User-friendly Explanations for Graph Neural Networks

Arijit Khan
Department of Computer Science
Aalborg University, Denmark
arijitk@cs.aau.dk

AALBORG UNIVERSITY

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE
School of Computer Science and Engineering

Systems@ETHzürich

Computer Science UCSB Education

A Brief History in Time …

- **University of California, Santa Barbara (UCSB)**
  - PhD (2008-2013)

Internship at **IBM TJ Watson, NY** (2010)
Internship at **Yahoo! Labs, Barcelona** (2012)

- **ETH Zurich, Switzerland**
  - Post-doc (2014-2015)

- **Nanyang Technological University (NTU), Singapore**
  - Assistant Professor (2016-2022)

- **Aalborg University**
  - Associate Professor (2022-now)

Computer Science UCSB

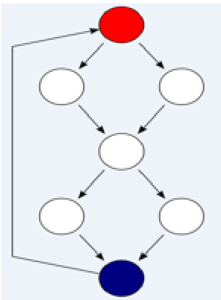
Systems@ETH zürich

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

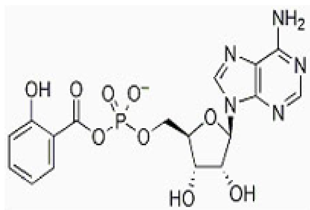
AALBORG UNIVERSITY

Graph Data is Everywhere

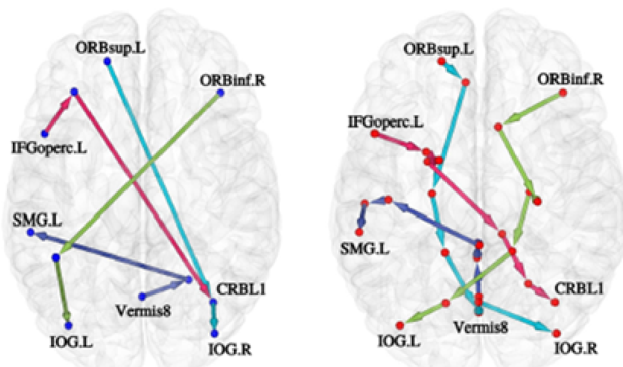
Graph database with many smaller graphs



Program flow/ call graph



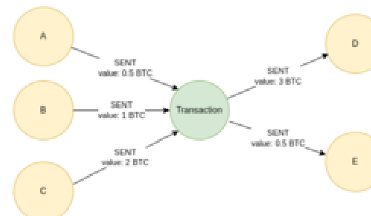
Chemical compound structure



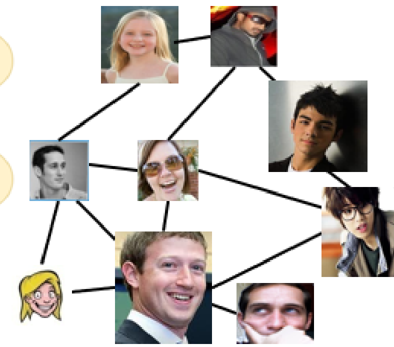
Human brain network

One large graph

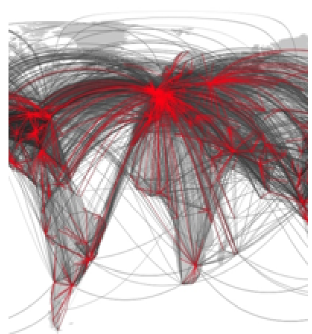
Biological network



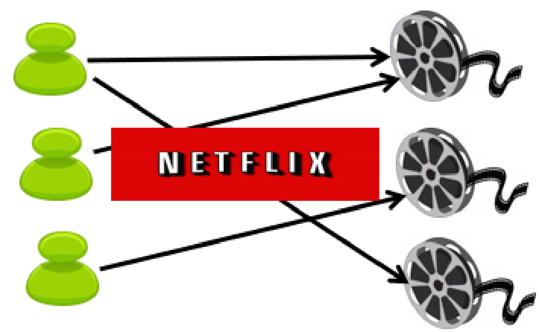
Financial transaction



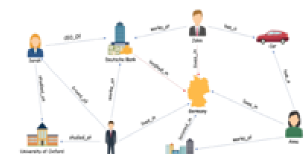
Social network



Transportation



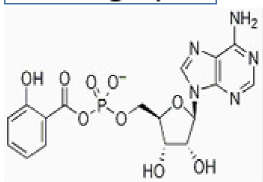
Recommendation



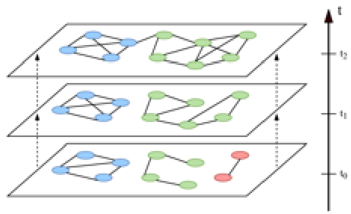
Knowledge graph

Graph Data in Many Forms

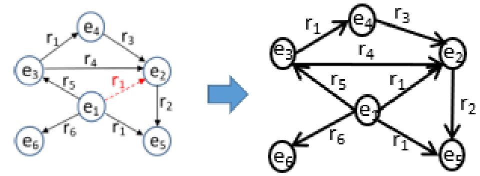
Static graph



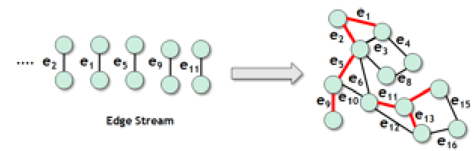
Dynamic / temporal / time-evolving graph



Historic / temporal snapshot graph

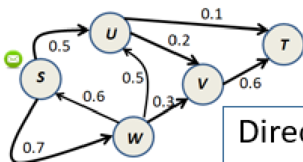


Dynamic / incremental update

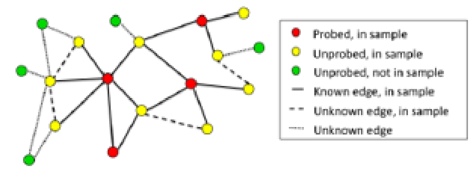


Graph stream

Uncertain graph

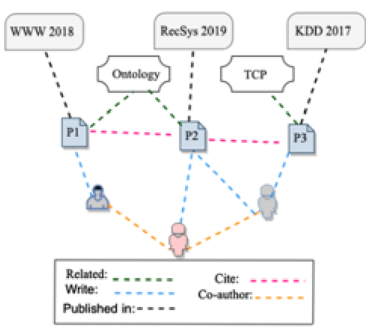


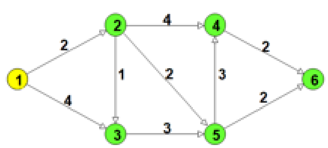
Incomplete / partially observed graph

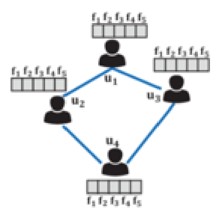


Directed/ weighted/ attributed graph

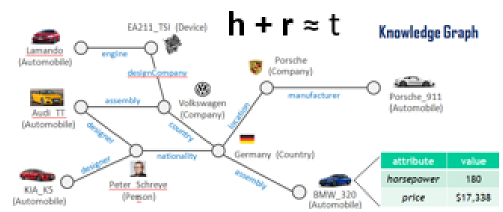
Heterogeneous graph



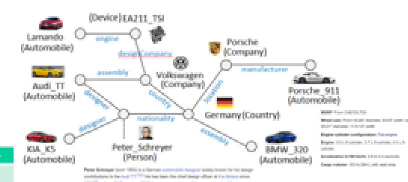




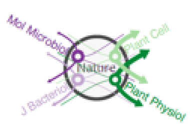
Knowledge graph



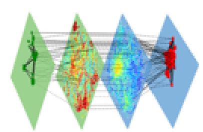
Multi-modal graph



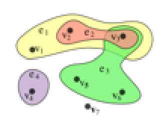
Higher-order graph



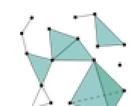
Higher-order interaction



Multilayer network

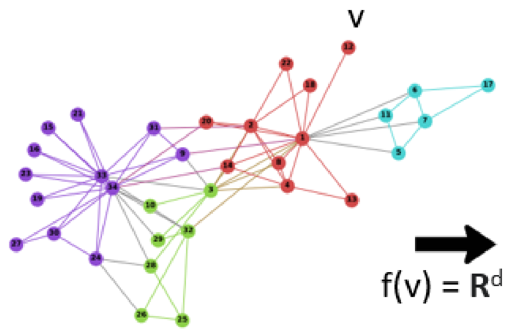


Hypergraph

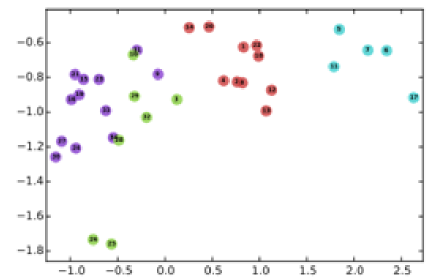


Simplicial complex

Graph Neural Network (GNN): Key Idea



f(v) = R^d



Representation learning or embedding

Downstream tasks (e.g., graph and node classification, recommendation, link prediction, question answering)

Learning could be end-to-end

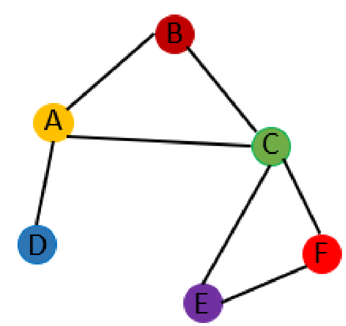
- End-to-end learning → ~~Feature Engineering~~
- Task-independent / task-dependent learning.
- Can capture graph structure and node, edge features.

Graph Convolutional Neural Network (GCN)

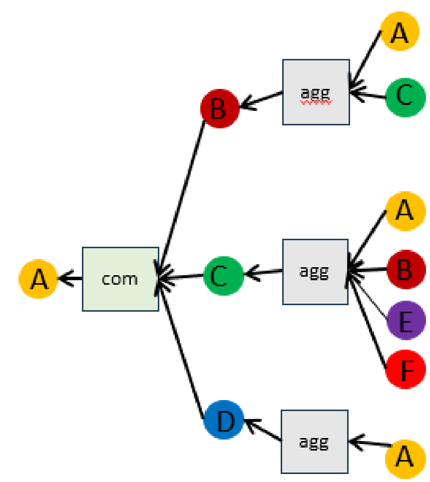
- Massage passing to use aggregation and combine functions repeated several times.

$$H^{(t+1)} = \sigma(\widetilde{D}^{-\frac{1}{2}} . \widetilde{A} . \widetilde{D}^{-\frac{1}{2}} . H^{(t)} . W^{(t)})$$

$\widetilde{A} = A + I_N$ $\quad H^{(0)} = X$ Input Node Features



Input graph



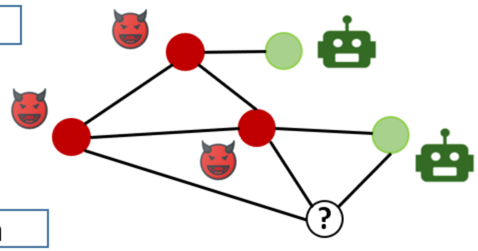
Technique in GCN

Representation Learning on Networks (WWW Tutorial, 2018)

Graph Neural Network (GNN): Downstream Tasks

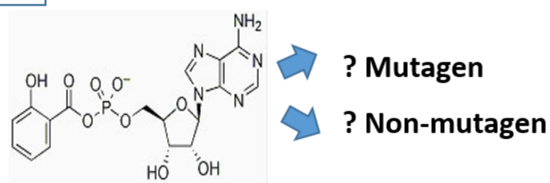
Node classification

Node-level task



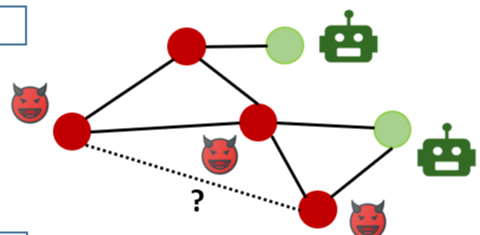
Graph classification

Graph-level task



Link prediction

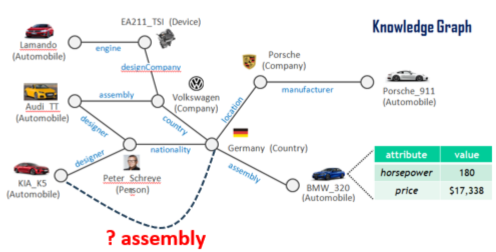
Edge-level task



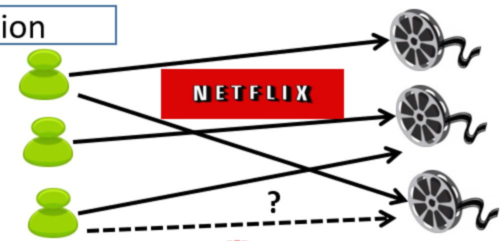
Triple classification

Edge-level task

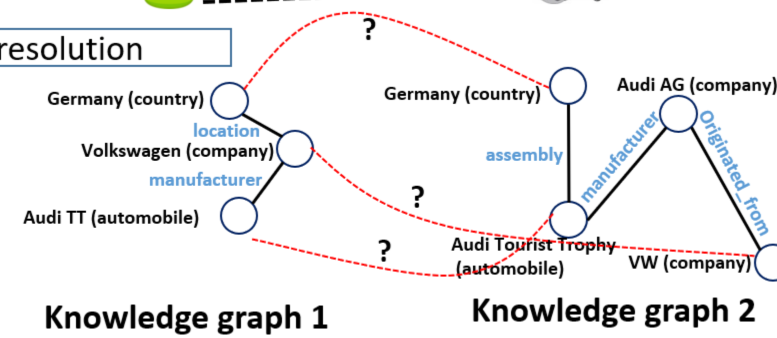
A Comprehensive Survey on Graph Neural Networks. IEEE Trans. Neural Networks Learn. Syst. 2021.



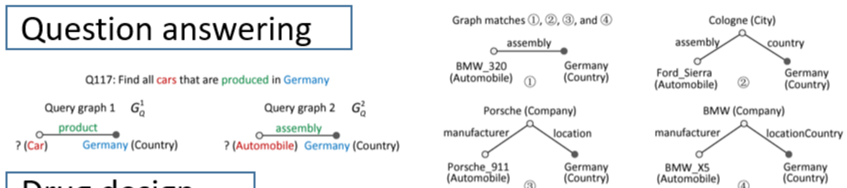
Recommendation



Entity resolution



Question answering



Drug design

- predicting missing links between drug and disease.

Synergy between Graph Data Management and Graph Machine Learning

Graph data construction and cleaning → Graph embedding → GNN training → Downstream tasks → Explainability

Graph machine learning and data science pipeline

- **How graph machine learning and graph data management benefit each other?**

- **Application of graph data management in graph machine learning**

  - Scalable Graph Embedding

  [PVLDB'23] *Distributed Graph Embedding with Information-Oriented Random Walks*

  - Dynamic updates in knowledge graph embedding

  [KBS'22] *Efficiently Embedding Dynamic Knowledge Graphs*

  ✓ - GNN Explainability

  [SIGMOD'24] *View-based Explanations for Graph Neural Networks*

  [ICDE'24] *Generating Robust Counterfactual Witnesses for Graph Neural Networks*

- **Application of graph machine learning in graph data management**

  - Knowledge graph question answering

  [ICDE'22] *Aggregate Queries on Knowledge Graphs: Fast Approximation with Semantic-aware Sampling*

  [ICDE'20] *Semantic Guided and Response Times Bounded Top-k Similarity Search over Knowledge Graphs*

Our Work in Graph Data Management and Machine Learning

- ✓ Big-Graphs: Querying, Mining, Streaming, Uncertainty, Systems, and Beyond
- ✓ User-friendly, efficient, and approximate querying and pattern mining using scalable algorithms, machine learning techniques, and distributed systems

Scalable, Approximate Graph Querying

Knowledge-Graph Search [SIGMOD 11, VLDB 13, ICDE 12 Tutorial, ACM SIGMOD Blog Post, Encyclopaedia of Big Data Technologies, ICDE 20, 22, KBS 22, CIKM 22, SIGMOD Record 23]

Graph Query-By-Example [ICDE 14, TKDE 15]

Graph Stream Summarization and Query [SNAM 17]

Spatial/ Road Network Query [TKDE 20, SIGMOD 20]

Uncertain Graph Processing

Reliability and Shortest Path [EDBT 14, VLDB 15 Tutorial, TKDE 18, 20, Book @Morgan&Claypool, VLDB 19, VLDB 21, TKDE 22]

Influence Maximization and Embedding [SDM 11, ICDE 16, CIKM 17, ICDE 18, SIGMOD 18, ICDE 23]

Densest subgraph [ICDE 23]

Complex Graph Mining

Novel Pattern Mining [SIGMOD 10, VLDB 20]

Graph Anomaly Detection [CIKM 12]

Graph-based Entity Resolution [CIKM 17, VLDB 18]

Graph Summary, Core Decomposition, Multi-layer graph, and Hypergraph [SIGMOD 19, VLDB 17 Tutorial, TKDD 22, VLDB 23]

Distributed Graph Systems

Information-centric Distributed Graph Embedding [VLDB 23]

Complementary Graph Partitioning [SIGMOD 12]

Graph Storage Decoupling and Smart Query Routing [USENIX ATC 18]

Vertex-Centric Graph Processing [VLDB 14 Tutorial, EDBT 17]

Graph Machine Learning ✓

Graph Neural Network Explainability [BigData 20, DSAA Tutorial 23, SIGMOD 24, ICDE 24]

Other Big Data Processing

Relational Data [SIGMOD 14]

Stream Data [SIGMOD 16]

Blockchain Network [WebConf 20, 21, WSDM 22, CIKM 22 Tutorial]

Crowd-Sourcing [CIKM 17, VLDB 18]

Graph Neural Network (GNN): Explainability [IEEE DSAA 2023 Tutorial]

- Explain the results of high-quality GNNs.
- **[Instance-level]** Understand which aspects of the input data drive the decisions of the GNN – discover critical nodes, edges, subgraphs, and their features that are responsible for GNN outcomes.
- **[Model-level]** Insight on how GNNs work – discover what input subgraph patterns lead to a certain prediction.

Importance

- Desirable to understand and explain the workings and results of black-box GNNs – bridge domain knowledge with GNN predictions, human-AI collaboration.
- Safety and well-being (e.g., autonomous car, AI in healthcare) – trust in deep learning models.
- Understand bias in machine learning (ML) algorithms – ML algorithms can amplify bias, model debugging.
- Robustness against adversarial examples – improve quality of GNN outputs.
- Legal requirements, e.g., GDPR – algorithms to explain their outputs.

Stakeholders

End users, domain experts, decision makers, policy makers, regulatory agencies, researchers, data scientists, and engineers

Challenges with GNN Explanations

[IEEE DSAA 2023 Tutorial]

- Many definitions, motivations, and requirements for explainability.

– trust, causality, transferability, informativeness, fair and ethical decision making, model debugging, recourse, mental model comparison, context-dependent, low-level mechanistic understanding of models, high-level human understanding, what makes users confident about the model.

- Comparing explanations is hard!

- Several quantitative and qualitative evaluation metrics or methods.

– **Quantitative:** faithfulness (fidelity+, fidelity-), sparsity, contrastivity, accuracy, stability.

– **Qualitative:** application-grounded, human-grounded, and functionally-grounded evaluation.

- Difficult to obtain ground-truth.

- **Other issues:** Evaluation via occlusion creates data outside training distribution, bias terms, redundant evidence, trivial correct explanations, weak GNN model, misaligned GNN architecture, problems due to graph data vs. grid data.

- Capture interplay of graph structure and features in GNN's decision making.

Challenges with GNN Explanations

[IEEE DSAA 2023 Tutorial]

• Many definitions, motivations, and requirements for explainabili

– trust, causality, transferability, informativeness, fair and ethi ging, recourse, mental model comparison, context-dependen models, high-level human understanding, what makes user

• Comparing explanations is hard!

• Several quantitative

– **Quantitati** ity, accuracy, stability.

– **Qu** , and functionally-grounded evaluation.

• Diffi

• **Other i** sion creates data outside training distribution, bias terms, redundant orrect explanations, weak GNN model, misaligned GNN architecture, problems d data vs. grid data.

• Capture interplay of graph structure and features in GNN's decision making.

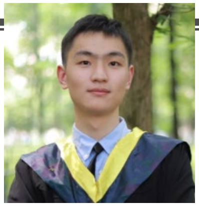
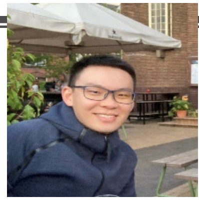
- **"Ability to explain or to present a model in understandable terms to humans"**

- Doshi-Velez and Kim 2017

Graph Neural Network (GNN) Explainability: Our Work

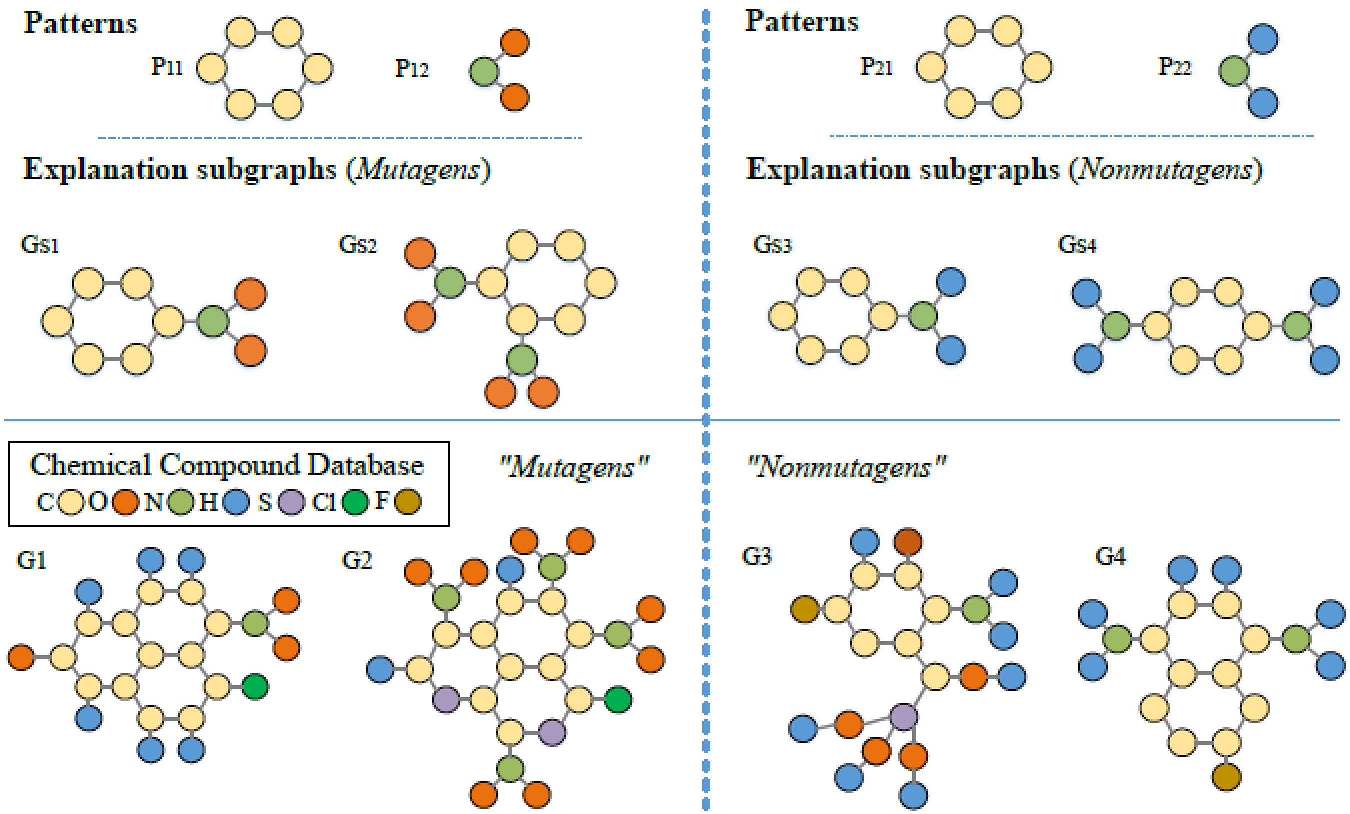
- **Benchmarking of GNN explainability methods** (IEEE BigData 2020, IEEE DSAA 2023 Tutorial)
- **GNN explanation usability** (SIGMOD 2024)
- **GNN explanation robustness** (ICDE 2024)
- **GNN counterfactual evidence** (under submission)
- **GNN explanation for model debugging** (under submission)
- **GNN explanation skyline / pareto optimality** (under submission)

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]

Tingyang Chen, Zhejiang University

Dazhuo Qiu, Aalborg University



GNN-based drug classification, with graph patterns and induced subgraphs that help understand the results: ''*which toxicophore occurs in mutagens*?''

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]

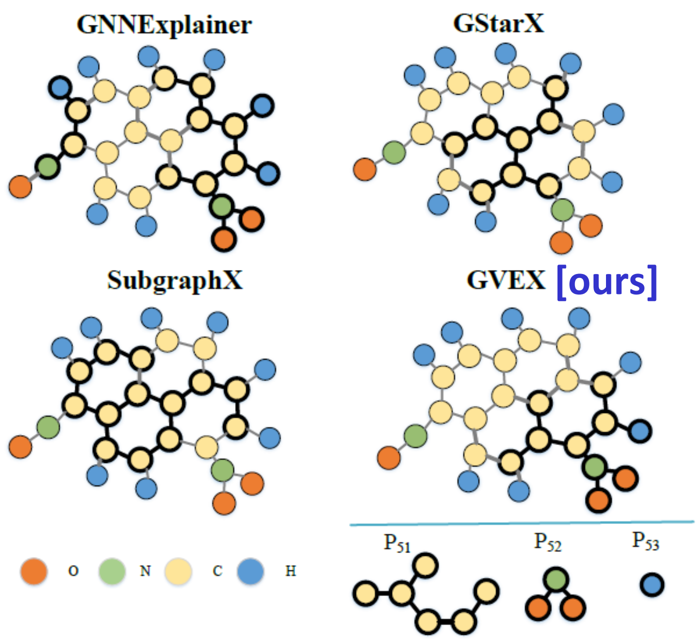
Challenges with Existing Approaches

Oversized explanation: Existing methods generate large explanation subgraphs.

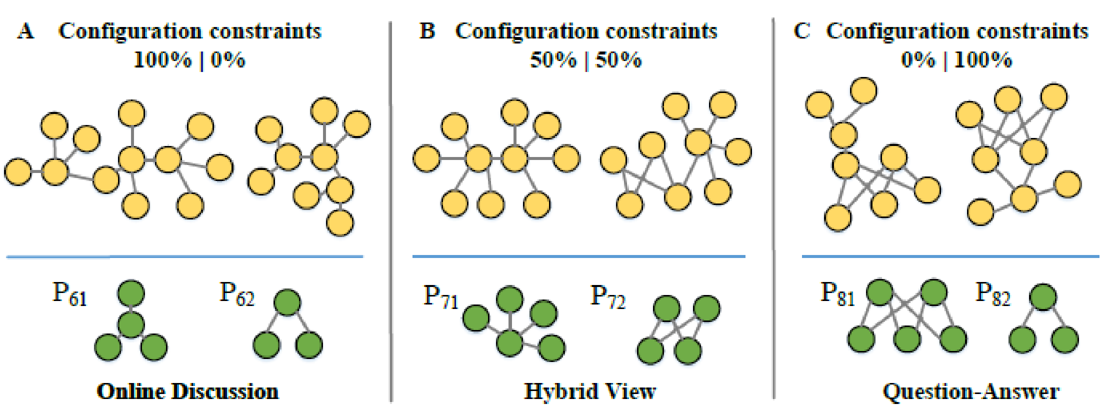
Lack of meaningful explanations for domain experts: Due the oversized explanations containing irrelevant/ repeated structures, it is hard to identify meaningful structures. Not queryable, hence not easy to access and inspect with domain knowledge.

Not configurable explanation based on user setting: Only explaining one class may omit the relevant information between classes.

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]

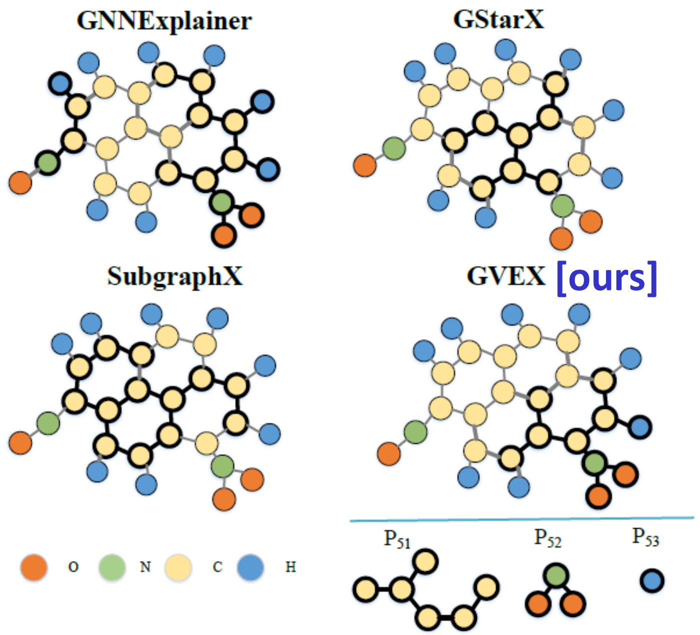


GNN-based drug classification



GNN-based social analysis: REDDIT-BINARY social network dataset, two classes - online-discussion threads and question-answer threads. Three different configuration scenarios.

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]

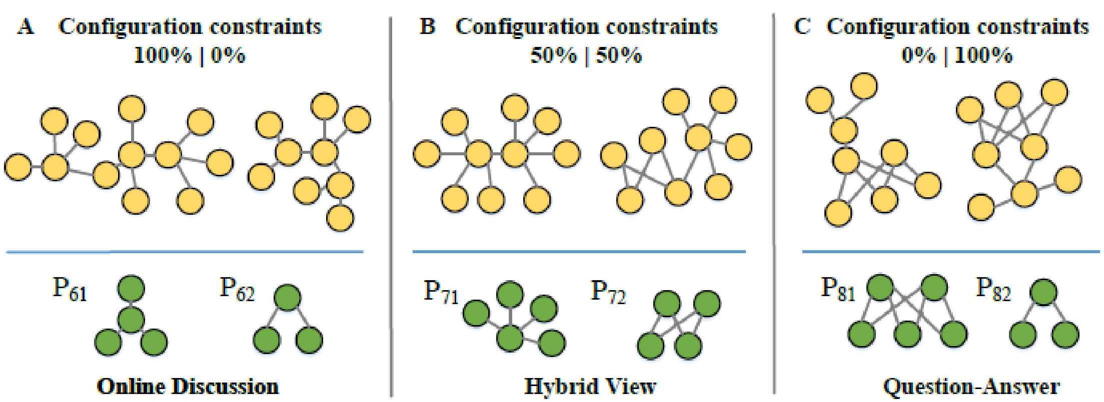


GNN-based drug classification

Two-tier Explanation

Lower-tier: An explanation subgraph that ensures same prediction (factual) and its removal changes G's prediction label (counterfactual).

Higher-tier: A set of graph patterns that summarizes the explanation subgraphs with coverage guarantee.



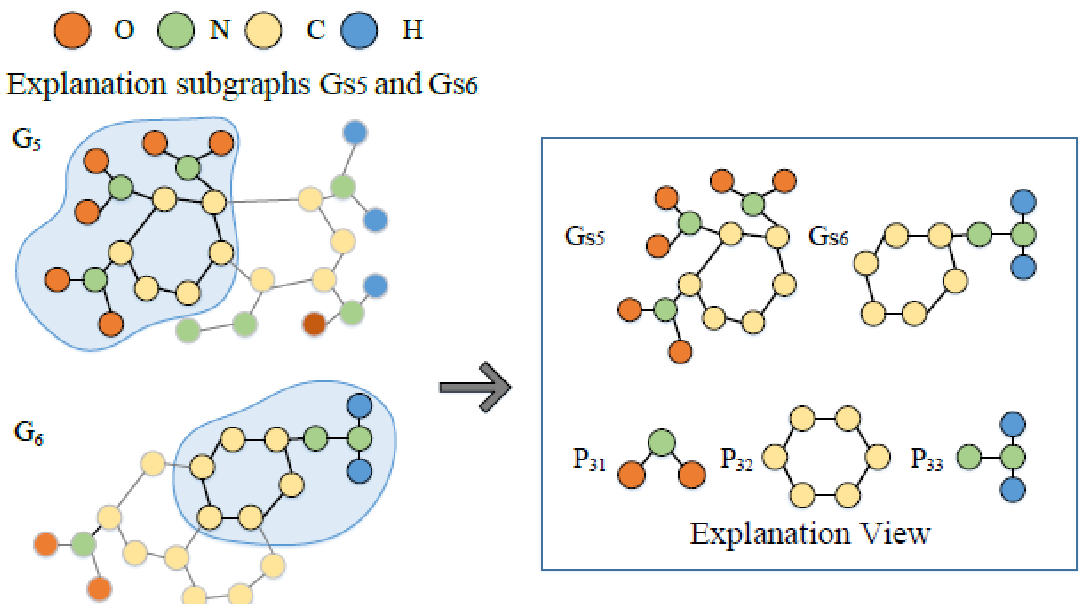
GNN-based social analysis: REDDIT-BINARY social network dataset, two classes - online-discussion threads and question-answer threads. Three different configuration scenarios.

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]

Explanation Subgraphs: Given a GNN M and a single graph G with label $M(G) = l$, an explanation subgraph G_s^l of G satisfies:

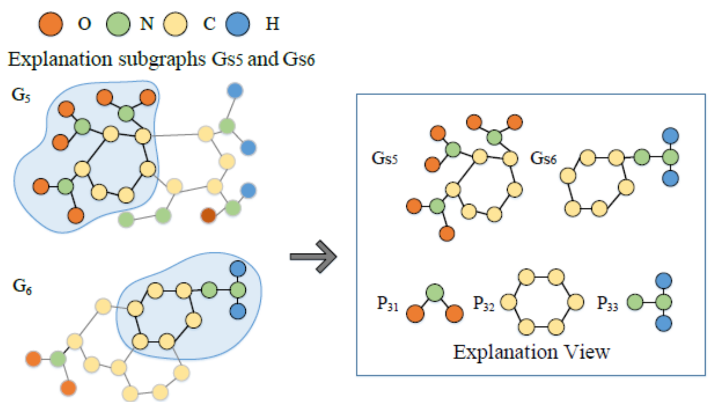
$$M(G) = M\left(G_s^l\right) = l \text{ and } M\left(G \backslash G_s^l\right) \neq l$$

Explanation Views: Given a graph database $\mathcal{G}$, a classifier M, and label l, an explanation view consists of **(1)**$\mathcal{G}_s^l$, a set of explanation subgraphs for the label group $\mathcal{G}^l$ and **(2)**$\mathcal{P}^l$, a set of patterns that cover the nodes of the explanation subgraphs.



An explanation view for a single class label: explanation subgraphs and patterns

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]



An explanation view for a single class label: explanation subgraphs and patterns

Quality of Explanation Views

Explainability: An explanation view has better explainability if its explanation subgraphs involve more nodes with features that can maximize their influence via a random walk-based message passing process of GNN.

$$f(\mathcal{G}^l_{\mathcal{V}}) = \sum_{G_{si} \in \mathcal{G}^l_s} \frac{I(V_{si}) + \gamma D(V_{si})}{|V_i|}$$

feature influence

neighbor diversity

Coverage: The set of explanation subgraphs $\mathcal{G}^l_s$ contains total n nodes, where $n \in [b_l, u_l]$, specified by the coverage constraint for label group $\mathcal{G}^l$.

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]

Explanation View Generation Problem

PROBLEM 1. *Given a graph database $\mathcal{G}$, a set of interested labels* $Ł$ *s.t.* $|Ł| = t$, *a* GNN $\mathcal{M}$, *and a configuration* C, *the* explanation view generation problem, *denoted as* EVG, *is to compute a set of graph views* $\mathcal{G}_{\mathcal{V}} = \{\mathcal{G}_{\mathcal{V}}^{l_1}, \ldots \mathcal{G}_{\mathcal{V}}^{l_t}\}$, *such that* $(i \in [1, t])$*:*

- *Each graph view* $\mathcal{G}_{\mathcal{V}}^{l_i} = (\mathcal{P}^{l_i}, \mathcal{G}_s^{l_i}) \in \mathcal{G}_{\mathcal{V}}$ *is an explanation view of* $\mathcal{G}$ *for* $\mathcal{M}$ *w.r.t.* $l_i \in Ł$;
- *Each* $\mathcal{G}_{\mathcal{V}}^{l_i}$ *properly covers the label group* $\mathcal{G}^{l_i}$; *and*
- $\mathcal{G}_{\mathcal{V}}$ *maximizes an aggregated explainability, i.e.,*

$$\mathcal{G}_{\mathcal{V}} = \arg\max \sum_{\mathcal{G}_{\mathcal{V}}^{l_i} \in \mathcal{G}_{\mathcal{V}}} f(\mathcal{G}_{\mathcal{V}}^{l_i})$$

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]

Explanation View Generation Problem: Hardness and Properties

View Verification

Given a graph database $\mathcal{G}$, configuration C, and a two-tier explanation structure $(\mathcal{P}, \mathcal{G}_s)$, the view verification problem is NP-complete when the GNN is fixed.

Explanation View Generation (EVG) Problem

For a fixed GNN, EVG is (1) $\sum_p^2 -complete$, and (2) remains NP-hard even when $\mathcal{G}$ has no edges.

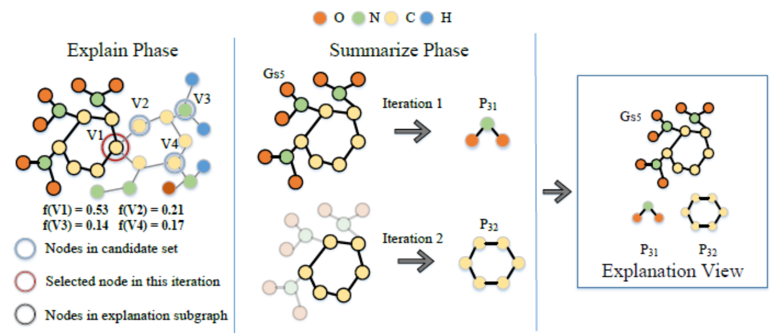
Monotone Submodularity

Given $\mathcal{G}, L, C$ and a fixed GNN, $f(\mathcal{G}_V)$ is a monotone submodular function of V_S.

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]

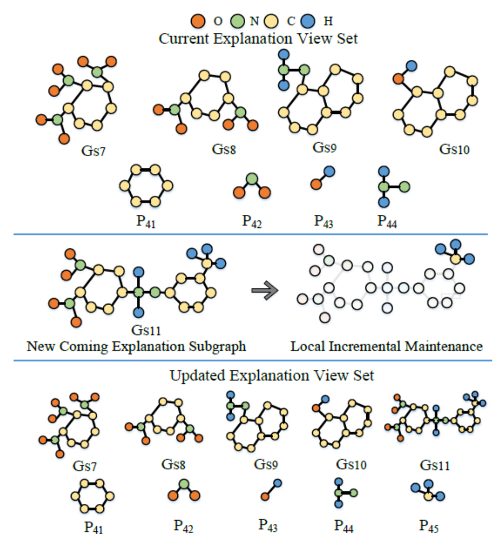
Approximation Algorithms and Quality Guarantees

Greedy, Approximation Algorithm: Given a configuration C, graph database $\mathcal{G}$, and a k-layer GNN over label set L, there is a $\frac{1}{2}$-approximate algorithm for generating explanation views.



Greedy, approximation algorithm

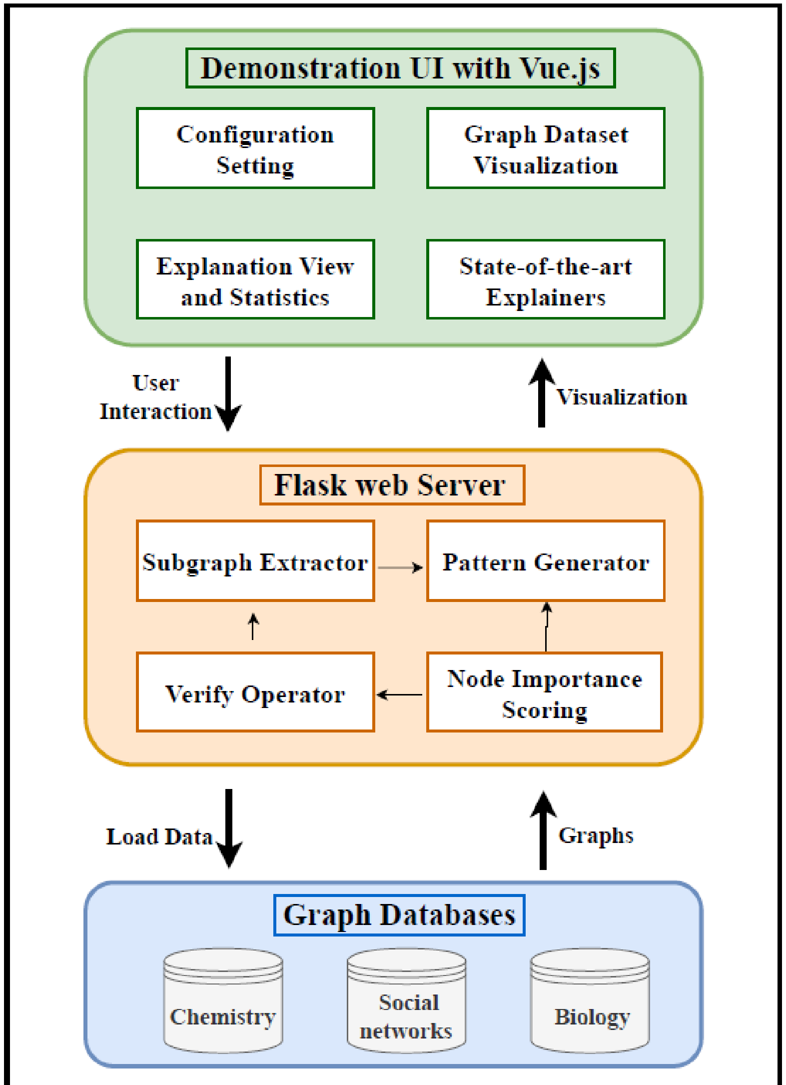
Streaming Algorithm: Given a configuration C, graph database $\mathcal{G}$, and a k-layer GNN, there is an online algorithm that maintains explanation views with a a $\frac{1}{4}$-approximation.



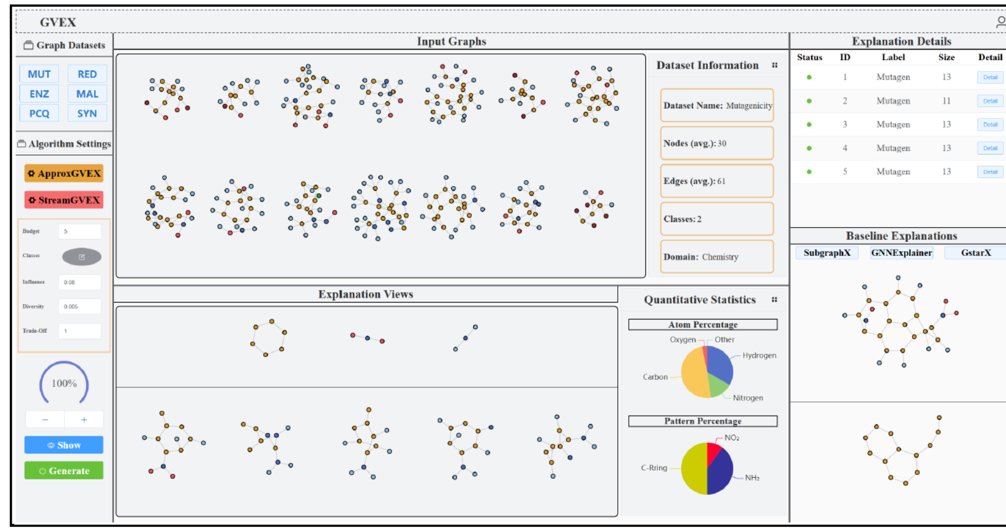
Streaming algorithm

Parallelization: Both algorithms can be parallelized.

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024 Demo]



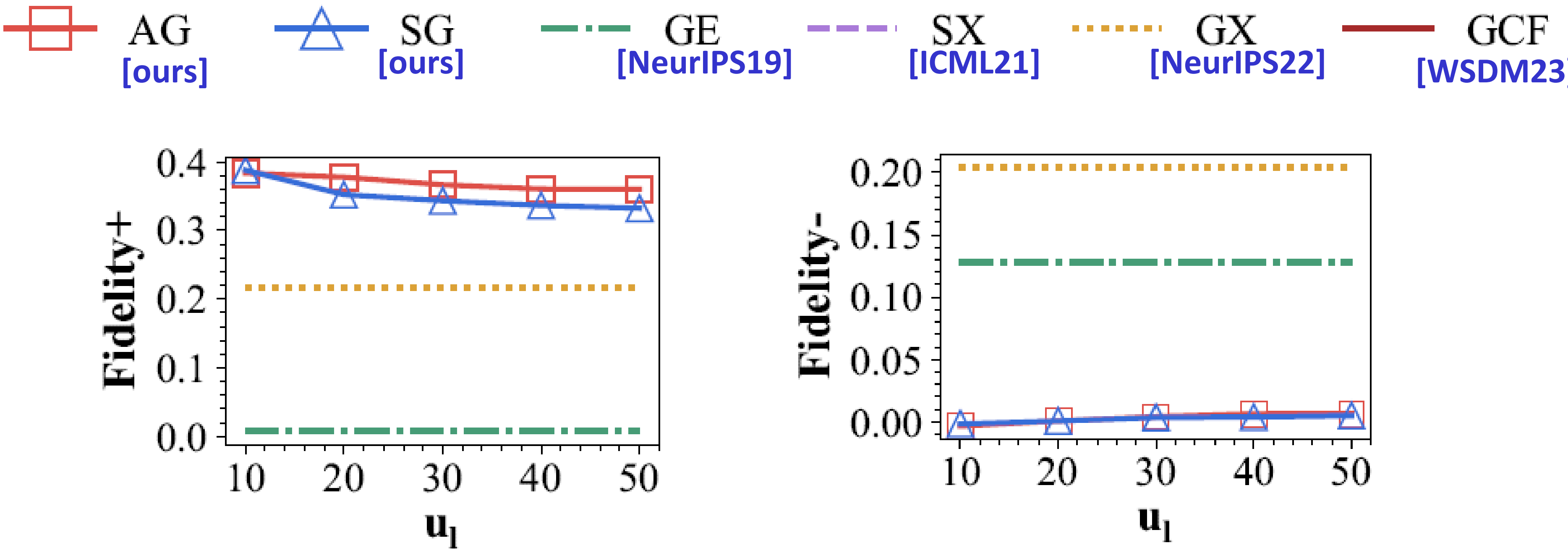
The architecture of the GVEX system



A screenshot of the GVEX frontend

https://youtu.be/q9d7ldulIuQ

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]

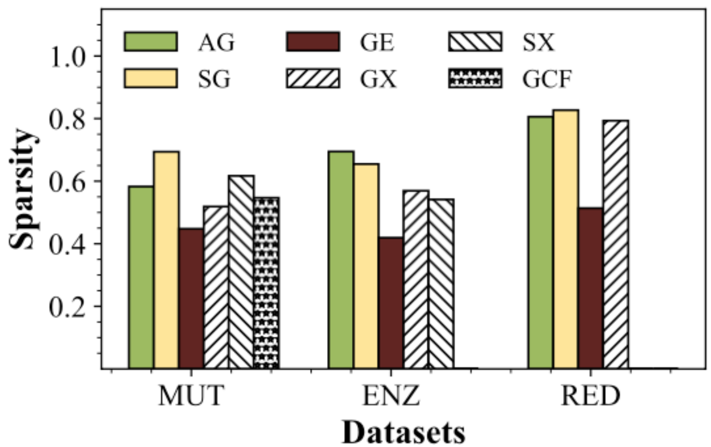


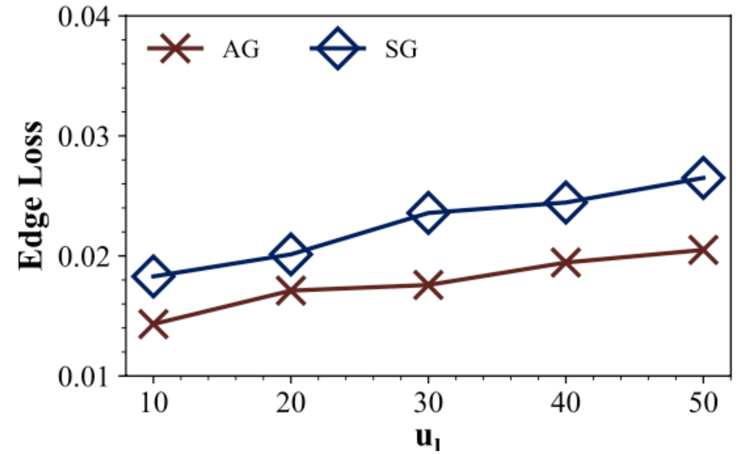
Fidelity+ quantifies the deviations caused by removing the explanation substructure from the input graph. Higher is better.

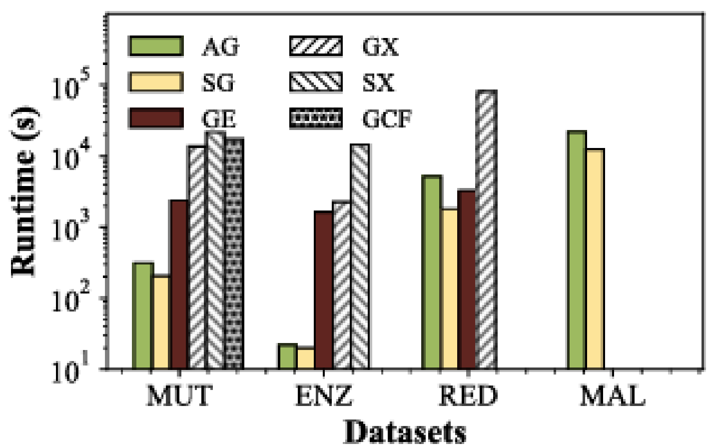
Fidelity- measures how close the prediction results of the explanation substructures are to the original inputs. Lower is better.

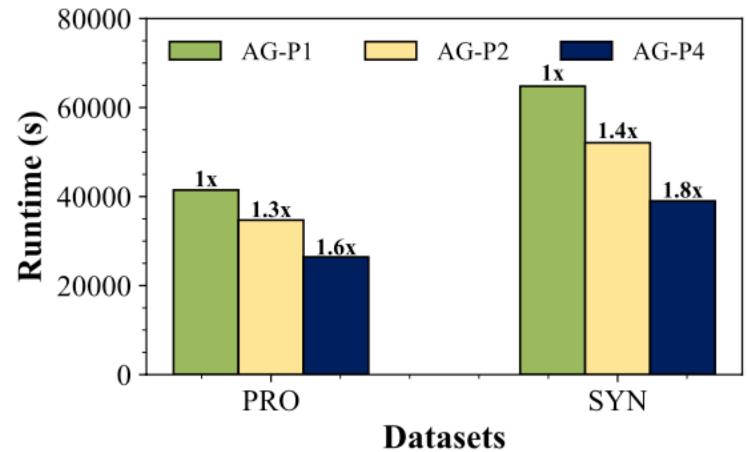
REDDIT-BINARY dataset.

GNN Explanation Usability: Graph View-based Explanation [SIGMOD 2024]

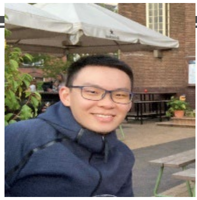






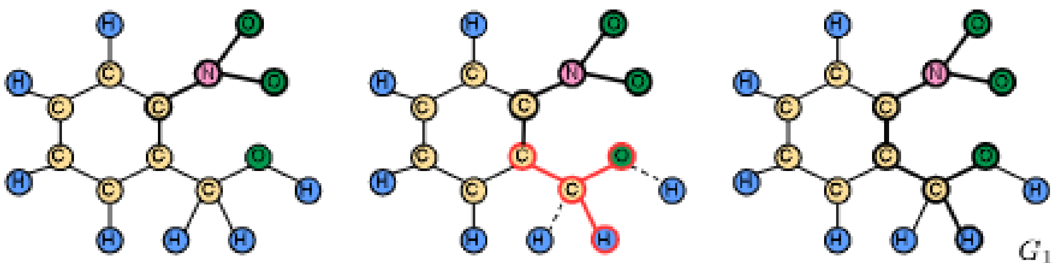


GNN Explanation Robustness [ICDE 2024]




Dazhuo Qiu, Aalborg University

Mengying Wang, Case Western Reserve University



G1: Left-Bold nodes and edges indicate a counterfactual. **Middle-**Red bold nodes and edges indicate a new counterfactual after deleting dotted edges. **Right**-A counterfactual robust to graph edits.

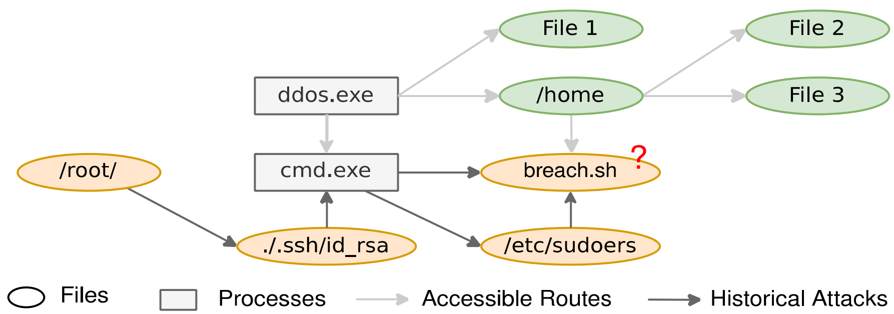
(Factual) Witness: A subgraph G_w that satisfies: $M(v, G) = M(v, G_w) = l$

Counterfactual Witness: A witness that satisfies: $M(v, G \backslash G_w) \neq l$

k - Robust Counterfactual Witness: By removing k edges from input graph G, G_w remains a counterfactual witness.

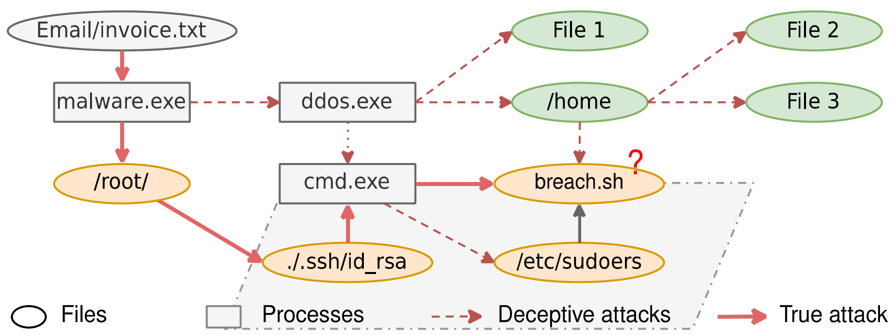
GNN Explanation Robustness [ICDE 2024]

Example - "Vulnerable Zone" in Cyber Networks



🔒 ***GNN-based Security System:***

- **Detection**: Train GNN based on historical attacks to classify files' vulnerability.
- **Protection**: Enhanced security for vulnerable files (colored orange).



☠️ ***Multi-Phase Cyber Attack Strategy:***

- **Phase 1: Deception Attacks**: Conduct deceptive but harmless attacks to Induce false invulnerable classification on target.
- **Phase 2: True Attack**: attack by exploiting reduced defenses on target.

> 💡 *How can we identify a "Vulnerable Zone" within cyber networks where, if protected, GNN predictions remain solid, even if other parts of the network are disturbed by deceptive attacks?*

GNN Explanation Robustness [ICDE 2024]

Robust Explanation Generation Problem: Hardness and Solution

- ***Factual Explanation (Witness)***:
  - $M(v, G) = M(v, G_s) = l$
- ***Counterfactual Explanation (CW)***:
  - $M(v, G) \neq M(v, G \setminus G_s) \neq l$
- ***Robust Explanation (k-RCW)***:
  - G_s remains consistent under disturbance.

We are the first to consider <u>all three criteria</u>! 😎

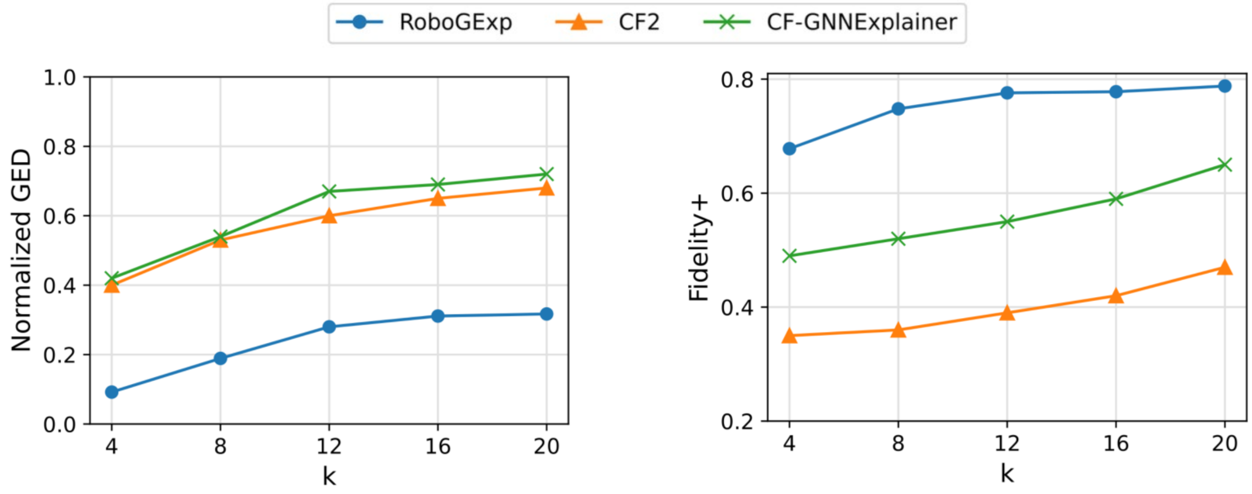
- ***Verification Problem***: Given G_s, decide if G_s is a k-RCW for a set of test nodes V_t, w.r.t a model M.
  - Witness verification ☞ <u>PTIME</u>.
  - CW verification ☞ <u>PTIME</u>.
  - $\boldsymbol{k}$-RCW verification ☞ <u>NP-hard</u>.
- ***Generation Problem***: Given a graph G and V_t, compute a $\boldsymbol{k}$-RCW if exists.
  - $\boldsymbol{k}$-RCW generation in general ☞ <u>co-NP-hard</u>
  - under ($\boldsymbol{k}$, **b**)-disturbances ☞ <u>PTIME</u>.

- We propose effective and efficient solution under (k, b)-disturbance and APPNP GNN. k: global budget, b: local budget. APPNP GNN (ICLR 2019).
- Parallelization scheme with graph partition.

GNN Explanation Robustness [ICDE 2024]

Experiment Results: Effectiveness

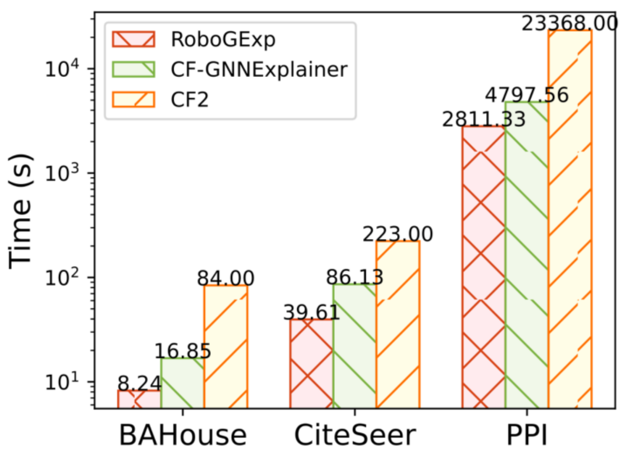
Methods	Counterfactual	Factual	Robustness
CF-GNNExp (AISTATS 2022)	✓		
CF^2 (WWW 2022)	✓	✓	
RoboGExp	✓	✓	✓

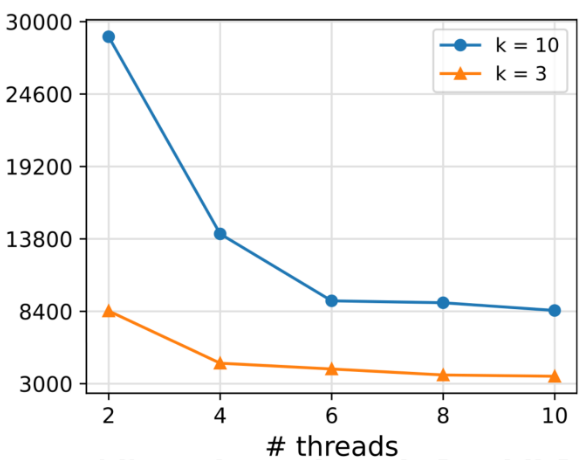


GNN Explanation Robustness [ICDE 2024]

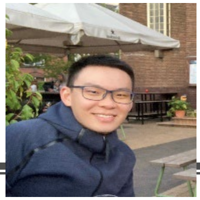
Experiment Results: Efficiency & Scalability

Methods	Counterfactual	Factual	Robustness
CF-GNNExp (AISTATS 2022)	✓		
CF^2 (WWW 2022)	✓	✓	
RoboGExp	✓	✓	✓

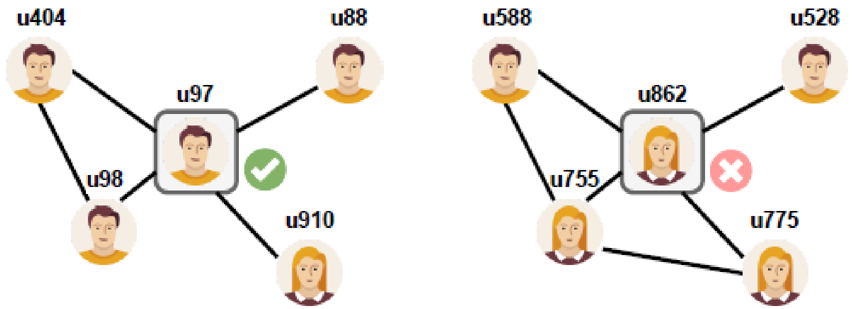




GNN Counterfactual Evidence [under submission]



Dazhuo Qiu, Aalborg University

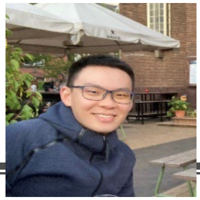


User	Gender	Age	Loan Duration (Month)	Loan Amount (Deutsch Mark)
u97	***Male***	***30-40***	***12-24***	***2k-3k***
u98	Male	30-40	36-48	2k-3k
u404	Male	40-50	12-24	2k-3k
u88	Male	30-40	12-24	3k-4k
u910	Female	30-40	36-48	4k-5k

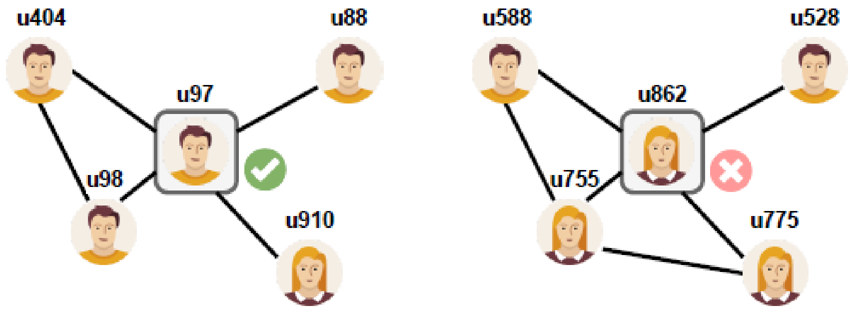
User	Gender	Age	Loan Duration (Month)	Loan Amount (Deutsch Mark)
u862	Female	***30-40***	***24-36***	***2k-3k***
u755	Female	30-40	24-36	1k-2k
u588	Male	40-50	12-24	1k-2k
u528	Male	30-40	36-48	2k-3k
u775	Female	18-30	24-36	1k-2k

Figure 1: Two customers $u97$ and $u862$ from the *German Credit* dataset [3] are shown at the centers. We present their 1-hop neighborhood structures, along with feature values for all these nodes. Three features, namely *age*, *loan duration*, and *loan amount* are relevant w.r.t. node classification, i.e., whether their loans would be approved or not; whereas *gender* is a sensitive feature. Notice that $u97$ and $u862$ share similar node feature values for *age*, *loan duration*, and *loan amount*. They also share similar 1-hop neighborhood structures with only one edge difference. Notably, $u97$ and $u862$ have different genders. A pre-trained GNN predicts different classes for $u97$ and $u862$, making each of them a counterfactual evidence of the other.

GNN Counterfactual Evidence [under submission]



Dazhuo Qiu, Aalborg University



User	Gender	Age	Loan Duration (Month)	Loan Amount (Deutsch Mark)
u97	***Male***	***30-40***	***12-24***	***2k-3k***
u98	Male	30-40	36-48	2k-3k
u404	Male	40-50	12-24	2k-3k
u88	Male	30-40	12-24	3k-4k
u910	Female	30-40	36-48	4k-5k

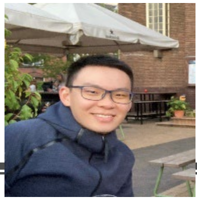
User	Gender	Age	Loan Duration (Month)	Loan Amount (Deutsch Mark)
u862	***Female***	***30-40***	***24-36***	***2k-3k***
u755	Female	30-40	24-36	1k-2k
u588	Male	40-50	12-24	1k-2k
u528	Male	30-40	36-48	2k-3k
u775	Female	18-30	24-36	1k-2k

Figure 1: Two customers $u97$ and $u862$ from the *German Credit* dataset [3] are shown at the centers. We present their 1-hop neighborhood structures, along with feature values for all these nodes. Three features, namely *age*, *loan duration*, and *loan amount* are relevant w.r.t. node classification, i.e., whether their loans would be approved or not; whereas *gender* is a sensitive feature. Notice that $u97$ and $u862$ share similar node feature values for *age*, *loan duration*, and *loan amount*. They also share similar 1-hop neighborhood structures with only one edge difference. Notably, $u97$ and $u862$ have different genders. A pre-trained GNN predicts different classes for $u97$ and $u862$, making each of them a counterfactual evidence of the other.

Counterfactual evidence for node classification

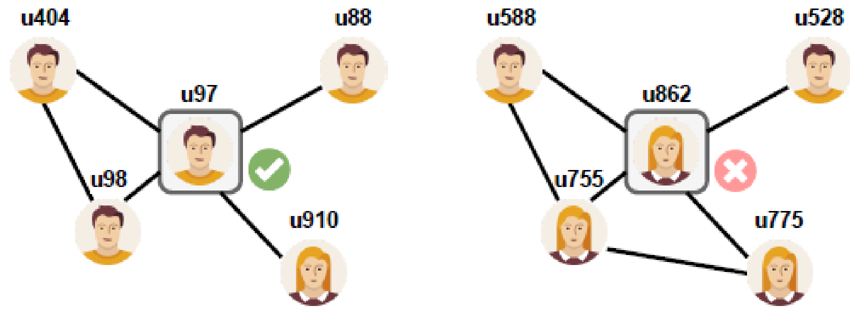
PROBLEM (TOP-1 LOCAL COUNTERFACTUAL EVIDENCE). *Given a query node $v \in V_{test}$, the top-1 counterfactual evidence, $\mathsf{LCE}_{opt}(v)$ is a node $u \in V_{test}$ that 1) has a different predicted label w.r.t. v; and 2) attains the highest similarity score $\mathsf{KS}(v,u)$ compared to all other nodes in the test set.*

$$\mathsf{LCE}_{opt}(v) = \underset{u \in V_{test},\ M(v) \neq M(u)}{\arg\max} \mathsf{KS}(v,u)$$



GNN Counterfactual Evidence [under submission]

Dazhuo Qiu, Aalborg University



User	Gender	Age	Loan Duration (Month)	Loan Amount (Deutsch Mark)
u97	***Male***	***30-40***	***12-24***	***2k-3k***
u98	Male	30-40	36-48	2k-3k
u404	Male	40-50	12-24	2k-3k
u88	Male	30-40	12-24	3k-4k
u910	Female	30-40	36-48	4k-5k

User	Gender	Age	Loan Duration (Month)	Loan Amount (Deutsch Mark)
u862	***Female***	***30-40***	***24-36***	***2k-3k***
u755	Female	30-40	24-36	1k-2k
u588	Male	40-50	12-24	1k-2k
u528	Male	30-40	36-48	2k-3k
u775	Female	18-30	24-36	1k-2k

Figure 1: Two customers $u97$ and $u862$ from the *German Credit* dataset [3] are shown at the centers. We present their 1-hop neighborhood structures, along with feature values for all these nodes. Three features, namely *age*, *loan duration*, and *loan amount* are relevant w.r.t. node classification, i.e., whether their loans would be approved or not; whereas *gender* is a sensitive feature. Notice that $u97$ and $u862$ share similar node feature values for *age*, *loan duration*, and *loan amount*. They also share similar 1-hop neighborhood structures with only one edge difference. Notably, $u97$ and $u862$ have different genders. A pre-trained GNN predicts different classes for $u97$ and $u862$, making each of them a counterfactual evidence of the other.

Counterfactual evidence for node classification

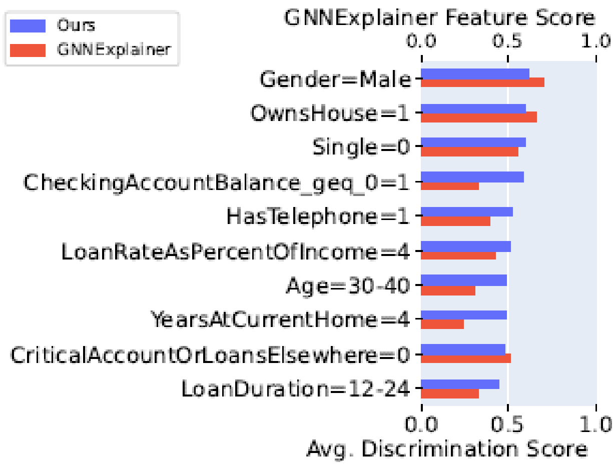
PROBLEM (TOP-1 LOCAL COUNTERFACTUAL EVIDENCE). *Given a query node $v \in V_{test}$, the top-1 counterfactual evidence, $\mathsf{LCE}_{opt}(v)$ is a node $u \in V_{test}$ that 1) has a different predicted label w.r.t. v; and 2) attains the highest similarity score $\mathsf{KS}(v, u)$ compared to all other nodes in the test set.*

$$\mathsf{LCE}_{opt}(v) = \underset{u \in V_{test},\ M(v) \neq M(u)}{\arg\max} \mathsf{KS}(v, u)$$

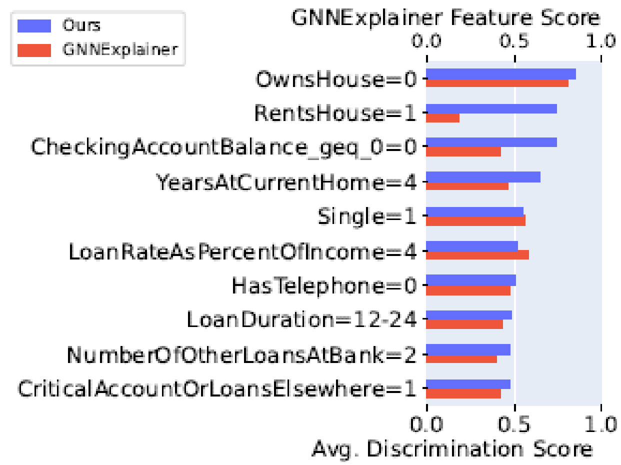
- We convert the counterfactual evidence finding problem to **nearest neighbor search problem** over the vector space.

- We demonstrate applications of our problem in:
  - GNN Explainability,
  - Revealing Unfairness of GNNs,
  - Verifying Prediction Errors,
  - Fine-tuning with Counterfactual Evidences.

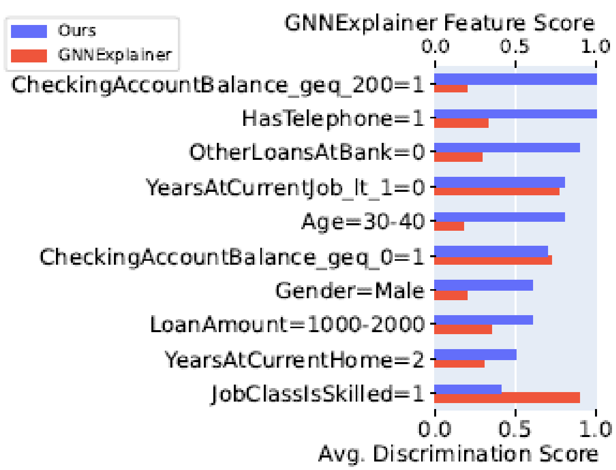
GNN Counterfactual Evidence [under submission]



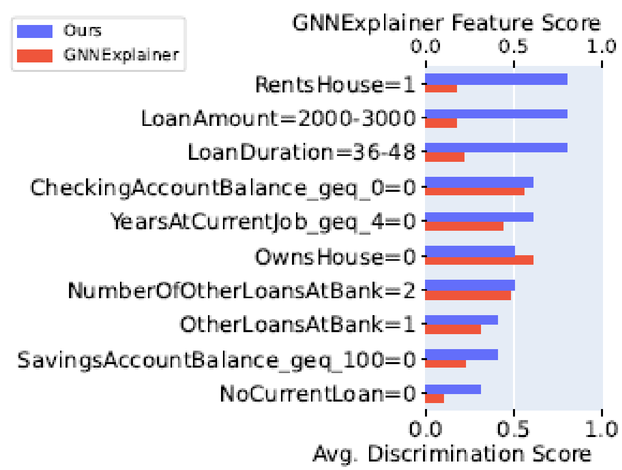
(a) $GCN, Class = 1$



(b) $GCN, Class = 0$



(c) $FairGNN, Class = 1$



(d) $FairGNN, Class = 0$

Loan Dataset: Feature importance across different GNN classifiers and classes. Class label =1 indicates predicted good customers by the GNN whose loans can be approved. Our findings reveal that "Gender = Male" and "Age = 30-40" are two critical factors based on which the classic GCN could predict a loan approval, whereas the gender bias is reduced when using the FairGNN (WSDM 2021).

Conclusion - GNN Explanability

- User-friendly, interactive, configurable, and robust explanation for graph neural networks (GNNs).

- Synergy between graph data management (e.g., graph view, usability, robustness, parallelization, fairness, vector search) and graph machine learning (e.g., GNN explanation).

Future Work

- Diversified GNN Explanations with Skylines/ Pareto Optimality

- GNN Explanations for Model Slicing and Debugging

- GNN Explanations beyond Classification (e.g., Graph Alignment)

Our Work in Graph Data Management and Machine Learning

✓ Big-Graphs: Querying, Mining, Streaming, Uncertainty, Systems, and Beyond

✓ User-friendly, efficient, and approximate querying and pattern mining using scalable algorithms, machine learning techniques, and distributed systems

✓

Knowledge-Graph Search [SIGMOD 11, VLDB 13, ICDE 12 Tutorial, ACM SIGMOD Blog Post, Encyclopaedia of Big Data Technologies, ICDE 20, 22, KBS 22, CIKM 22, SIGMOD Record 23]

Graph Query-By-Example [ICDE 14, TKDE 15]

Graph Stream Summarization and Query [SNAM 17]

Spatial/ Road Network Query [TKDE 20, SIGMOD 20]

Scalable, Approximate Graph Querying

Reliability and Shortest Path [EDBT 14, VLDB 15 Tutorial, TKDE 18, 20, Book @Morgan&Claypool, VLDB 19, VLDB 21, TKDE 22]

Influence Maximization and Embedding [SDM 11, ICDE 16, CIKM 17, ICDE 18, SIGMOD 18, ICDE 23]

Densest subgraph [ICDE 23]

Uncertain Graph Processing

Novel Pattern Mining [SIGMOD 10, VLDB 20]

Graph Anomaly Detection [CIKM 12]

Graph-based Entity Resolution [CIKM 17, VLDB 18]

Graph Summary, Core Decomposition, Multi-layer graph, and Hypergraph [SIGMOD 19, VLDB 17 Tutorial, TKDD 22, VLDB 23]

Complex Graph Mining

Information-centric Distributed Graph Embedding [VLDB 23]

Complementary Graph Partitioning [SIGMOD 12]

Graph Storage Decoupling and Smart Query Routing [USENIX ATC 18]

Vertex-Centric Graph Processing [VLDB 14 Tutorial, EDBT 17]

Distributed Graph Systems

Graph Neural Network Explainability [BigData 20, DSAA Tutorial 23, SIGMOD 24, ICDE 24]

Graph Machine Learning

Relational Data [SIGMOD 14]

Stream Data [SIGMOD 16]

Blockchain Network [WebConf 20, 21, WSDM 22, CIKM 22 Tutorial]

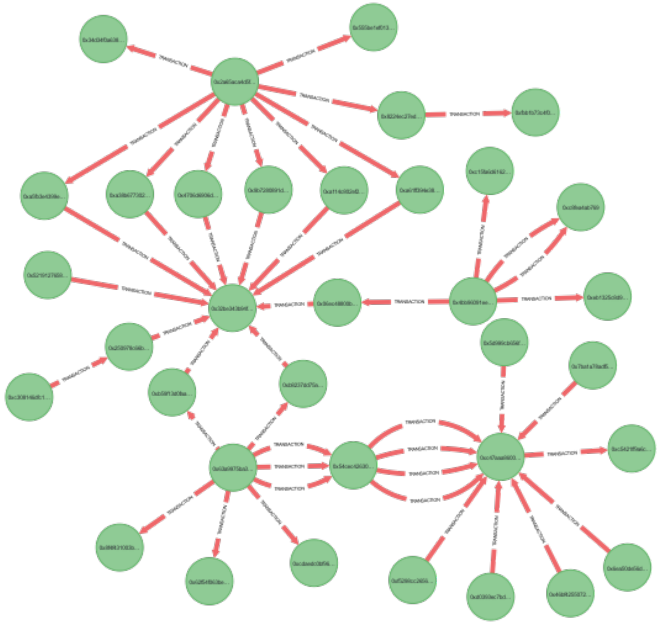
Crowd-Sourcing [CIKM 17, VLDB 18]

Other Big Data Processing

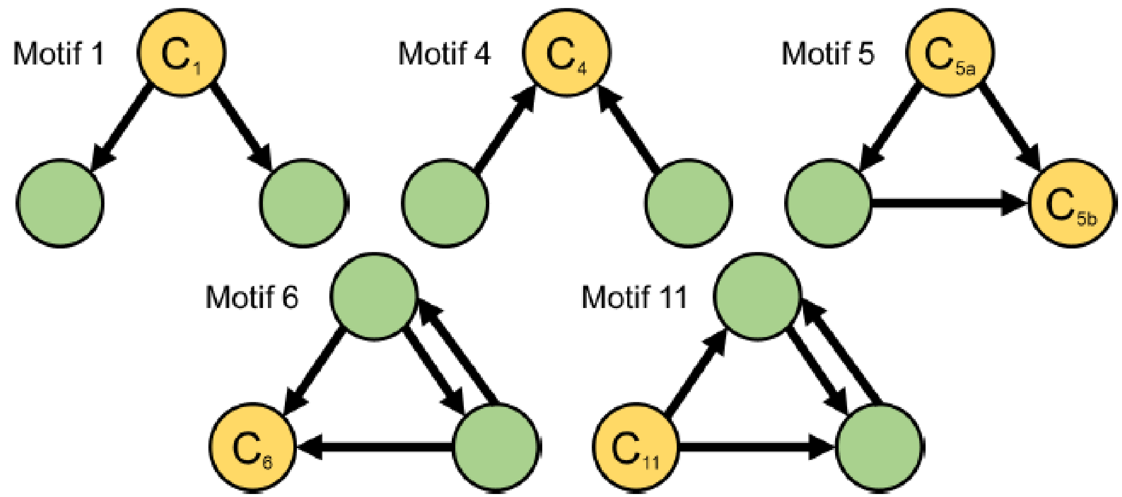
Approximate (Semantic) Subgraph Search for Knowledge Graphs Querying

- **NESS** (SIGMOD 2011), **NEMA** (PVLDB 2013) → *Neighborhood-based Fast, Approximate Graph Search*
- **GQBE** (ICDE 2014, TKDE 2015, ICDE 2016) → *Graph Query-By-Example [User-Friendliness]*
- **SGQ** (ICDE 2020), **AGQ** (ICDE 2022, CIKM 2022) → *Semantic-Guided and Response-Time Bounded Graph Search, Aggregate Queries on KG Embedding*
- **DKGE** (KBS 2022) → *Dynamic KG Embedding*
- **DistGER** (PVLDB 2023) → *Distributed Graph Embedding*
- **KG-QA** (SIGMOD Record 2023) → *Survey*
- **LLM Logic Consistency** (under submission) → *LLM + KG*

Subgraph Search: Market manipulators in LunaTerra StableCoin Collapse



Ethernet transaction network



Five three-node motifs exhibiting buy and sell behaviors. Nodes labeled C denote the center where a center with an in-degree = 2 indicates buy behavior and an outdegree = 2 indicates sell behavior.

Address/Motif Center	C_1	C_4	C_{5a}	C_{5b}	C_6	C_{11}
Celsius	-	81.2	-	78.9	-	-
hs0327.eth	88.2	67.0	96.2	69.7	81.6	-
Smart LP: 0x413	-	68.4	-	-	95.1	-
Token Millionaire 1	84.6	89.7	-	86.2	74.2	88.9
Token Millionaire 2	69.7	99.5	-	98.7	98.9	37.6
masknft.eth	90.9	90.7	-	82.1	93.3	92.2
Heavy Dex Trader	71.3	96.2	-	-	81.2	-
Oapital	91.6	78.9	60.5	58.5	71.8	92.5
Hodlnaut	39.9	98.9	-	90.6	99.4	-

Influential addresses found by our method that match with ground truth

Frontiers in Blockchain (2024)

Takeaway

Graph data construction and cleaning → Graph embedding → GNN training → Downstream tasks → Explainability

Graph machine learning and data science pipeline

- Synergy between graph data management and graph machine learning

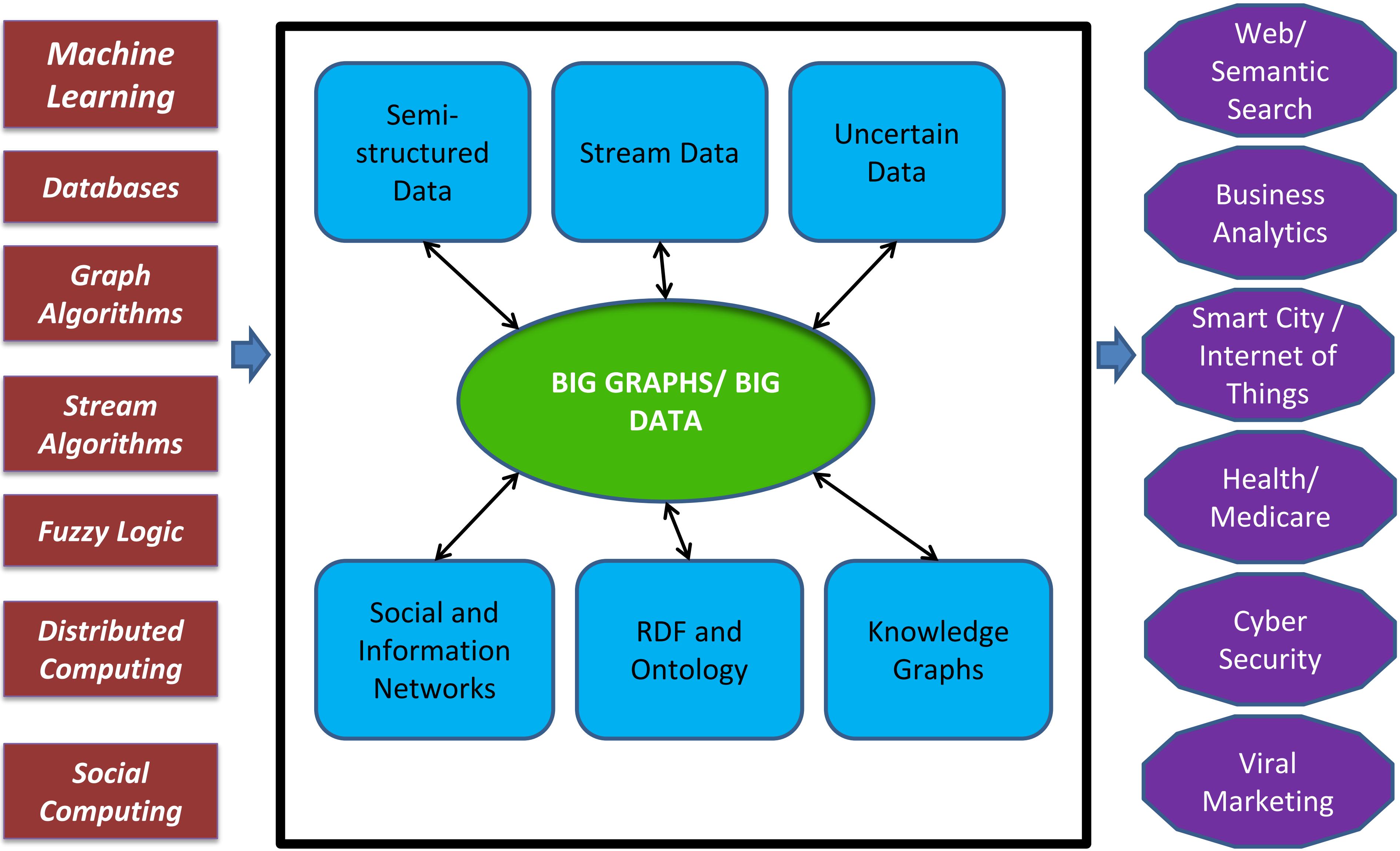
<u>Graph Machine Learning for Graph Data Management:</u>

- **Embedding based question answering**

<u>Graph Data Management for Graph Machine Learning:</u>

- **GNN Explanation: usability and robustness**
- **Scalable distributed graph embedding systems**

Big Picture



Machine Learning
Databases
Graph Algorithms
Stream Algorithms
Fuzzy Logic
Distributed Computing
Social Computing
Semi-structured Data
Stream Data
Uncertain Data
BIG GRAPHS/ BIG DATA
Social and Information Networks
RDF and Ontology
Knowledge Graphs
Web/ Semantic Search
Business Analytics
Smart City / Internet of Things
Health/ Medicare
Cyber Security
Viral Marketing

Future Direction: LLM+KG

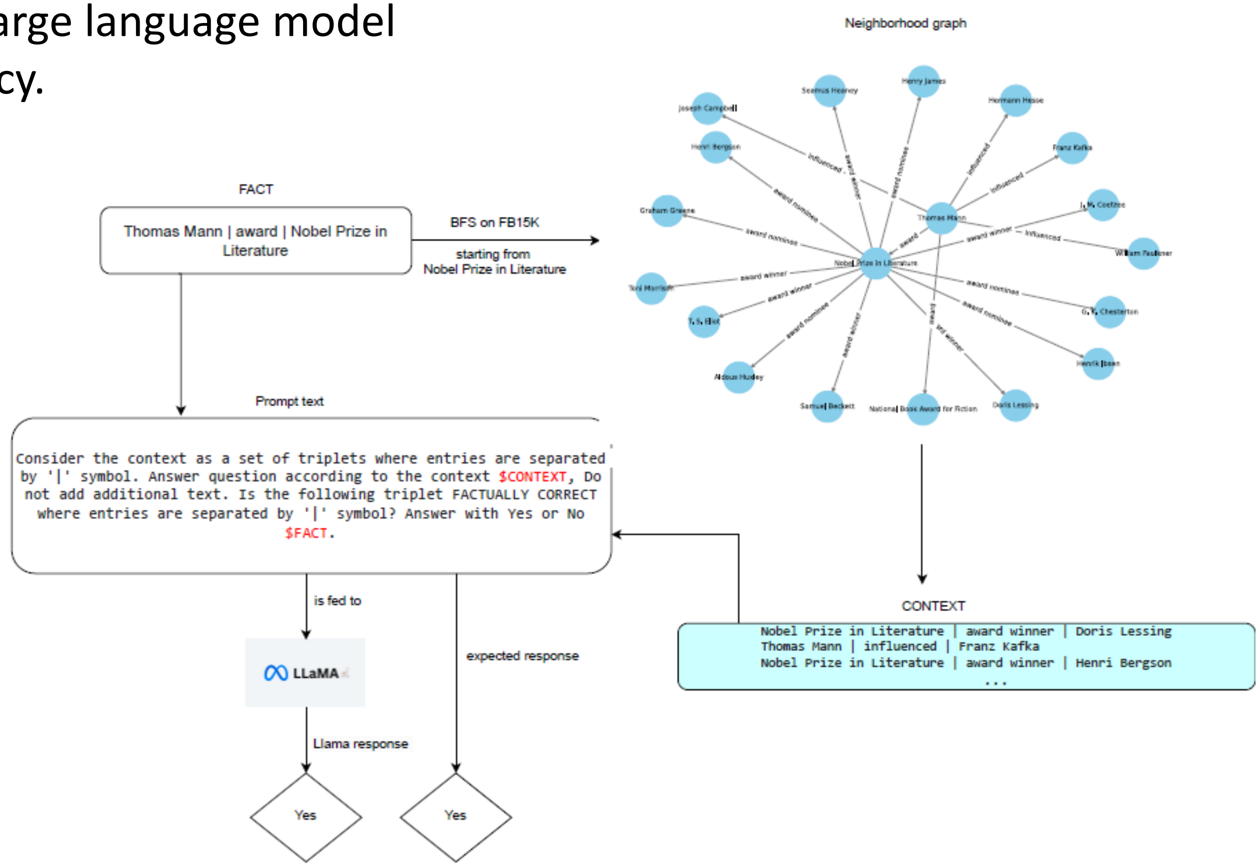
- **Retrieval-augmented generation (RAG)** to use KG context in order to improve a large language model (LLM)'s accuracy and consistency.

- **GraphRAG:**
  Graph-based retrieval-augmented generation

- **Complex Fact Checking**
- **Code Explanation**
- **Schema Matching**

Graph data management + Graph machine learning

Explainability of LLM



LLM-based Query Processing with KG Context