Eventually they are capable of exhibiting discriminatory behaviors against certain subgroups. For example, a recruiting tool for STEM jobs believes that men are more qualified and shows bias against women [3], facial recognition performs extremely poorly for female with darker skin [4], recognition accuracy is very low for subgroup of people in pedestrian detection of self-driving cars [2]. The fairness problem might

opaque and hard to comprehend. This is problematic and makes it difficult to identify whether these models make decisions based on right and justified reasons, or due to biases. In addition, this makes it challenging to design bias detection and mitigation approaches.

In this article, we summarize *fairness in deep learning* work from the computational perspective, and do not discuss work from social science, law and many other disciplines [1]. Particularly we show that interpretability could significantly contribute to better understandings of the reasons that affect fairness. We also review fairness mitigation strategies categorized into three stages of deep learning life-cycle. Finally, we propose open challenges and future research directions. Throughout this article, *we don't differentiate between deep learning and DNN* (Deep neural network)

Methodology

Back to Overview

Review previous works	• I will review existing related work more deeply
Come up with hypothesis	• Try to combine these concepts together.
Do expriments	• Use expriments to varify my hypothesis

Back to Overview

Q&A Session

Thank you for listening!