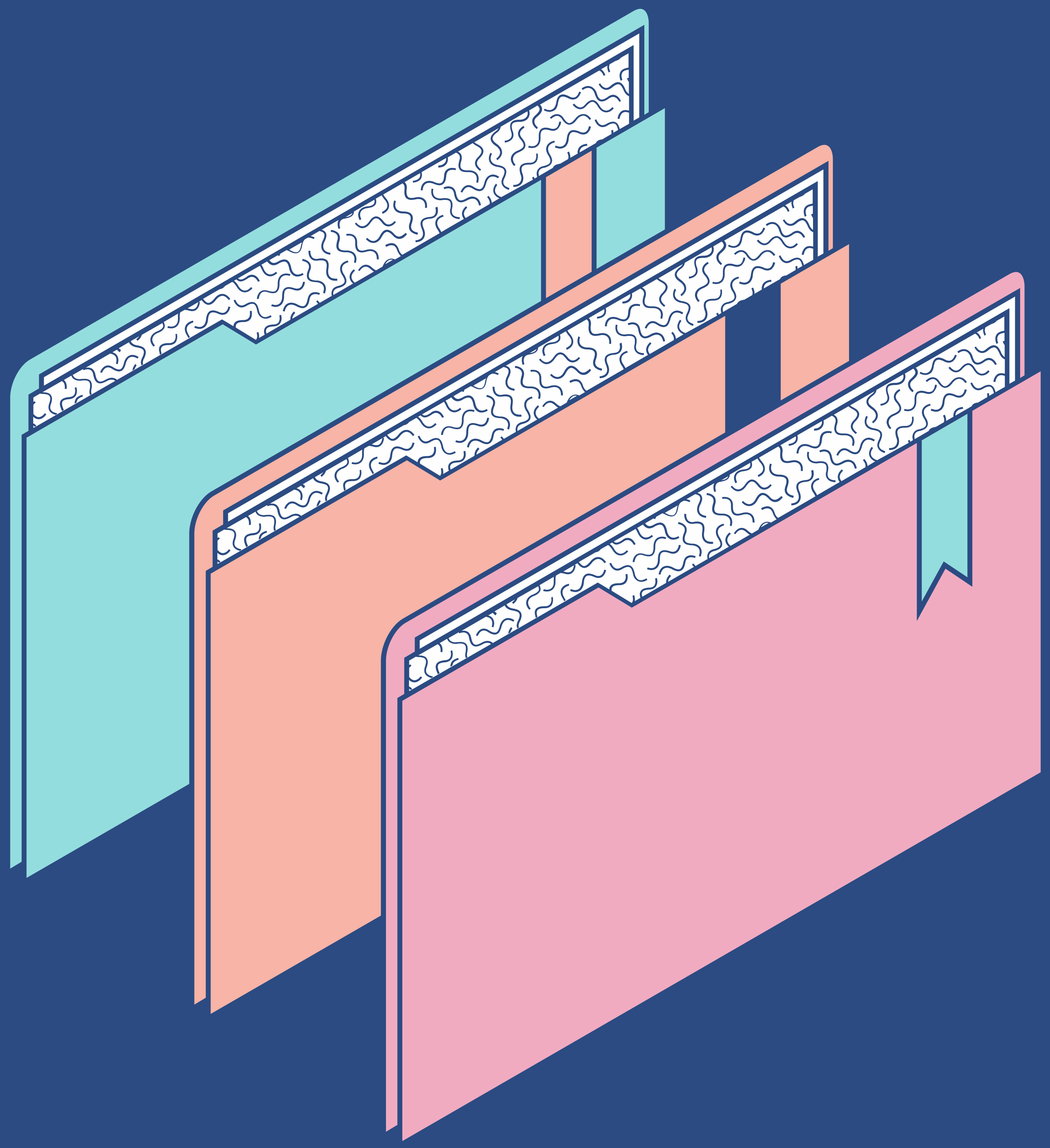
Fairness & Graph AD

Fairness in graph anomaly detection.

Shouju Wang



Agenda

KEY TOPICS DISCUSSED IN
THIS PRESENTATION

- Breif introduction
- Overview of fairness in deep learining
- Fairness notion in graph data mining*
- Optimazing methods*
- Conclusion

WHAT IS FAIRNESS?

fairness

noun [U]

UK /ˈfeə.nəs/ US /ˈfer.nəs/

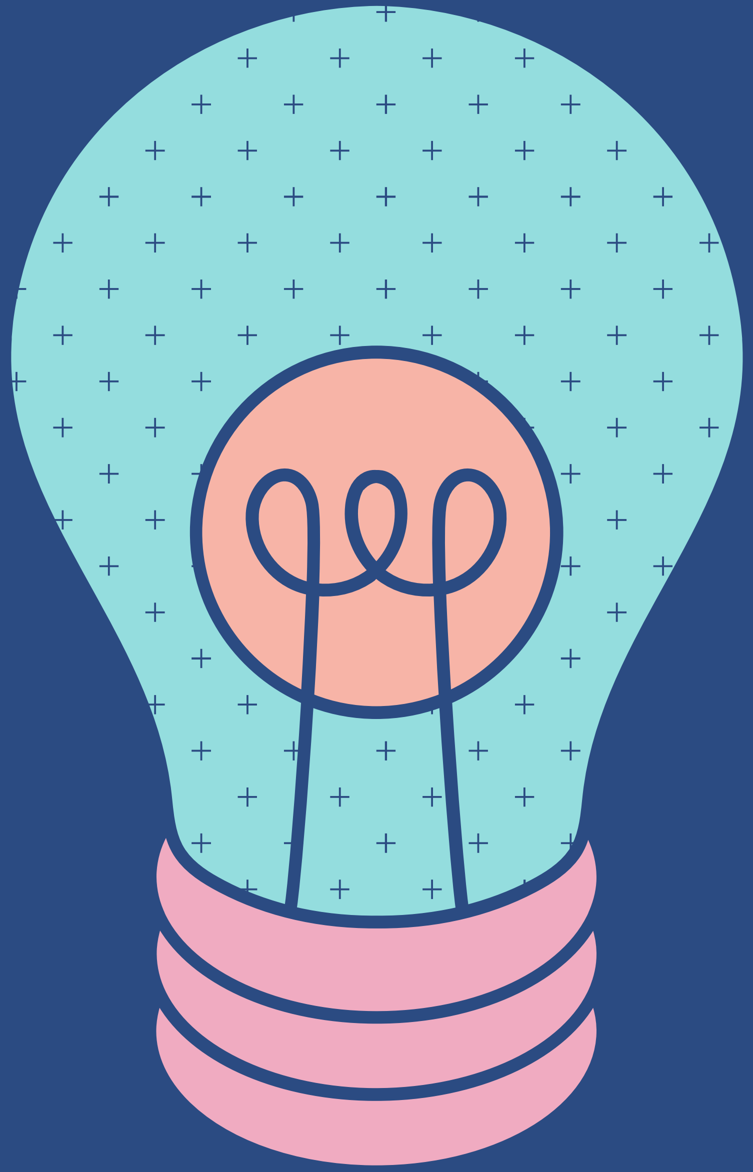
fairness ***noun*** **[U] (FAIR TREATMENT)**

Add to word list

C1

the quality of treating people equally or in a way that is right or reasonable:

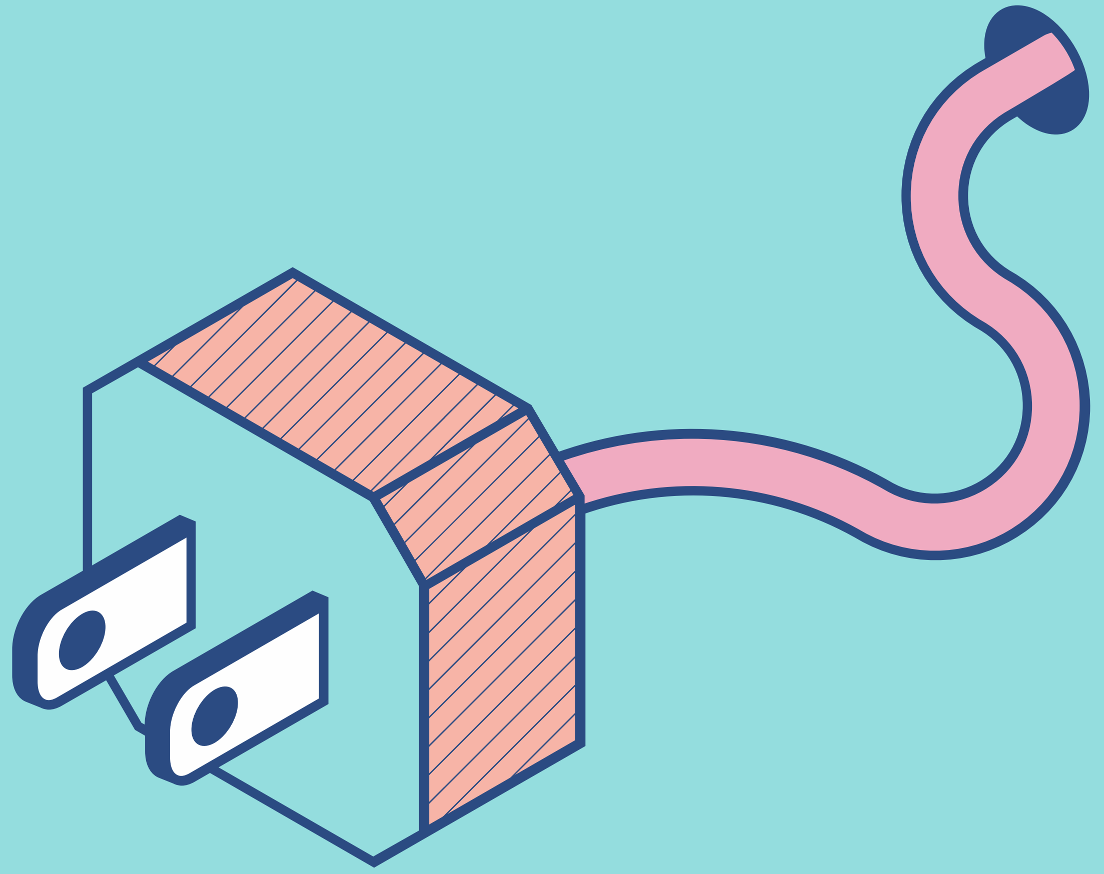
- *He had a real sense of fairness and hated injustice.*
- *The ban on media reporting has made some people question the fairness of the election (= ask whether it was fair).*

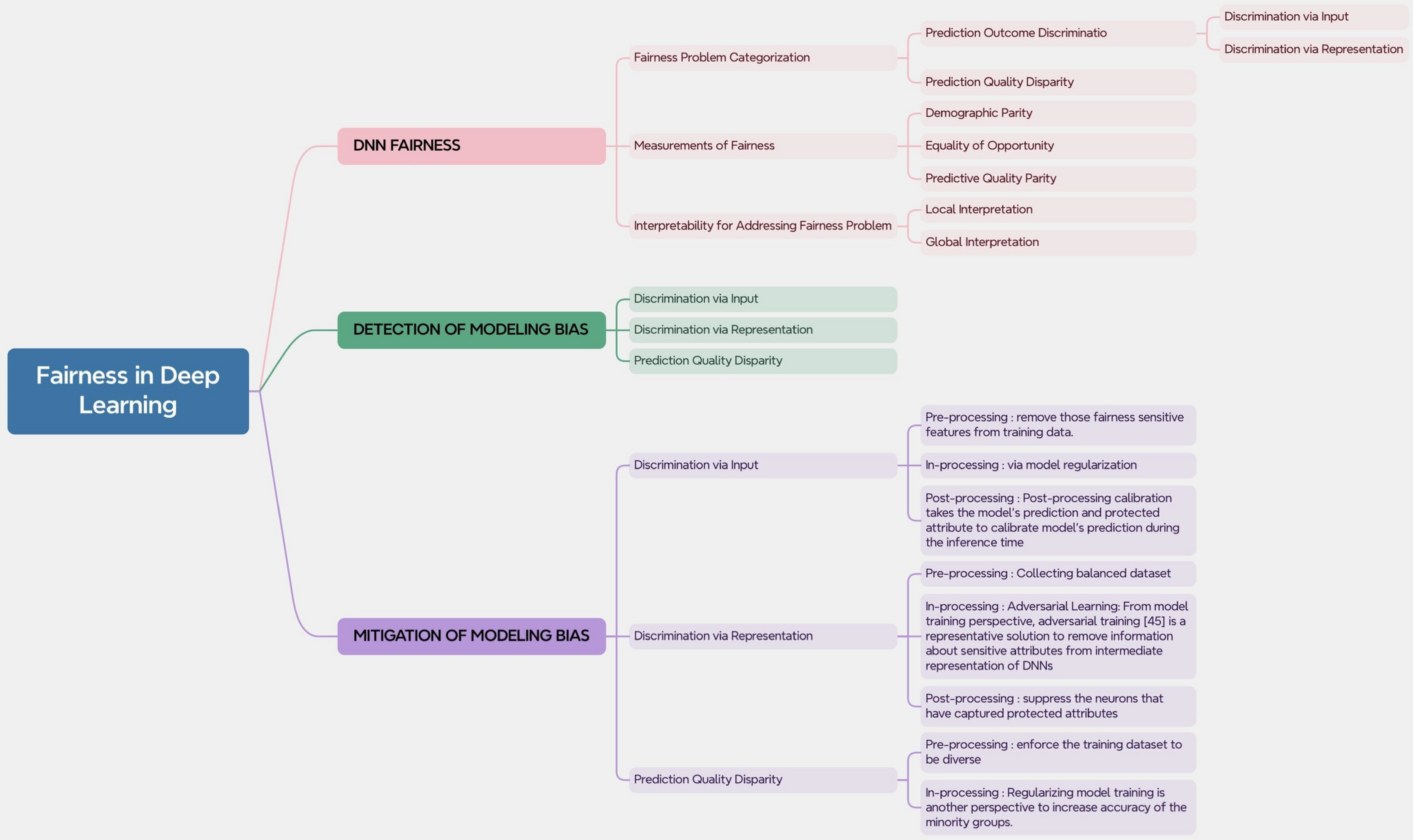


Overview of Fairness in Deep Learining

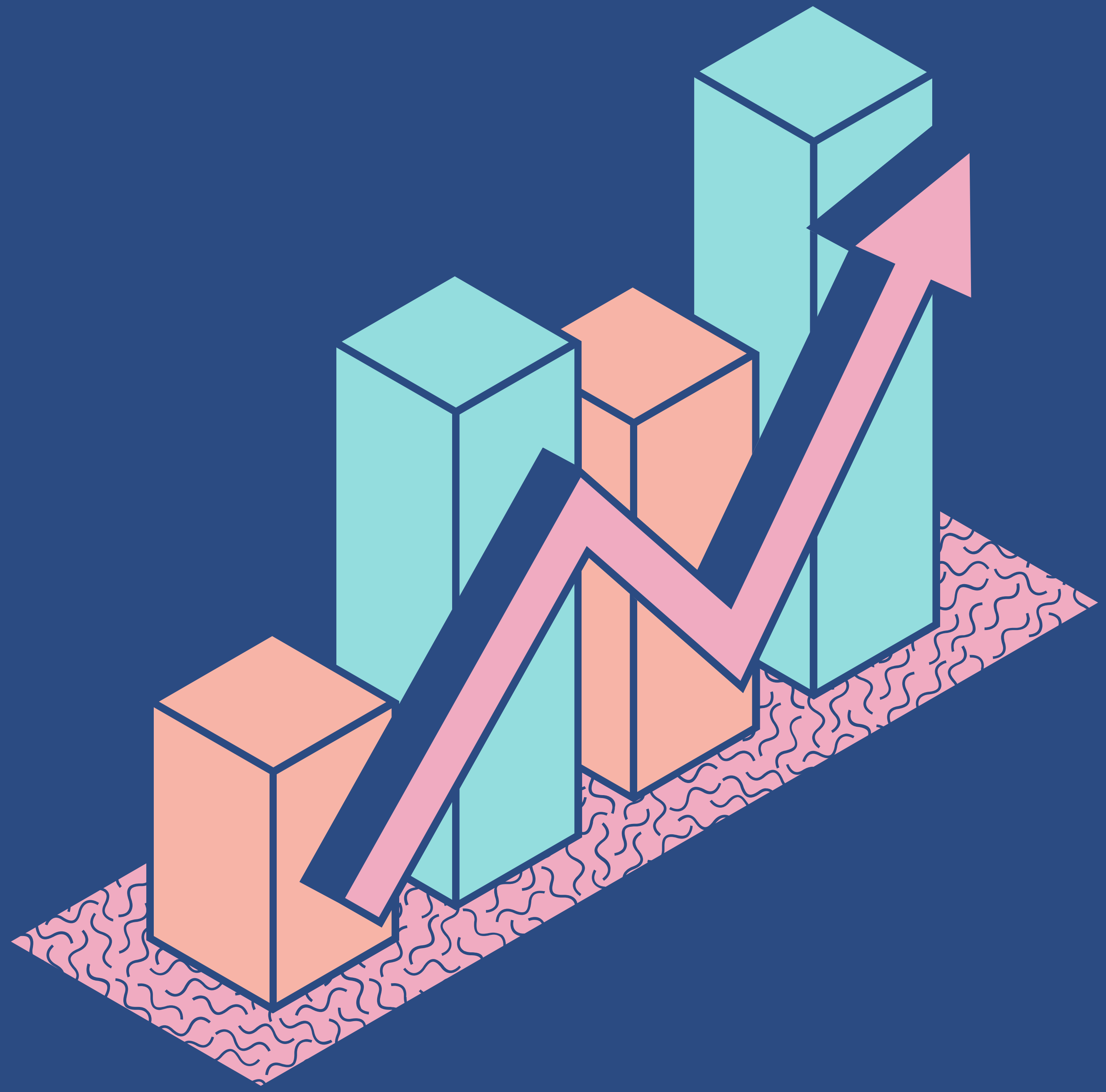
Why deep learning?

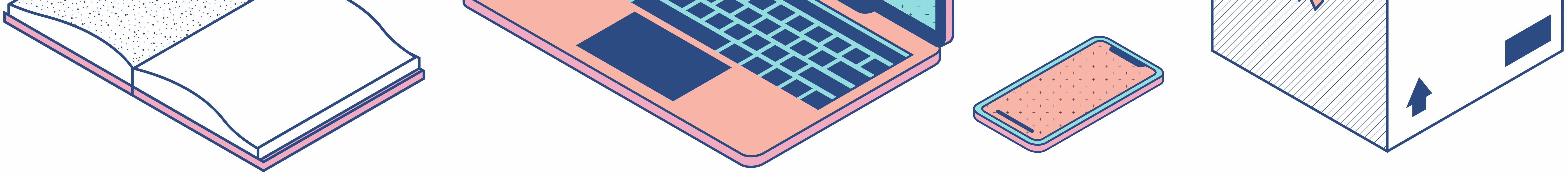
- State-of-art performance
- automatically learn from data
- data containing human bias
- have the risk of amplifying societal stereotypes by over associating protected attributes




Fairness in Deep Learning
DNN FAIRNESS
Fairness Problem Categorization
Prediction Outcome Discriminatio
Discrimination via Input
Discrimination via Representation
Prediction Quality Disparity
Measurements of Fairness
Demographic Parity
Equality of Opportunity
Predictive Quality Parity
Interpretability for Addressing Fairness Problem
Local Interpretation
Global Interpretation
DETECTION OF MODELING BIAS
Discrimination via Input
Discrimination via Representation
Prediction Quality Disparity
MITIGATION OF MODELING BIAS
Discrimination via Input
Pre-processing : remove those fairness sensitive features from training data.
In-processing : via model regularization
Post-processing : Post-processing calibration takes the model's prediction and protected attribute to calibrate model's prediction during the inference time
Discrimination via Representation
Pre-processing : Collecting balanced dataset
In-processing : Adversarial Learning: From model training perspective, adversarial training [45] is a representative solution to remove information about sensitive attributes from intermediate representation of DNNs
Post-processing : suppress the neurons that have captured protected attributes
Prediction Quality Disparity
Pre-processing : enforce the training dataset to be diverse
In-processing : Regularizing model training is another perspective to increase accuracy of the minority groups.

Fairness in Graph data mining





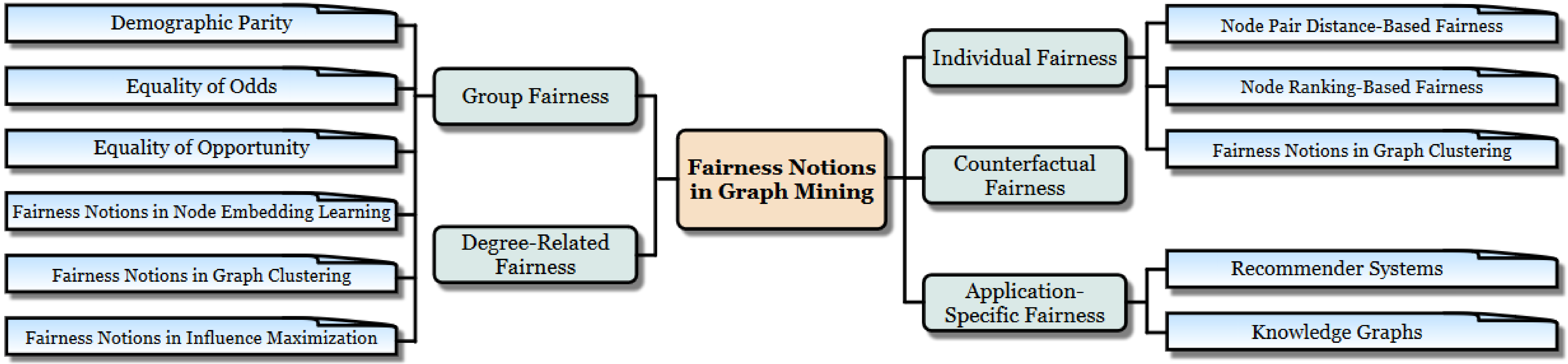


Fig. 2: Taxonomy of algorithmic fairness notions in graph mining algorithms.

Group Fairness

Generally speaking, group fairness requires that the algorithm should not yield discriminatory predictions or decisions against individuals from any specific sensitive subgroup

Demographic Parity

Demographic Parity is first introduced as a notion of group fairness based on binary sensitive feature(s) in binary classification tasks

Equality of Odds

In general, the algorithm predictions are enforced to be independent with the sensitive feature(s) conditional on the ground truth class labels.

Equality of Opportunity

*I*n binary classification tasks, equality of opportunity only requires the positive predictions to be independent of sensitive feature(s) for individuals with positive ground truth labels.

Notation

TABLE 1: Notations and the corresponding descriptions.

Notations	Definitions or Descriptions
$\lvert \cdot \rvert$	Cardinality operator for any set.
$E[\cdot]$	Expectation operator.
$< \cdot, \cdot >$	Inner product operator.
$\mathcal{G}$	The graph data.
$\mathcal{V}$	The set of nodes.
$\mathcal{E}$	The set of edges.
$\mathcal{X}$	The set of node features.
$\mathcal{A}$	The seed set in influence maximization.
$\mathcal{N}_{v_i}$	The one-hop neighboring node set of v_i.
$\mathcal{V}_i$	The node set of the i-th sensitive subgroup.
$\mathbf{A}$	The adjacency matrix of graph $\mathcal{G}$.
$\mathbf{A}^\top$	The transpose of adjacency matrix.
$\mathbf{X}$	The node feature matrix of graph $\mathcal{G}$.
$\mathbf{z}_i$	The embedding of node v_i.
v_i	The i-th node.
n	The size of the node set $\mathcal{V}$.
d	The number of node features.
c	The class number for node classification.

Demographic Parity

- criterion of demographic parity

$$P(\hat{Y} = 1 | S = 0) = P(\hat{Y} = 1 | S = 1).$$

- To quantify, formulation of ΔDP is given as

$$\Delta_{DP} = |P(\hat{Y} = 1 \mid S = 0) - P(\hat{Y} = 1 \mid S = 1)|.$$

Demographic Parity

- For graph anomaly detection

(indices values can be the same within each tuple). The criterion of demographic parity is given as $\delta_{i,j} = \delta_{k,l}, \ \forall i, j, k, l$, where $\delta_{i,j}$ is the average linking probability of node pairs spanning across the i-th and the j-th sensitive subgroup, and it is formally defined as

$$\delta_{i,j} = \frac{1}{N_{i,j}} \sum_{v \in \mathcal{V}_i} \sum_{v' \in \mathcal{V}_j} P(f_{\text{link}}(v, v') = 1). \qquad (3)$$

[137], [162]. $\Delta_{DP}^{\text{link}}$ is defined as the largest absolute difference [162] between all pairs of $\delta_{i,j}$ and $\delta_{k,l}$, which is given as

$$\Delta_{DP}^{\text{link}} = \max_{\forall i,j,k,l} |\delta_{i,j} - \delta_{k,l}|. \qquad (4)$$

Equality of Odds

Equality of Odds in Node Classification. In node classification, suppose that the predictions of the graph mining algorithm $\hat{Y}$, the ground truth labels Y, and the sensitive feature S are all binary. Equality of odds requires that

$$P(\hat{Y}=1|S=0,Y=y)=P(\hat{Y}=1|S=1,Y=y) \quad (8)$$

holds for both $y = 0$ and $y = 1$. In other words, Eq. (8) enforces predictions to bear equal TPR (i.e., True Positive Rate) and FPR (i.e., False Positive Rate) for the two sensitive subgroups. To quantify how well the equality of odds is satisfied, the largest difference of TPR (and FPR) between any

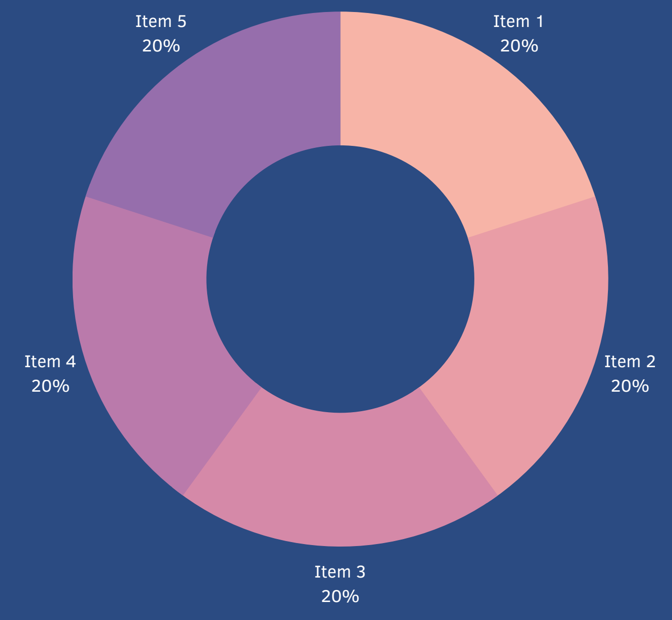
Equality of Opportunity

Equality of Opportunity in Node Classification. The criterion of equality of opportunity is given as

$$P(\hat{Y}=1|S=0,Y=1)=P(\hat{Y}=1|S=1,Y=1). \quad (9)$$

cation scenario is job candidate selection. We then introduce a commonly employed quantitative metric Δ_{EO} for equality of opportunity in node classification. Specifically, Δ_{EO} measures how far the prediction deviates from the ideal situation that satisfies equality of opportunity. Δ_{EO} is formally given as

$$\Delta_{EO} = |P(\hat{Y}=1 \mid Y=1, S=0) - P(\hat{Y}=1 \mid Y=1, S=1)|. \quad (10)$$



TECHNIQUES FOR IMPROVING FAIRNESS

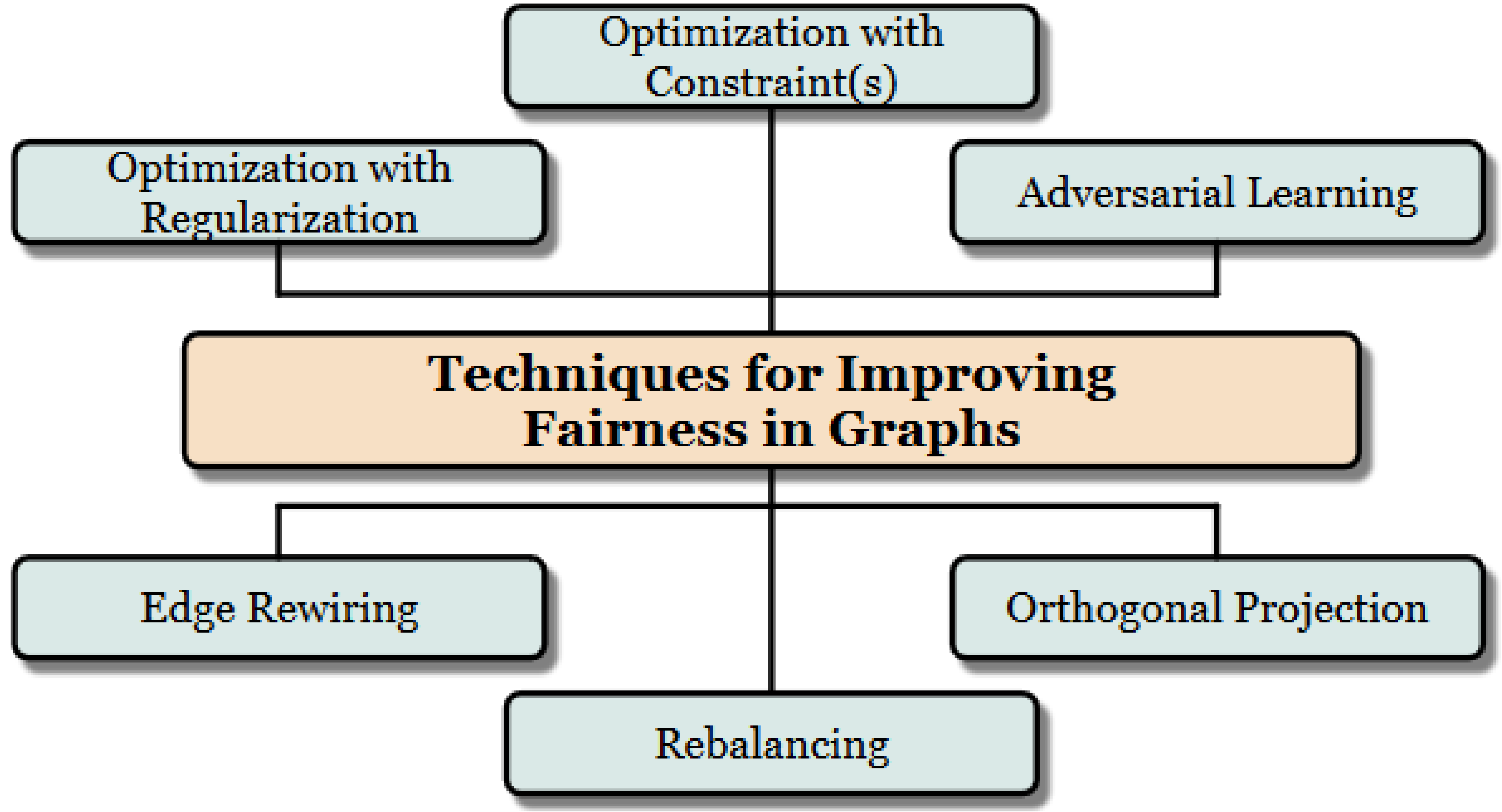


Fig. 3: A taxonomy of the commonly used techniques to improve fairness in graph mining.

OPTIMIZATION WITH REGULARIZATION

algorithm output. Formally, the total objective $\mathscr{L}$ is given as

$$\mathscr{L} = \mathscr{L}_{\text{utility}} + \lambda \mathscr{L}_{\text{fair}}, \tag{26}$$

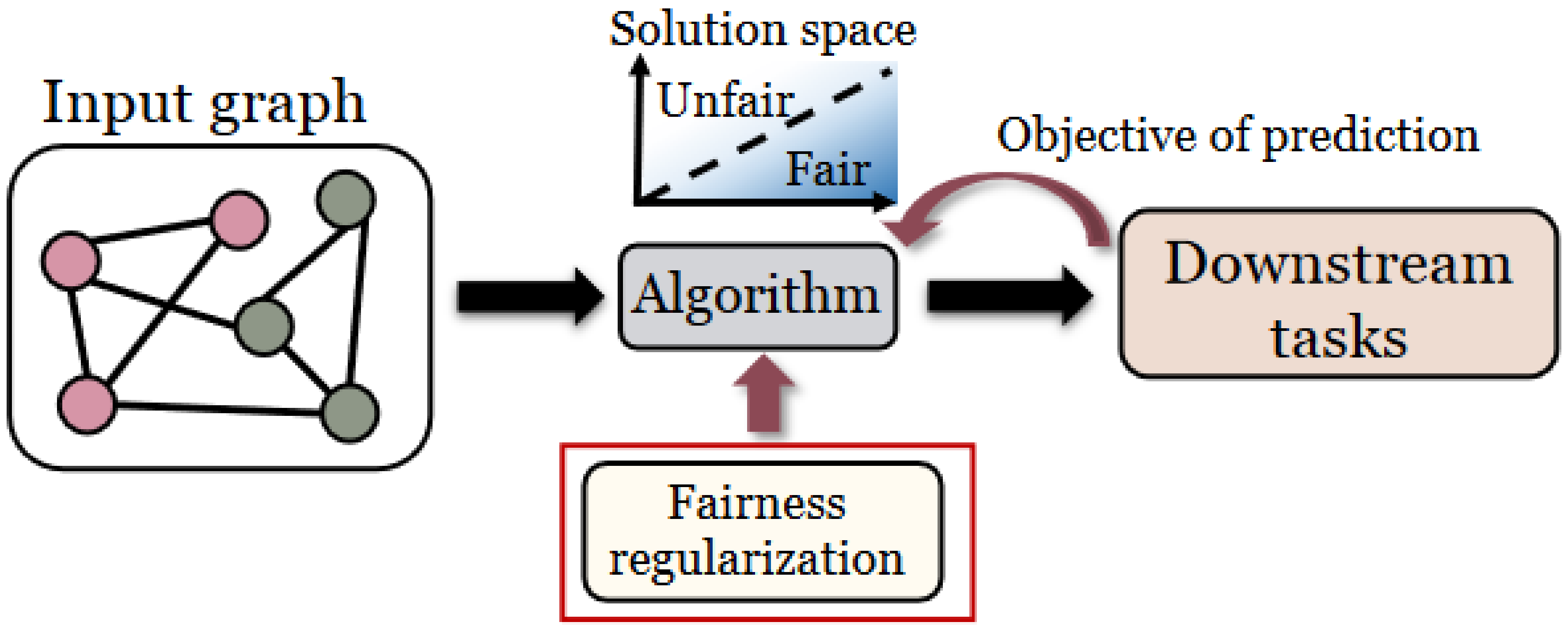
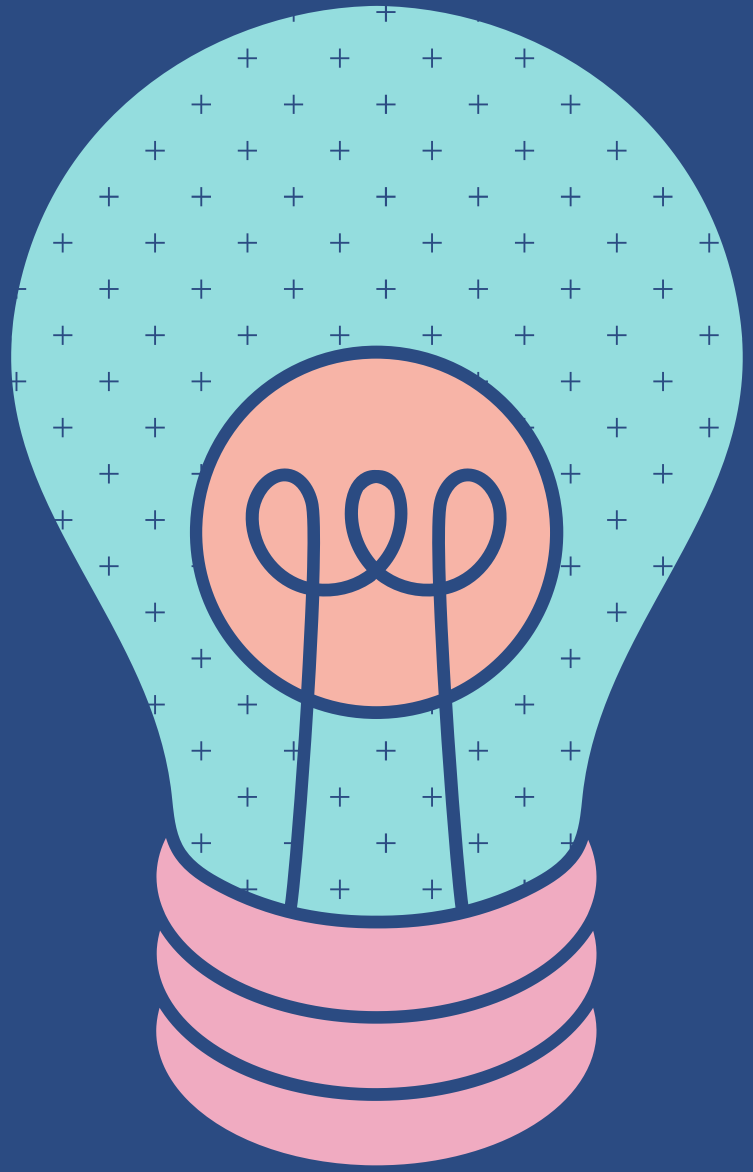


Fig. 4: The pipeline of optimization with regularization. The fairness regularization encourages the optimization result to stay in the area with a higher level of fairness (i.e., more blue) of the solution space.



FAIROD, CORRELATION, HIN OVERVIEW

as a replacement. To incorporate the fairness regularizer methods (*i.e.*, FAIROD, CORRELATION, and HIN) into the GAD methods, we made the following modifications to the GAD methods: (1) $\mathcal{L} = \mathcal{L}_o + \lambda\mathcal{L}_{FairOD} + \gamma\mathcal{L}_{ADCG}$ for FAIROD; (2) $\mathcal{L} = \mathcal{L}_o + \lambda\mathcal{L}_{HIN} + \gamma\mathcal{L}_{ADCG}$ for HIN; and (3) $\mathcal{L} = \mathcal{L}_o + \lambda\mathcal{L}_{corr}$ for CORRELATION, where $\mathcal{L}_o$ denotes the original loss of the GAD method, and λ and γ are hyperparameters.

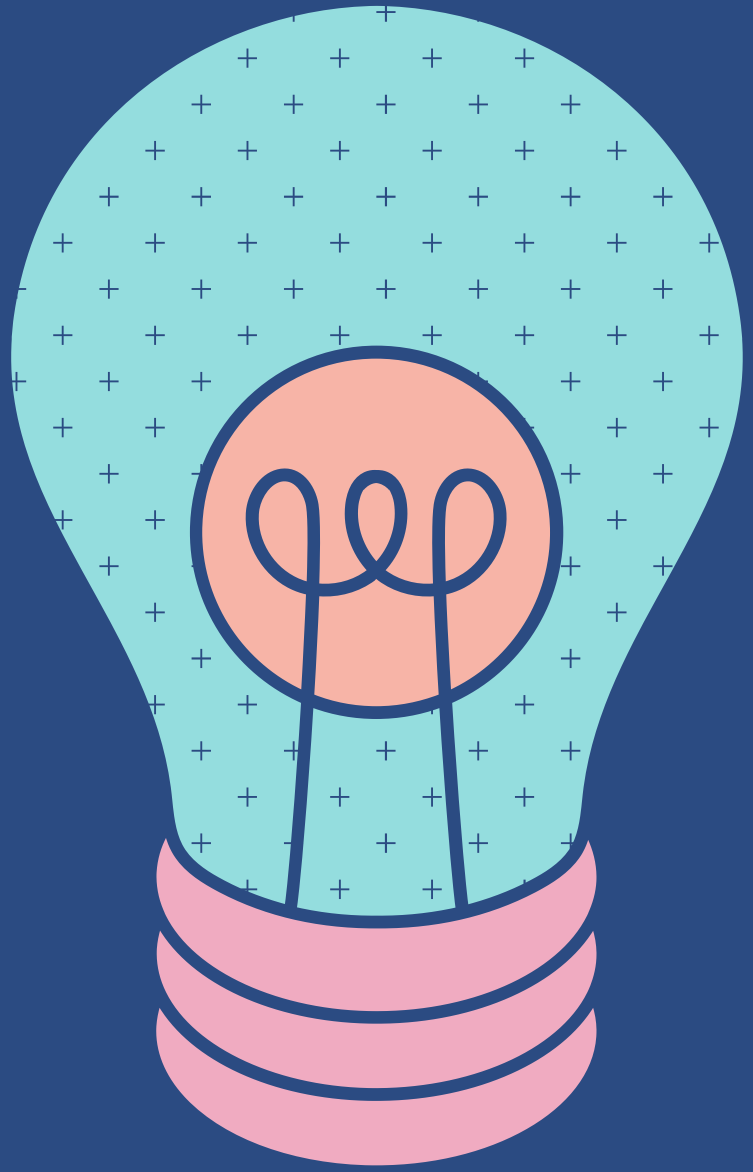
EG.FAIROD, CORRELATION, HIN

$$\mathcal{L}_{FairOD} = \left| \left(1 - \frac{1}{n}\right)^2 \frac{\left(\sum_{i=1}^{n} Err(v_i)\right)\left(\sum_{i=1}^{n} S(v_i)\right)}{\sigma_{Err}\sigma_S} \right|,$$

$$\mathcal{L}_{ADCG} = \sum_{s\in\{0,1\}} \left(1 - \sum_{\{v_i:S(v_i)=s\}} \frac{2^{BaseErr(v_i)} - 1}{\log_2\left(1 + IDCG_{S=s} \cdot DIFF(v_i)\right)}\right),$$

$$\mathcal{L}_{corr} = \left| \frac{(\mathbf{Err} \cdot \mathbf{S})}{\sqrt{(\mathbf{Err} \cdot \mathbf{Err})(\mathbf{S} \cdot \mathbf{S})}} \right|,$$

$$\mathcal{L}_{HIN} = \sum_{y\in\{0,1\}} \left(\frac{\sum_{\{v:S(v)=1\}} Pr(\hat{y}_v = y)}{|\{v : S(v) = 1\}|} - \frac{\sum_{\{v:S(v)=0\}} Pr(\hat{y}_v = y)}{|\{v : S(v) = 0\}|} \right)^2,$$



REBALANCING

Rebalancing aims to reduce the distribution difference of certain properties (e.g., the appearance rate of a node in random walks and frequency of recommendations of an item) between advantaged and disadvantaged nodes.

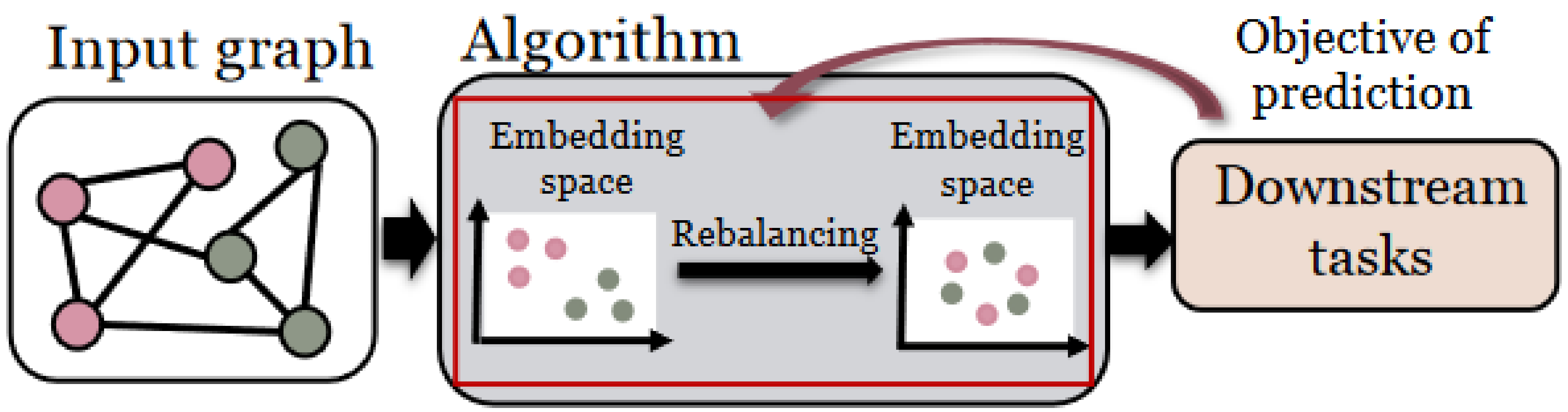
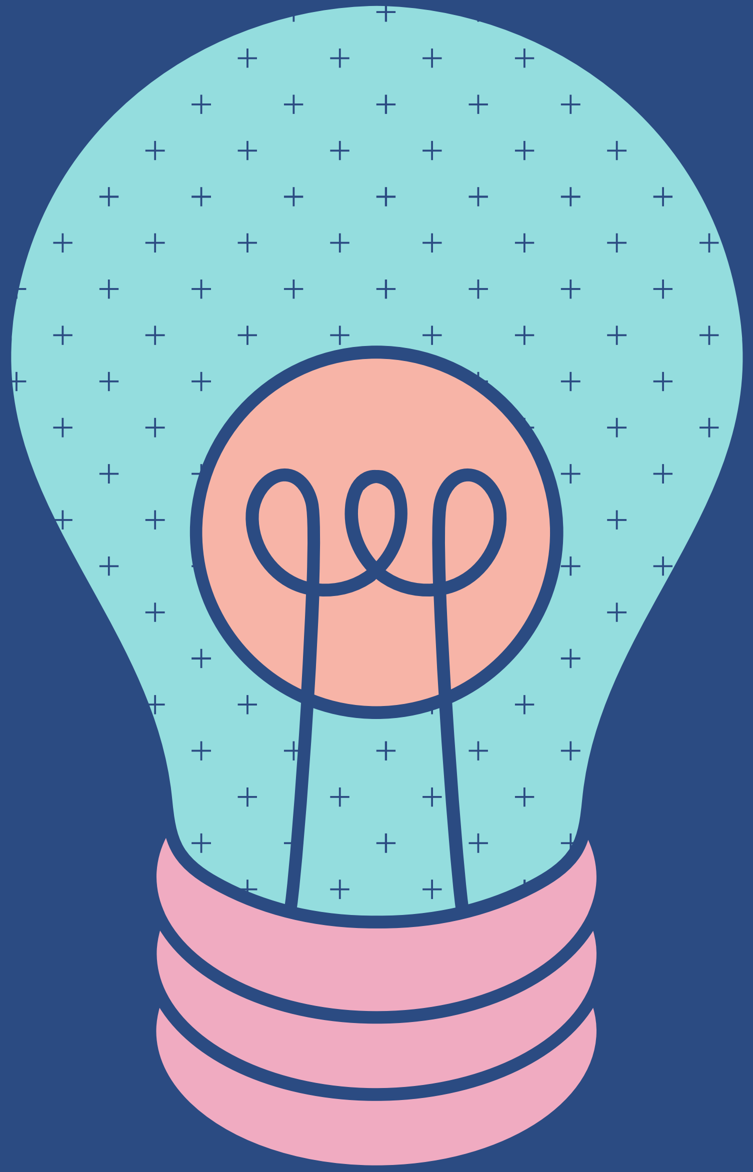


Fig. 6: The pipeline of rebalancing-based approaches. Generally, rebalancing aims to make certain characteristics to be as balanced as possible between the advantaged and disadvantaged individuals.



REBALANCING

eg. Fairwalk, Crosswalk

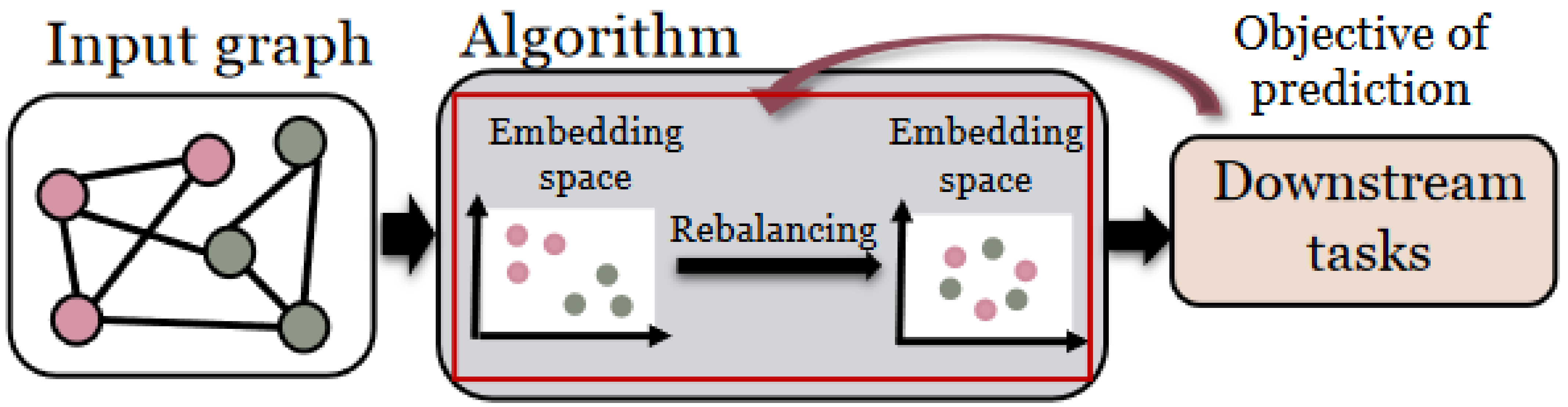
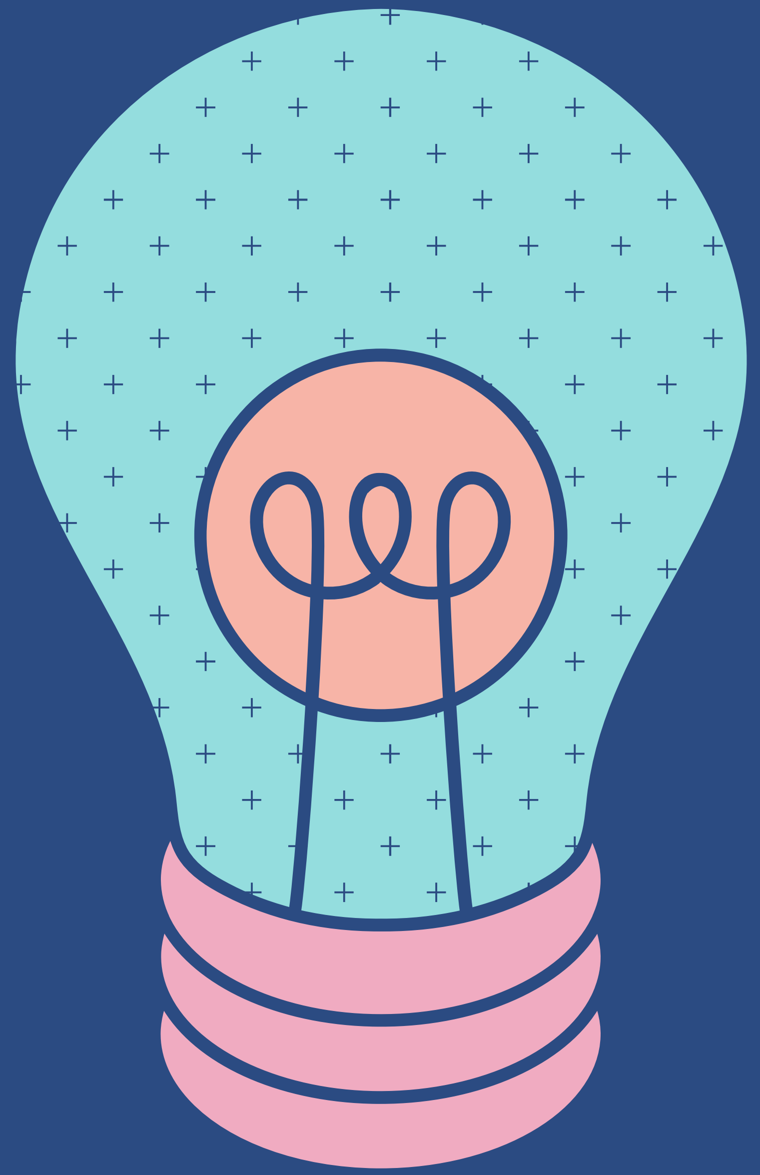


Fig. 6: The pipeline of rebalancing-based approaches. Generally, rebalancing aims to make certain characteristics to be as balanced as possible between the advantaged and disadvantaged individuals.



EDGE REWIRING

Biases exhibited in the node embeddings and algorithm predictions could also be attributed to the biased network topology.In this regard, modifying the graph topology through edge rewiring is a common debiasing strategy.

15

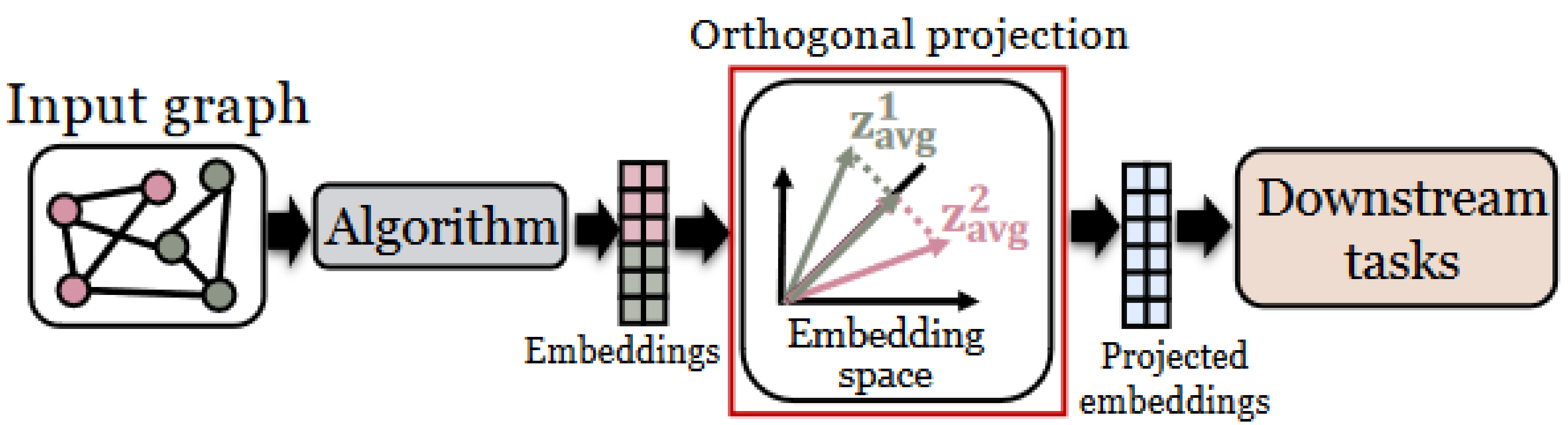
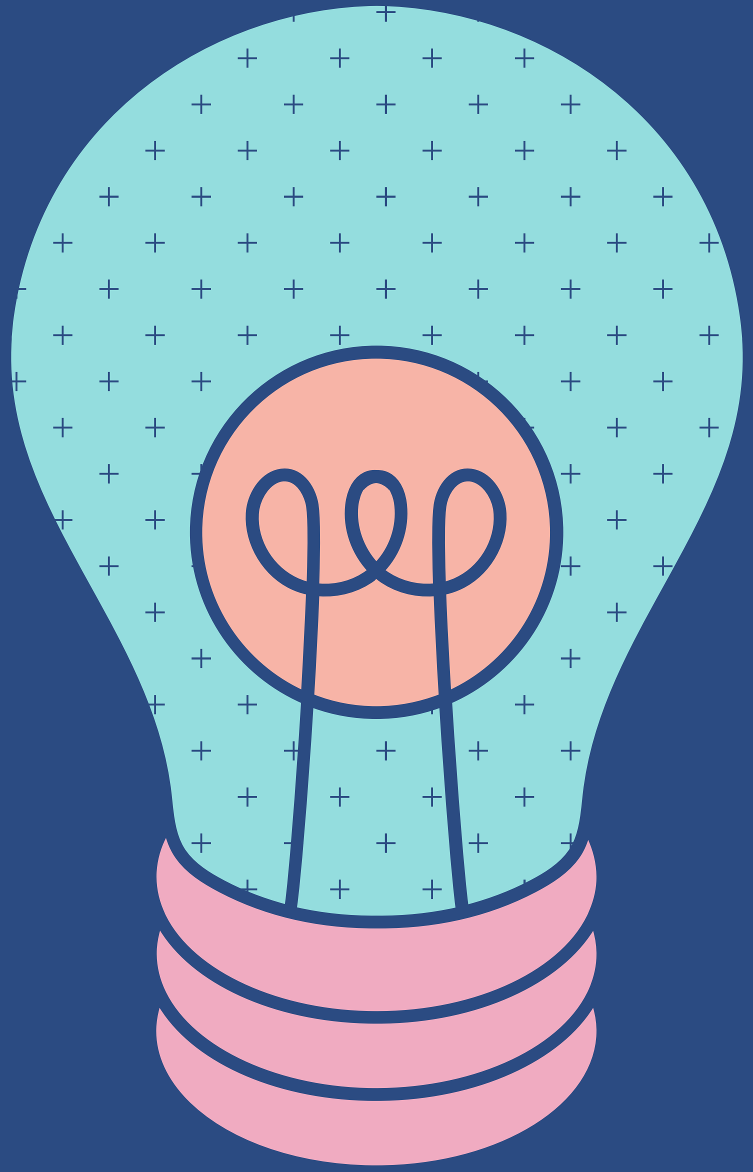


Fig. 9: The pipeline of orthogonal projection-based approaches. The embeddings of all individuals are projected onto the same hyperplane in the embedding space.



Conclusion

1 STEP

Optimization with regularization is the most widely used technique due to its simplicity and flexibility.

2 STEP

Optimization with constraint(s) is less explored in deep learning

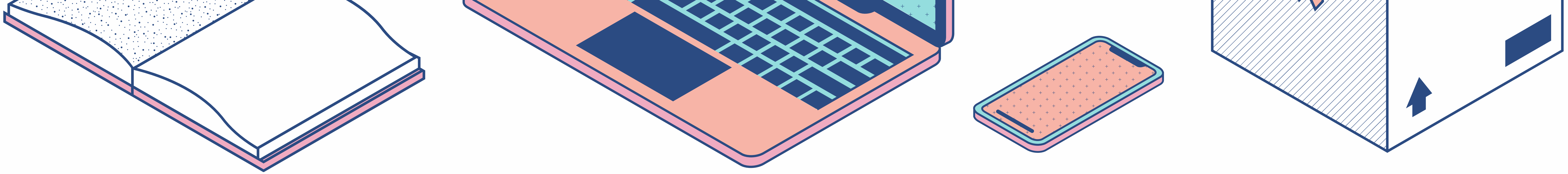
3 STEP

The strategy of rebalancing is mostly designed to be tailored for specific application scenarios

4 STEP

Both adversarial learning and orthogonal projection aim to explicitly remove the sensitive information from the algorithm output.

Technology is an effective tool that can make education more meaningful and engaging for teachers and students alike.



Traditional Teaching and Learning

- Physical learning materials and equipment like paper, pens and chalkboard
- Limited access to education materials and information
- Teaching and learning typically occurs in an in-person classroom setting

Teaching and Learning with Technology

- Wider access to education materials and information
- More available channels and tools for communication as well as collaboration
- Enables a more personalized kind of learning for students

Do you have any questions?

Thanks for your attention.