

- SHAP 개발자인 Scott Lundberg (MS) 가 투고한 article

Be careful when interpreting predictive models in search of causal insight



However, in this article, we discuss how using predictive models to guide this

kind of policy choice can often be misleading.

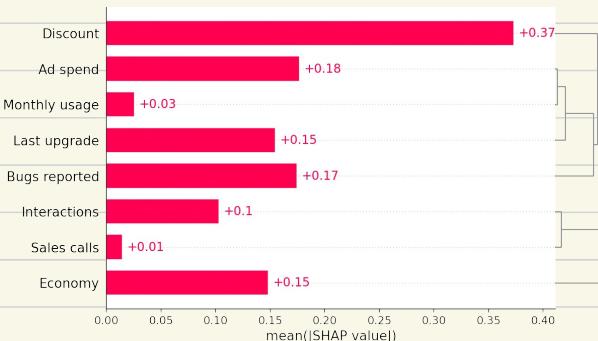
그저 관측된 feature의 상관관계만
트哧려져 할 수 있음

- correlation 과 causation 은 근본적인 차이가 있기 때문에

■ Example

: 학습, 광고 지출, 고객의 월별 사용량, 마지막 업그레이드 시점, 고객이 보고한 버그, 고객과의 대화,
고객과의 영업 상담, 경제 활동을 기반으로 서비스 개선 여부 예측하는 문제

features



outcome



SHAP 분석 결과

- 더 많은 버그를 보고한 사용자는 서비스 개선할 가능성 높음
- 할인폭이 큰 사용자는 서비스를 개선할 가능성 낮음 ???

■ The challenges of estimating causal effects

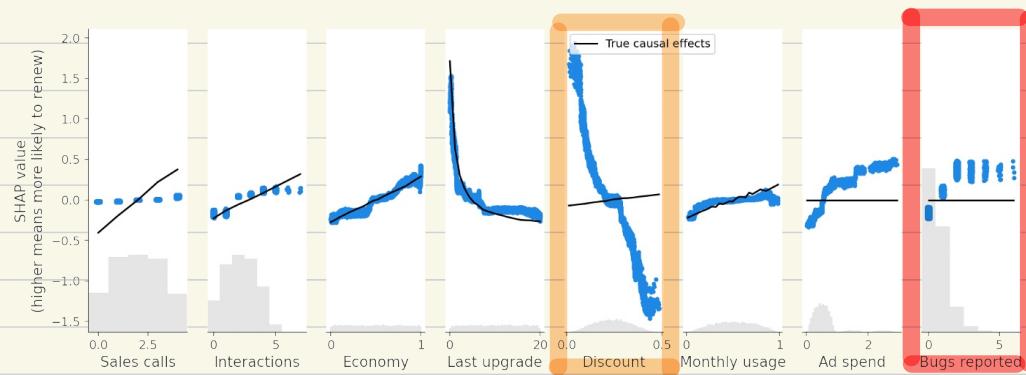
- 인과관계를 이해하기 위해서 causal graph 그려볼 것

- 관측된 feature 와 숭겨진 feature 모두 그려봄

solid

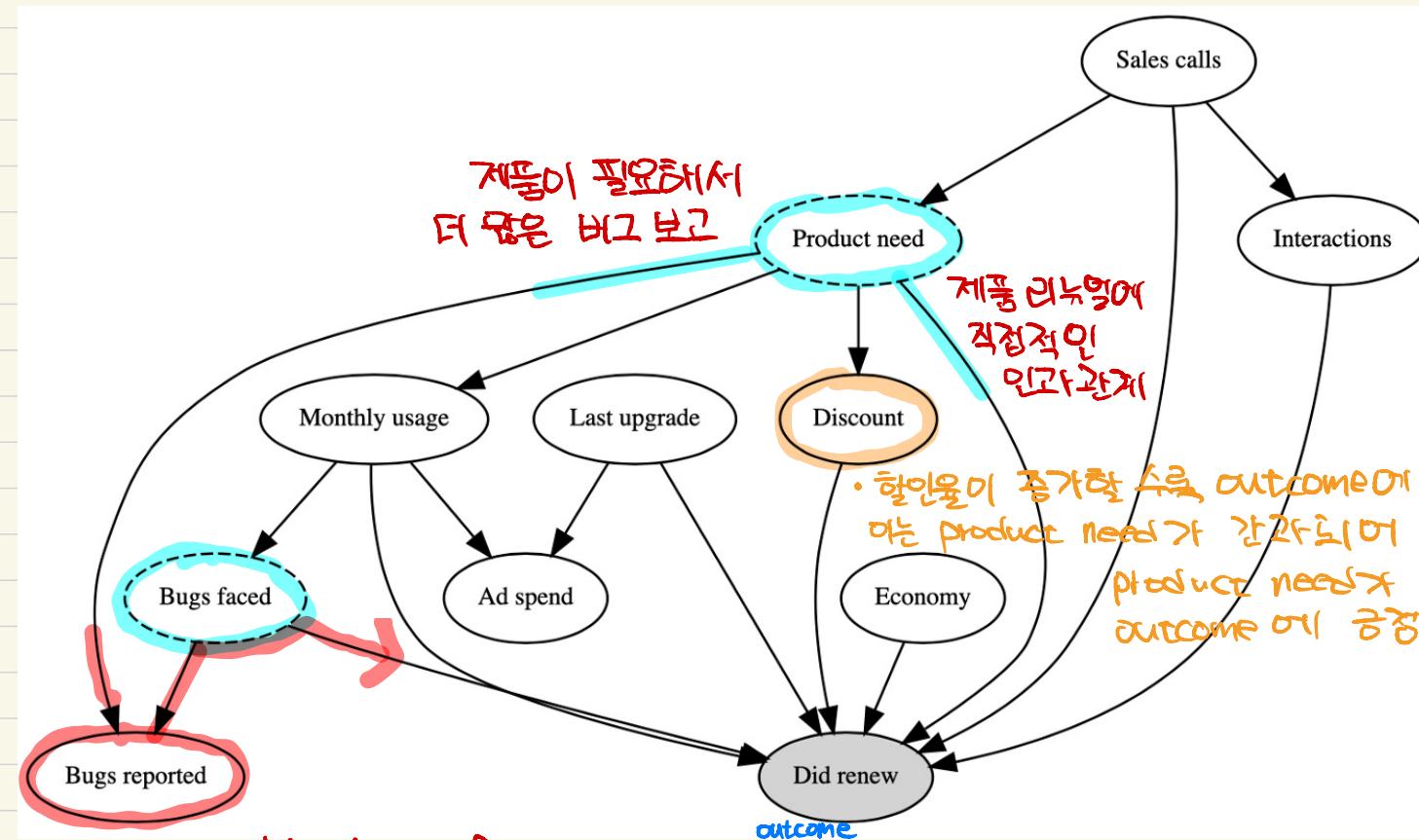
dashed

Unmeasured confounding feature



- SHAP 값
- 실제 인과 효과

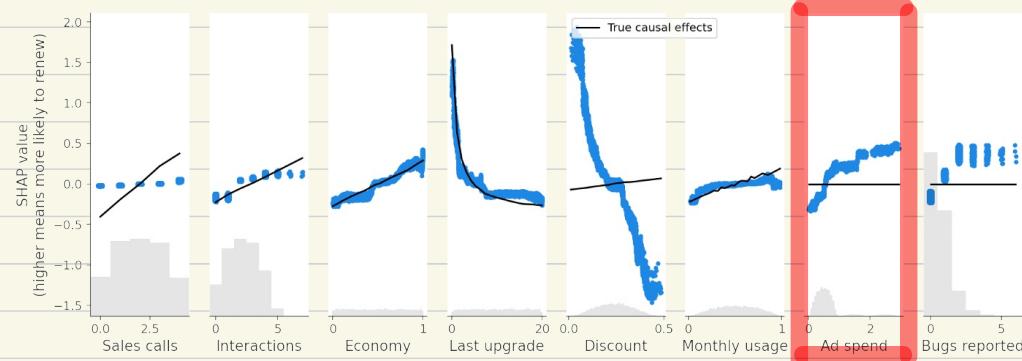
- ML 모델에서 bugs reported 영향 있는 것으로 보임
 - Discounts는 실제 인과관계로 반대 경향
- 관측되지 않은 교란 요인



부정적인 영향을 주나 발생한 문제로
변수를 바꾸면 할인율은
outcome에 긍정적인 영향을 줌

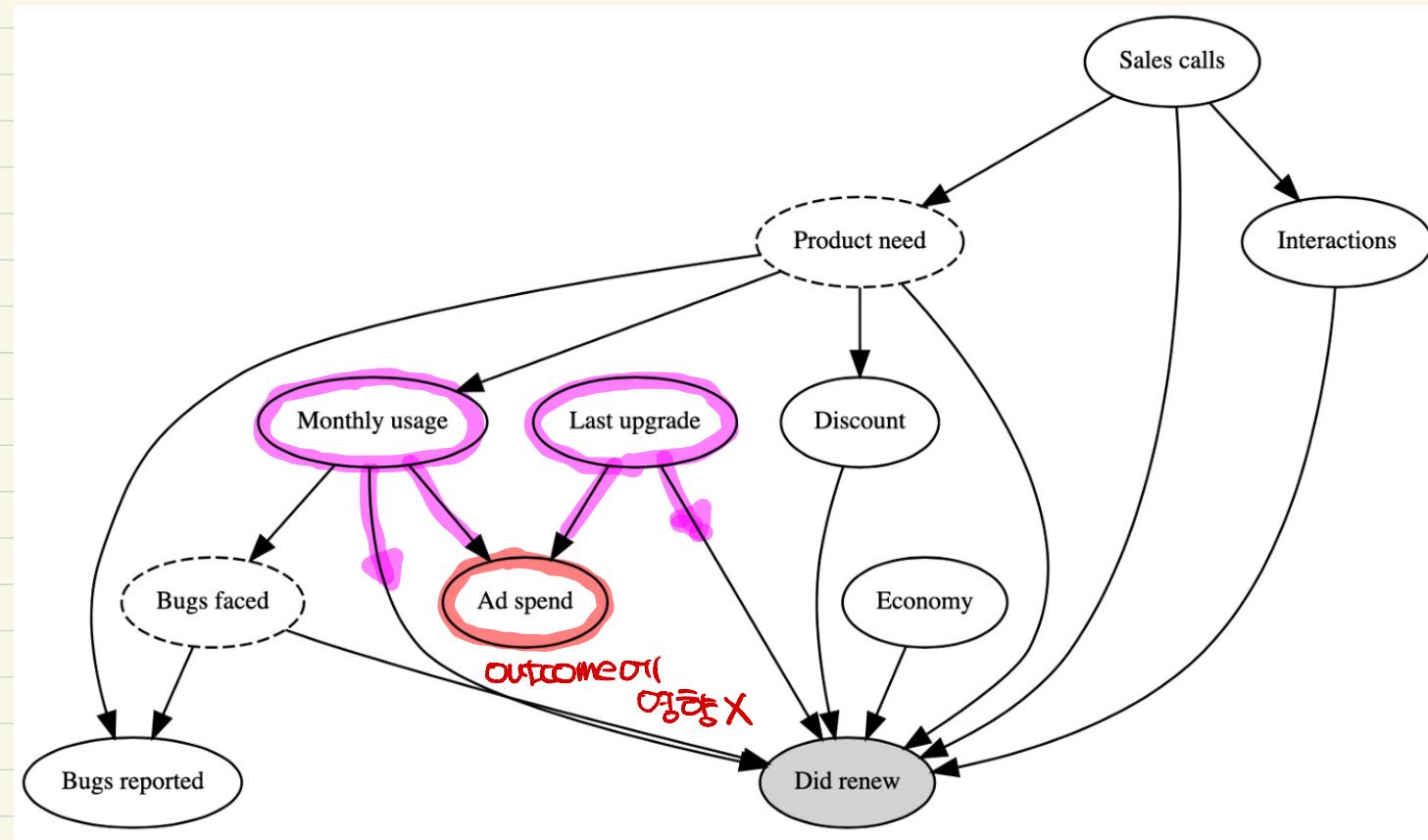
- 관측되지 않은 bugs faced로 인해 리뉴얼에 직접적인 인과관계로 인식
- 관측되지 않은 product need가 bugs reported에 영향을 받기 때문에 outcome에 큰 영향을 주는 것으로 오해

Measured confounding feature



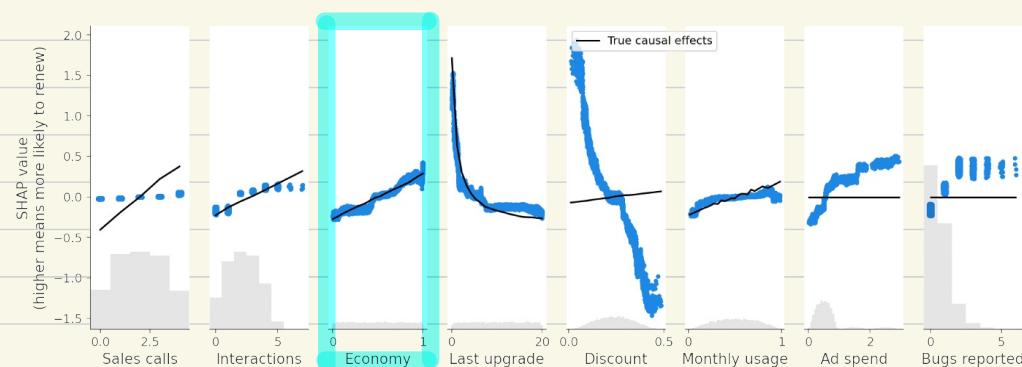
- SHAP 값
- 실제 인과 효과

- 실제로는 outcome에 영향이 없는 Ad spend가 ML 모델에서는 영향이 있는 것으로 보여짐

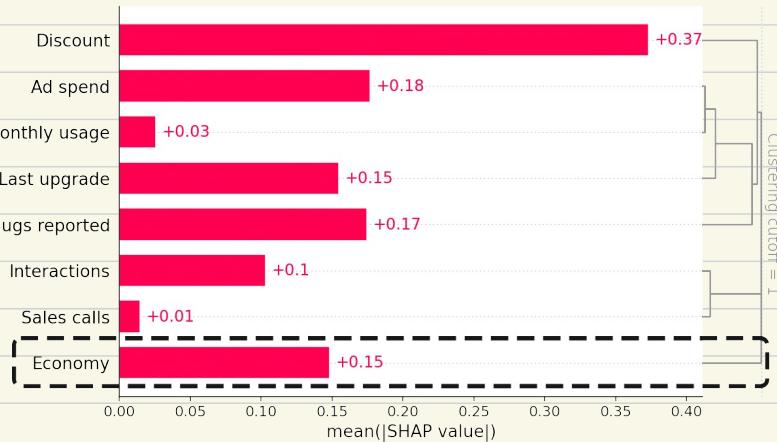
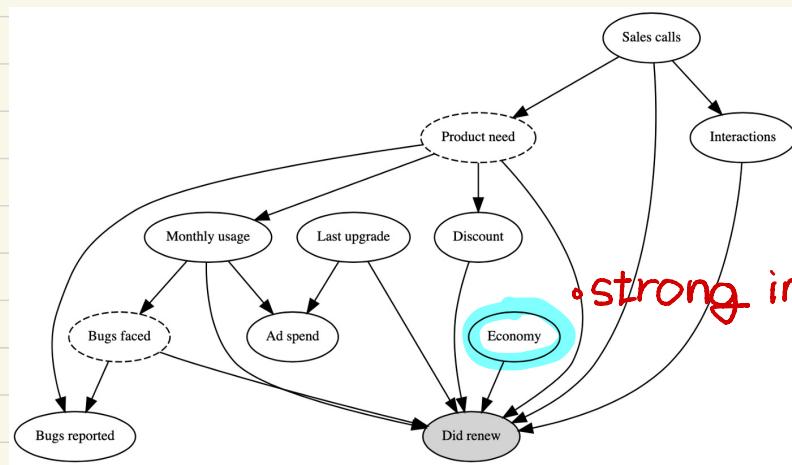


- Ad spend는 outcome에 영향이 없지만 outcome에 직접적인 인과성이 있는 Monthly usage와 Last upgrade에 영향을 주기에 의해 소지가 발생함

ML 모델이 인과성에 대한 답을 할 수 있는 경우

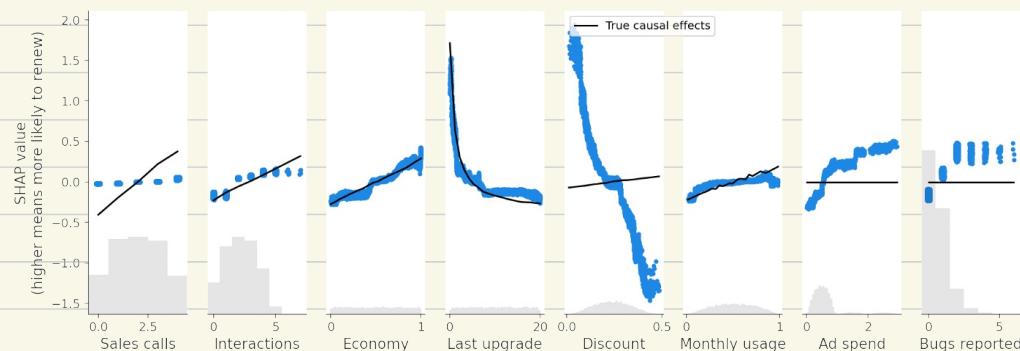


● SHAP 값
— 실제 인과 효과

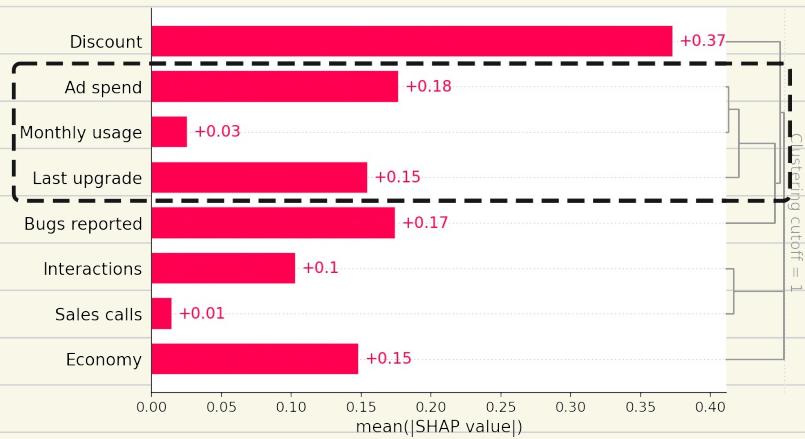
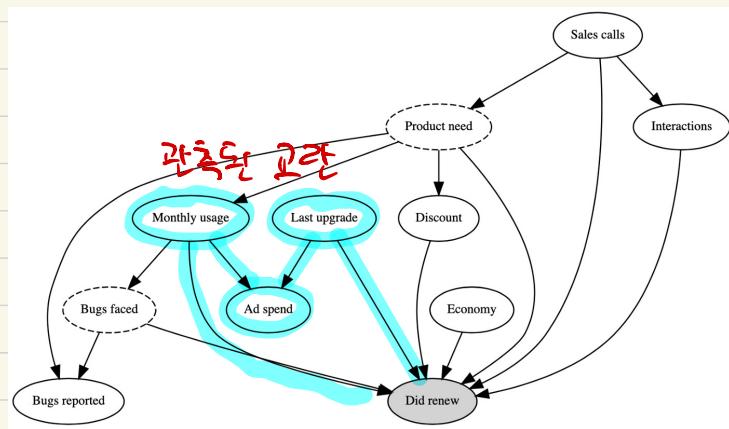


ML 모델이 인과성에 대한 답을 할 수 없지만 인과 추론에 도움이 될 수 있는 경우

- Observed confounding



- SHAP 값
- 실제 인과 효과



- Ad spend는 outcome에 직접적 영향은 없지만 outcome에 영향이 있는 monthly usage와 last upgrade와 상관성이 있음
- ML 모델은 똑같이 예측을 잘 할 수 있다며 과적합을 피하기 위해 상관성이 있는 여러 feature 중 한 개의 feature를 사용하려는 경향이 있음 → causal feature 대신 상관 관계가 있는 feature를 사용할 수 있음
- Ad spend 증가와 outcome의 상관관계(?) 보여 outcome을 증대하기 위해서 Ad spend를 증가시키기 원하는 결과 얻지 못함
- 이와 같은 관찰된 교란의 경우 ① 관심있는 feature의 교란을 해제하고 ② 그 feature 변경에 따른 평균 인과 효과(ex, 인과 효과의 평균 기울기)를 추정하고자 하는 ML 모델 사용
(double/debiased)

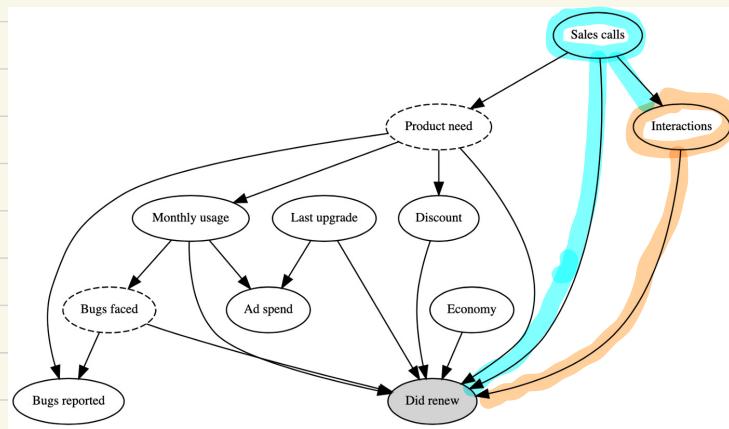
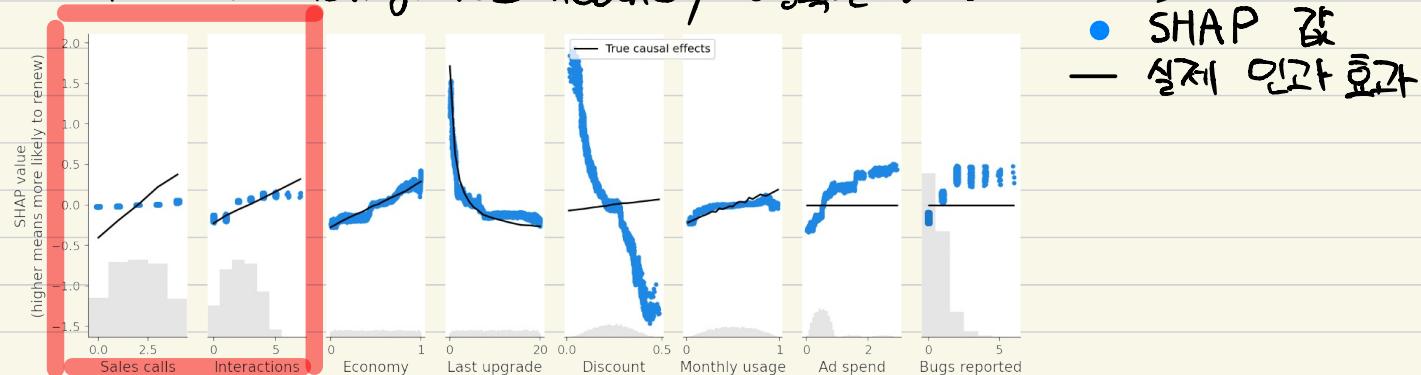
Double ML

1. Train a model to predict a feature of interest (i.e. Ad Spend) using a set of possible confounders (i.e. any features not caused by Ad Spend).
$$\text{interest feature} = \hat{f}_1(\text{confounder features})$$
2. Train a model to predict the outcome (i.e. Did Renew) using the same set of possible confounders.
$$\text{outcome} = \hat{f}_2(\text{confounder features})$$
3. Train a model to predict the residual variation of the outcome (the variation left after subtracting our prediction) using the residual variation of the causal feature of interest.

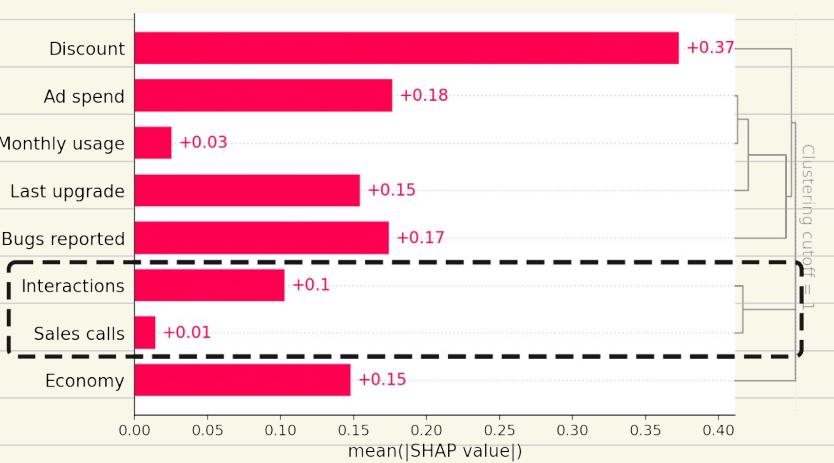
=> 이를 위한 패키지는 EconML, CausalML 등이 있음

ML 모델이 인과성에 대한 답을 할 수 없지만 인과 추론에 도움이 될 수 있는 경우

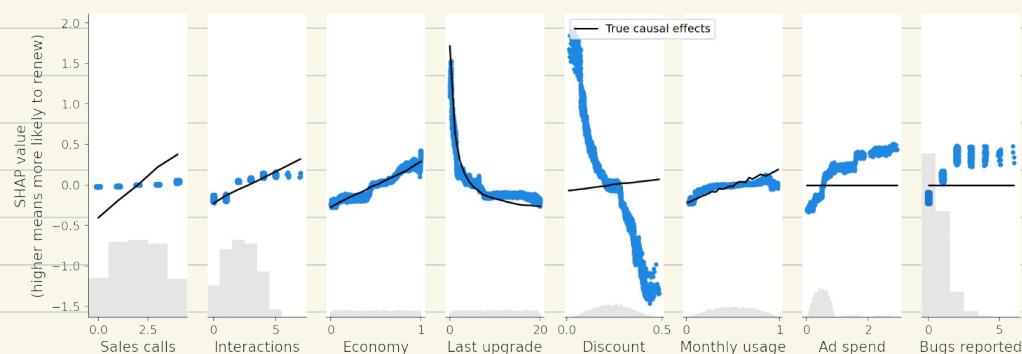
- Non-confounding redundancy (중복될 빠고만 인자)



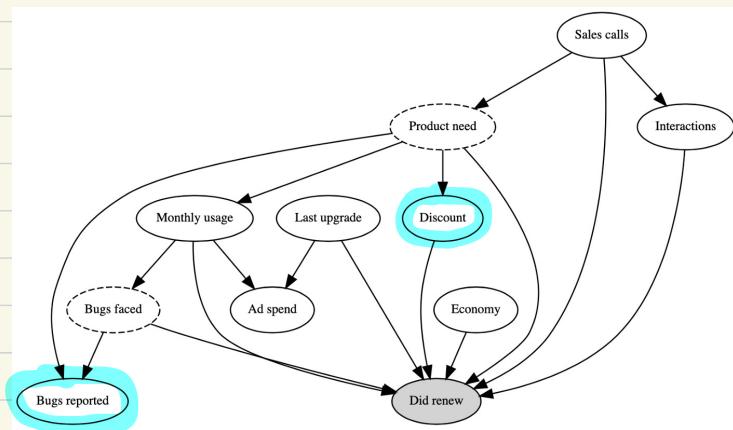
- Sales call은 outcome에 직접적인 원인이며, interaction을 통해서도 outcome에 영향을 미침
- ML 모델에 Sales call과 interaction이 모두 포함될 경우 causal effect가 분산됨
- 예시에서는 outcome의 직접적인 원인인 sales call 대신 interaction에 기여도가 분산됨
- Non-confounding redundancy는 중복될 feature를 제거하여 수정 가능



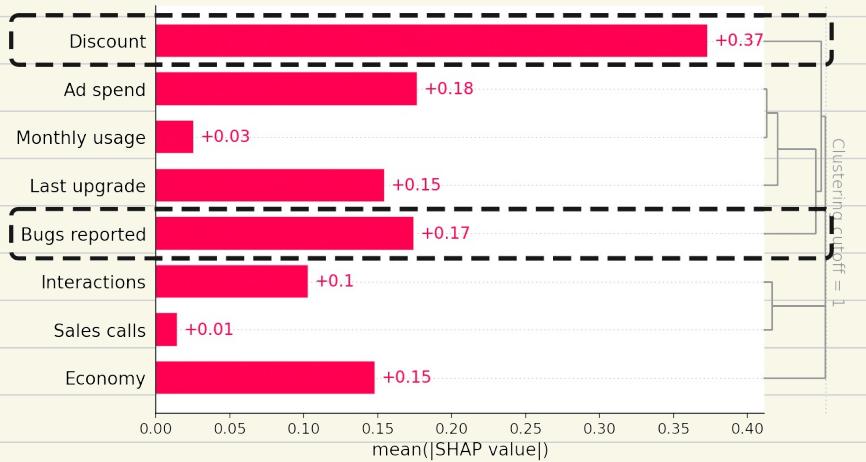
ML 모델이 인과성에 대한 답을 할 수 없으며 인과 추론에도 도움이 되지 않는 경우



- SHAP 값
- 실제 인과 효과



- Double ML 방법은 모든 고려 인자가 관측 가능한 경우에만 작동
- 관측되지 않은 고려 요인은 인과 추론 할 수 없음



Summary

Flexible predictive models like XGBoost or LightGBM are powerful tools for solving *prediction* problems. However, they are not inherently causal models, so interpreting them with SHAP will fail to accurately answer *causal* questions in many common situations. Unless features in a model are the result of experimental variation, applying SHAP to predictive models without considering confounding is generally not an appropriate tool to measure causal impacts used to inform policy. SHAP and other interpretability tools can be useful for causal inference, and SHAP is integrated into many causal inference packages, but those use cases are explicitly causal in nature. To that end, using the same data we would collect for prediction problems and using causal inference methods like double ML that are particularly designed to return causal effects is often a good approach for informing policy. In other situations, only an experiment or other source of randomization can really answer *what if* questions. Causal inference always requires us to make important assumptions. The main point of this article is that the assumptions we make by interpreting a normal predictive model as causal are often unrealistic.

=> 그럼에도 불구하고 ML 모델로 인과성을 설명하려면 ML 모델이 직접적인 인과를 배울 수 있도록 설계해야하지 않을까?

- non-confounding redundancy

3가지 항목

1. 기상/기후
2. 시간적 여유
3. 경영활동

→ 2011년과 2012년의
경영활동 차이는?

