



УДК 519.23:004.855

О.С. БАЛАБАНОВ

Інститут програмних систем НАН України, Київ, Україна,
e-mail: olexastep@gmail.com.

РЕГРЕСІЙНИЙ АНАЛІЗ ТА КАУЗАЛЬНІ МОДЕЛІ (ОГЛЯД)

Анотація. Розглянуто принципи застосування регресійного аналізу в прогнозно-аналітичних задачах і особливості прогнозування наслідків (ефектів) керування об'єктом моделювання. Виокремлено типи (групи) задач: «пасивний» прогноз цільової змінної; ретрореконструкція значення цільової змінної; прогноз каузального ефекту; оцінювання структурних коефіцієнтів та коефіцієнтів каузального впливу; діагностика та виявлення причино-наслідкових зв'язків; контрфактуальний аналіз. Стандартна процедура регресії призначена для пасивного прогнозу. Особливість задачі активного (каузального) прогнозування полягає в тому, що він стосується зміненого «світу», який виникає після втручання (примусового встановлення значень певних змінних). Модель об'єкта після втручання відрізняється від оригінальної моделі локальними деформаціями. В деяких обставинах для вироблення коректного активного прогнозу достатньо правильно вибрати підходящий набір регресорів. У складніших випадках це вимагає багатокрокових стратегій і процедур, які враховують знання про структуру каузальних відношень. Описуються, зокрема, узагальнена версія критерію «чорного ходу» (back-door), розширений метод інструментальної змінної для лінійних моделей тощо. Показано роль каузальної інформації та латентних конфаундерів.

Ключові слова: регресія, каузальний ефект, прогноз, регресійний коефіцієнт, каузальний зв'язок, підбір регресорів.

DOI 10.34229/KSA2522-9664.26.4.1

ВСТУП

Регресійний аналіз (РА) є одним з найбільш поширених засобів аналізу даних та емпіричних досліджень. У численних публікаціях з регресійного аналізу викладено техніку регресії, його математико-статистичні основи, тестування моделей, обчислення довірчих інтервалів, аналіз залишків тощо [1–9]. Тим часом, у постановці задачі та інтерпретації результатів деякі дослідники досі покладаються на інтуїцію, звичку або традиції в їхньому оточенні. Іноді навіть не декларують суттєві припущення. Сучасна методологія і тактики застосування РА недостатньо відображені у вітчизняній літературі. Широко відома теза «кореляція не є каузальністю» не дає конструктивних відповідей. Упродовж останніх трьох десятиліть відбувся революційний прогрес каузальних моделей та методів їхнього виведення на основі даних. Нині у найпрестижніших журналах і на наукових конференціях регулярно оприлюднюються результати досліджень з розвитку, адаптації та застосування РА в економетриці, соціології, біоінформатиці, медицині, епідеміології, у штучному інтелекті тощо. Начасі систематизувати свіжі результати, пояснити формулювання задач та інтерпретацію результатів. У цьому огляді зроблено спробу частково задовольнити вказану потребу. Розглянуто формування набору регресорів, тлумачення

© О.С. Балабанов, 2026