24th February 2025, Lecture Note: Causalality (continued); Fair Ranking – Part I: Score-based Ranking

Madhura Dongare, Resham Patil

CS 516: Responsible Data Science and Algorithmic Fairness; Spring 2025
Abolfazl Asudeh; www.cs.uic.edu/∼ asudeh/teaching/archive/cs516spring25/

1 Introduction

Causal graphs help us understand relationships between variables, distinguishing between correlation and causation. This lecture explores different causal structures, the impact of confounding and collider bias, and the importance of blocking backdoor paths to ensure valid causal inferences.

This lecture examined the critical differences between ranking and classification problems. Unlike classification, ranking creates a competitive environment where decisions about one item affect all others. We explored how ranking takes sets of elements as input rather than individual items, and why this fundamental difference creates unique fairness challenges. Through examples from college admissions, search engines, and hiring, we demonstrated how the interdependent nature of ranking decisions requires specialized approaches to fairness.

2 Types of Causal Structures

2.1 Fork (Common Cause)

Structure: $X \leftarrow Z \rightarrow Y$

X and Y are related through a common cause, Z, which introduces a confounding effect. This means that the observed relationship between X and Y is not necessarily causal but rather influenced by their shared dependence on Z. As a result, the probability of Y given an intervention on X, denoted as $P(Y \mid do(X))$, is not the same as the conditional probability $P(Y \mid X)$, highlighting the presence of confounding. To address this issue, conditioning on Z removes the confounding effect, ensuring that X and Y become independent and allowing for a more accurate causal interpretation.

2.2 Chain (Indirect Causal Path)

Structure: $X \rightarrow Z \rightarrow Y$

In a chain structure, X causes Z, which in turn causes Y, forming an indirect causal link between X and Y. This means that any effect X has on Y operates through Z rather than directly. In this case, the probability of Y given an intervention on X, $P(Y \mid do(X))$, is equal to the conditional probability $P(Y \mid X)$, indicating that X and Y are causally connected through the mediating variable Z. However, if we condition on Z, the indirect causal effect is removed, breaking the connection between X and Y.

2.3 Collider (Spurious Association)

Structure: $X \rightarrow Z \leftarrow Y$

In a collider structure, X and Y are independent, but conditioning on Z creates an artificial association between them, leading to collider bias. Normally, no information flows between X and Y, meaning they remain uncorrelated. However, when Z is conditioned on, it introduces a spurious correlation between X and Y, making it seem as though they are related when they are not. This unintended association can distort causal analysis and lead to incorrect conclusions about the relationship between variables.

3 Key Concepts

3.1 Confounding Effect

The confounding effect arises in fork structures where a common cause, Z, influences both X and Y, creating a spurious association between them. This confounding can lead to misleading conclusions about the relationship between X and Y. However, by conditioning on Z, the confounding effect is removed, ensuring that any observed association between X and Y is not due to their shared dependence on Z but rather a genuine causal link, if one exists.

3.2 Collider Bias

Conditioning on a collider (Z) can falsely create a correlation between two variables (X and Y) that are otherwise independent. This occurs when both X and Y influence Z, and conditioning on Z introduces a spurious relationship between X and Y. For example, if the education level (X) and work experience (Y) of job applicants both affect their hiring status (Z), conditioning on the hiring status creates a misleading correlation between education and experience. In this case, the correlation between education and experience is not due to any direct relationship between the two variables but is instead a result of their shared influence on the hiring decision.

3.3 Backdoor Path

A backdoor path occurs when there is a directed path from X to Y, but the edge leading to X is reversed. This creates a structure where information flows through a non-causal route, allowing confounding to influence the relationship between X and Y. For example, in the structure $X \leftarrow Z \rightarrow Y$, the edge connecting X to Z is referred to as the backdoor edge. This backdoor path enables the flow of information between X and Y that is not due to a direct causal relationship, potentially distorting the true connection between the variables.

3.4 Blocking Backdoor Paths

To remove confounding effects, block backdoor paths by conditioning on variables that block the flow of non-causal information.

Examples of blocking: Consider: $X \leftarrow Z \rightarrow Y$

To remove the confounding effect, we need to compute the correlation between X and Y conditioned on Z $P(X, Y \mid Z)$.

Consider: $X \leftarrow Z \rightarrow W \rightarrow Y$

Compute $P(X, Y \mid Z)$ or $P(X, Z \mid W)$ to remove the confounding effect. In both cases, the flow of information is stopped.

Important: Consider: $X \leftarrow Z \rightarrow R \leftarrow W \rightarrow Y$

The path connecting X and Z is not a backdoor path, since it is not possible to have a flow of information from X to Y as there is a collider. If you condition on a collider, it creates collider bias and should be avoided.

4 Studying Causal Effects

Step 1: Identifying Backdoor Paths

Ensure that all non-causal correlations are blocked by conditioning on appropriate variables. The goal is that after blocking confounding, $P(Y \mid do(X))$ should be equal to $P(Y \mid X)$, ensuring a purely causal relationship.

Step 2: Evaluating Disparity and Discrimination

Argument A: No direct effect between X and Y after conditioning on confounding variables (C) and intermediate chain variables (Ψ):

$$X \perp Y \mid C, \Psi$$

Argument B: If only indirect effects remain and they are socially acceptable, the disparity is not considered discriminatory.

5 The Berkeley Case: Example of Discrimination in Causality

Scenario: Gender (X) influences department choice (Z), which in turn affects admission (Y).

$$X \to Z \to Y$$

Analysis: After conditioning on department choice, gender has no direct causal effect on admission.

Conclusion: If department choice is a socially acceptable mediator, gender disparity in admissions is not inherently discriminatory.

6 Potential Pitfalls in Causal Reasoning (Masking)

6.1 Condition A: No Direct Effect

Violation Example: If department funding depends on state residence, admission rates may correlate with department choice due to state policies rather than gender.

$$\text{Causal Graph: } X \to Z \leftarrow R \to Y$$

Issue: R (state policies) is an unobserved confounder creating collider bias when conditioning on Z.

6.2 Condition B: Acceptability of Indirect Effects

Violation Example: If certain departments discourage female applicants through biased advertising, gender disparities persist indirectly through applicant self-selection.

$$\text{Causal Graph: } X \to W \to Z \to Y$$

Issue: If W (fear of discrimination) is unobserved, it falsely justifies gender disparities.

7 Policy-Level Implications

Systematic discrimination can be intentionally masked by leveraging causal graphs.

Example: If policymakers adjust department funding based on past gender disparities, they reinforce biased outcomes while appearing neutral.

Takeaway: Policymakers must account for all backdoor paths and unobserved confounders to avoid reinforcing hidden biases.

8 Fundamental Distinction: Ranking vs. Classification

Discussions of fairness in ranking must begin by understanding how ranking fundamentally differs from classification tasks, because these differences have profound implications for how we conceptualize and implement fairness.

8.1 Key Differences

In classification, we take a single element as input and make an independent decision about it. But in ranking, we're dealing with an entire set of elements, and we need to order them relative to each other. This creates a competitive environment where decisions are inherently interdependent.

The input structure for classification is a single element, while ranking takes a set of elements as input. Classification decisions are made independently for each item, while ranking decisions are made in conjunction with others. This means changing one score in a ranking system potentially affects everyone else's position, not just that individual.

Error sensitivity also differs dramatically. In classification, small errors might not change the outcome. For example, if I'm grading students with a pass/fail threshold of 50, it doesn't matter if someone scores 65, 67, or 68 - they all pass. But in ranking, even tiny changes can significantly alter the entire list.

This competitive aspect is critical to understand. Think about Olympic runners - it doesn't just matter how fast you run in absolute terms; what matters is whether others run slower than you. Your grade in a class depends only on your performance, but your ranking depends on how you compare to others.

9 Types of Ranking Problems

Ranking problems appear in various contexts, each with distinct characteristics and fairness challenges.

9.1 Fixed Data, Current Performance

In this scenario, we have a fixed dataset and rank items based on historical or current data using a fixed preference function. Examples include ranking billionaires based on net worth, ranking universities based on publication metrics, and ranking students based on GPA and standardized test scores. For these problems, the ranking criteria are typically well-defined and transparent, though they may still contain biases.

9.2 Fixed Data, Expected Future Performance

Here, we have fixed data but rank items based on predicted future outcomes. We might rank university applicants based on expected success rate, job candidates based on predicted job performance, or loan applicants based on predicted repayment likelihood. These problems introduce additional complexity because the predictions themselves may contain biases or amplify existing inequalities.

9.3 Variable Data, Variable Preference

In this most complex scenario, both the data and the preference functions change dynamically. The classic example is information retrieval/search results, where each user has different preferences, queries filter and change the relevant data, and users apply personal preference functions to results. These systems must balance fairness considerations with personalization, creating unique challenges.

10 Fairness Issues in Ranking

Ranking systems can perpetuate or amplify various forms of bias. Understanding these biases is essential for developing fair ranking approaches.

10.1 Types of Bias

Pre-existing bias exists independently of algorithms and has its origins in society. For example, SAT math scores show significant disparities across racial groups (White: 534, Black: 428). These differences reflect deeper racial and class inequalities experienced early in life.

Technical bias arises from technical constraints or considerations, such as position bias where items at the top of rankings receive disproportionate attention because in Western cultures, we read from top to bottom and left to right. This creates an inherent advantage for items positioned high in the list.

Emergent bias develops over time through system use. Users tend to trust items at the top of search results, which shapes their expectations and creates self-reinforcing feedback loops, leading to "winner-takes-all" situations where popular items become increasingly popular.

10.2 Group Structure Considerations

When addressing fairness in ranking, we must consider the cardinality of sensitive attributes (binary or multinary), the number of sensitive attributes (single or multiple), and how we handle multiple attributes (independently or considering their intersections). These considerations significantly impact our approach to fairness in ranking systems.

11 Conclusion

Causal inference is a powerful tool, but it requires careful analysis to ensure accurate conclusions. One of the key steps in this process is blocking backdoor paths, which is essential for making valid causal interpretations. Identifying and addressing unobserved confounders is also crucial, especially when evaluating policies and fairness, as they can introduce hidden biases that distort the results. Additionally, collider bias should be avoided when making causal claims, as it can create misleading associations between variables. In both policy-making and research, it is vital to ensure that the remaining causal effects are not only valid but also socially acceptable, to prevent unjust outcomes and ensure fairness in decision-making.

Ranking differs fundamentally from classification through its competitive nature where changes to one item affect all others. We've explored three types of ranking problems (fixed data/current performance, fixed data/future predictions, and variable data/preferences) and identified three forms of bias (pre-existing, technical, and emergent). Understanding these distinctions is essential for developing ranking systems that balance utility with fairness across demographic groups.

References

[1] M. Zehlike, K. Yang, J. Stoyanovich. "Fairness in ranking, part i: Score-based ranking." ACM Computing Surveys, 55.6, 2022, pp. 1-36.

[2] E. Pitoura, K. Stefanidis, G. Koutrika. "Fairness in rankings and recommendations: an overview." The VLDB Journal, 2022, pp. 1-28.