Causal Inference with Graphical Neural Networks

Your Name

April 9, 2024

1 Introduction

2 do Operator

The **do** operator is a way to represent interventions in a causal model. It is a way to represent the effect of an intervention on a variable. As an example, consider the following model involving smoking.

If a person's fingernails (N) have turned yellow, this implies a higher probability that they are a heavy smoker (S) and hence have a higher probability of developing lung cancer (C). But, simply dyeing a persons fingernails yellow does not impact their probability of developing lung cancer.

So, in terms of **do** calculus, we can denote the process of setting a variable N to have a value *yellow* by $\mathbf{do}(N = \textit{yellow})$. We note that

$$P(C \mid N = \textit{yellow}) \neq P(C \mid \mathbf{do}(N = \textit{yellow})).$$

With this in mind, we now define the **do** operator.

Theorem 2.1 ([1]) *In a causal diagram Γ with nodes $X_1, \ldots, X_n$ and joint distribution $P(X_1, \ldots, X_n)$, the result of doing $X_i = x_i$ on the joint distribution is*

$$P(X_1, \ldots, X_n \mid \boldsymbol{do}(X_i = x_i)) = \frac{P(x_1, \ldots, x_n)}{P(x_i \mid \text{par}(x_i))} = \prod_{j \neq i} P(x_j \mid \text{par}(x_j)).$$

In this, we have $\text{par}(x_i)$ represent values of the parent nodes of $\text{PAR}(X_i)$ of X_i in Γ. The probabilities on the right hand side of the above equation are what we call *preintervention*. This means they use the original probabilities from the original model before doing $X_i = x_i$.

It is important to note that the above equation is how we calculate the probability of several events happening given one event has happened. What if we want to get the probability of a single event happening, given we do a single event? That leads to the following corollary.

Corollary 2.1.1 *If X and Y are random variables in a causal diagram Γ and* $\mathrm{PAR}(X)$ *are the parents of X, then*

$$P(y \mid \boldsymbol{do}(x)) = \sum_{\text{par}} \frac{P(x,\, y,\, \text{par})}{P(x \mid \text{par})},$$

where the sum runs over all values par *that the variables* $\mathrm{PAR}(X)$ *can take. If X has no parents, then*

$$P(y \mid \boldsymbol{do}(x)) = \frac{P(x,\, y)}{P(x)} = P(y \mid x).$$

3 Background

4 Methodology

5 Experiments

5.1 Dataset Description

5.2 Experimental Setup

5.3 Results

6 Discussion

7 Conclusion

Acknowledgments

References

[1] Judea Pearl. Causal inference in statistics: An overview. *Statistics Surveys*, 3:96–146, 2009.