

Spis treści |

O autorze	15
O recenzentach	16
Podziękowania	17
Słowo wstępne	19
Przedmowa	21

CZĘŚĆ 1. Przyczynowość — wprowadzenie

ROZDZIAŁ 1

Związki przyczynowe? Przecież jest uczenie maszynowe, więc po co zawracać sobie tym głowę?	29
Krótka historia przyczynowości	30
Dlaczego przyczynowość? Zapytaj dzieci!	31
Interakcje ze światem	31
Zakłócenia — związki, które nie są prawdziwe	32
Jak nie stracić pieniędzy... i ludzkich istnień	35
Dylemat marketera	35
Pobawmy się w doktora!	36
Asocjacje w realnym świecie	37
Podsumowanie	37
Bibliografia	38

ROZDZIAŁ 2

Judea Pearl i drabina przyczynowości	39
Od asocjacji do logiki i wyobraźni. Drabina przyczynowości	39
Asocjacje	42
Ćwiczenie	44
Czym są interwencje?	47
Zmienianie świata	48
Korelacja i przyczynowość	50

Czym są kontrfakty?	52
Zanurzymy się w dziwactwa (zapis formalny)	53
Podstawowy problem wnioskowania przyczynowego	54
Obliczanie kontrfaktów	54
Czas na kodowanie!	56
Dodatek. Czym jest uczenie maszynowe z perspektywy przyczynowości? ...	57
Przyczynowość a uczenie ze wzmocnieniem	57
Przyczynowość a uczenie półnadzorowane i nienadzorowane	58
Podsumowanie	58
Bibliografia	59

ROZDZIAŁ 3

Regresja, obserwacje i interwencje	61
Wprowadzenie. Dane obserwacyjne a regresja liniowa	61
Regresja liniowa	62
Wartości p i istotność statystyczna	65
Interpretacja geometryczna regresji liniowej	66
Odwrocenie kolejności	67
Czy zawsze należy kontrolować wszystkie dostępne współzmiennne?	69
Poruszanie się po labiryncie	69
Jeśli nie wiesz, dokąd zmierzasz, możesz trafić gdzie indziej	70
Pójdźmy dalej!	73
Kontrolować czy nie kontrolować?	73
Modele regresyjne a modele strukturalne	73
Modele SCM	73
Regresja liniowa a modele SCM	74
Szukanie powiązania	74
Regresja a skutki przyczynowe	76
Podsumowanie	78
Bibliografia	78

ROZDZIAŁ 4

Modele grafów	79
Grafy, grafy, grafy	79
Rodzaje grafów	80
Reprezentacje grafów	83
Grafy w Pythonie	85
Czym jest model grafów?	87
Skierowane grafy acykliczne w świecie związków przyczynowych	88
Definicje przyczynowości	88
Grafy DAG a przyczynowość	89
Formalna definicja grafów DAG	90
Ograniczenia grafów DAG	90

Źródła grafów przyczynowych w świecie rzeczywistym	91
Odkrywanie związków przyczynowych	91
Wiedza ekspercka	91
Połączenie technik odkrywania związków przyczynowych i wiedzy eksperckiej	92
Dodatek. Czy można opisywać związki przyczynowe bez grafów DAG?	92
Układy dynamiczne	92
Cykliczne modele SCM	92
Podsumowanie	93
Bibliografia	93

ROZDZIAŁ 5

Rozwidlenia, łańcuchy i kolidery	95
Grafy i rozkłady oraz sposoby mapowania między nimi	95
Jak opisywać niezależność?	96
Wybór właściwego kierunku	97
Warunki i założenia	98
Łańcuchy, rozwidlenia i kolidery	102
Łańcuch zdarzeń	102
Łańcuchy	103
Rozwidlenia	104
Kolidery lub struktury v	106
Przypadki niejednoznaczne	108
Rozwidlenia, łańcuchy, kolidery i regresja	109
Tworzenie zbioru danych dla łańcucha	110
Tworzenie zestawu danych dla rozwidlenia	112
Tworzenie zbioru danych dla kolidera	112
Dopasowanie modeli regresji	113
Podsumowanie	116
Bibliografia	116

CZĘŚĆ 2. Wnioskowanie związków przyczynowych

ROZDZIAŁ 6

Węzły, krawędzie i statystyczna (nie)zależność	119
Zadbaj o separację d!	120
Trening czyni mistrza — separacja d	121
Najpierw estymandy!	124
Żyjemy w świecie estymatorów	124
Czym są estymandy?	124

Kryterium back-door	126
Czym jest kryterium back-door?	127
Kryterium back-door a estymandy równoważne	127
Kryterium front-door	130
Czy GPS może nas wyprowadzić na manowce?	130
Londyńskie taksówki i magiczny kamień	131
Otwarcie frontowych drzwi	132
Trzy proste kroki w kierunku kryterium front-door	134
Kryterium front-door w praktyce	134
Czy są jakieś inne kryteria? Zastosujmy rachunek do!	140
Trzy zasady rachunku do	141
Zmienne instrumentalne	142
Podsumowanie	144
Odpowiedź	144
Bibliografia	145

ROZDZIAŁ 7

Czteroetapowy proces wnioskowania przyczynowego	146
Wprowadzenie do bibliotek DoWhy i EconML	147
Ekosystem analizy przyczynowej Pythona	147
Dlaczego DoWhy?	149
Czym jest pakiet DoWhy?	150
A co z biblioteką EconML?	150
Krok 1. Modelowanie problemu	151
Utworzenie grafu	151
Tworzenie obiektu CausalModel	153
Krok 2. Identyfikacja estymand	154
Krok 3. Wyznaczanie oszacowań	156
Krok 4. Zestaw walidacyjny. Testy obalające	156
Jak walidować modele przyczynowe?	157
Wprowadzenie do testów obalających	158
Pełny przykład	160
Krok 1. Zakodowanie założeń	161
Krok 2. Wyznaczenie estymandy	163
Krok 3. Wyznaczenie oszacowania	163
Krok 4. Obalenie oszacowania	165
Podsumowanie	169
Bibliografia	169

ROZDZIAŁ 8

Modele przyczynowe. Założenia i wyzwania	170
Jestem królem świata! Czy rzeczywiście tak jest?	170
Gdzieś pośrodku	171
Identyfikowalność	172
Brak grafów przyczynowych	172
Za mało danych	173
Nieweryfikowalne założenia	175
Słoń w pokoju — nadzieja czy beznadzieja?	175
Zjedźmy słonia	175
Dodatniość	177
Wymiennność	179
Podmioty wymienne	180
Wymiennność a zakłócenia	180
...i inne	181
Modułowość	181
SUTVA	183
Spójność	183
Nazywaj mnie po imieniu — relacje pozorne	184
Nazwy, nazwy, nazwy	184
Czy powinienem zapytać Ciebie, czy kogoś, kogo tu nie ma?	185
Stwórzcie graf DAG!	185
Dodatkowe informacje o stronniczości wyboru	187
Podsumowanie	189
Bibliografia	189

ROZDZIAŁ 9

Wnioskowanie związków przyczynowych i uczenie maszynowe — od dopasowywania do metalearnerów	191
Podstawy I. Dopasowywanie	192
Rodzaje dopasowywania	192
Efekty interwencji — ATE w porównaniu z ATT i ATC	193
Estymatory dopasowywania	194
Implementacja dopasowywania	196
Podstawy II. Współczynniki skłonności	201
Dopasowywanie w praktyce	201
Zmniejszenie wymiarowości za pomocą współczynników skłonności ...	202
Dopasowywanie współczynników skłonności (PSM)	203
Odwrotne ważenie prawdopodobieństwa (IPW)	204
Wiele twarzy współczynników skłonności	204
Formalizacja techniki IPW	205
Implementacja IPW	205
IPW — względy praktyczne	206

S-Learner — samotny stróż	207
Diabeł tkwi w szczegółach	207
Mamo, tato, poznajcie CATE	208
Żarty na bok. Pozdrowienia dla heterogenicznego tłumu	209
Machanie flagą założeń	210
Jesteś jedyny. Modelowanie z wykorzystaniem techniki S-Learner	211
Dane o niewielkiej objętości	217
Słabe punkty modelu S-Learner	218
T-Learner. Razem możemy więcej	218
Wymuszenie właściwego podziału zmiennych	219
T-Learner w czterech krokach i wzory	219
Implementacja modelu T-Learner	220
X-Learner. Krok dalej	222
Wyciskanie cytryny	222
Rekonstrukcja modelu X-Learner	223
X-Learner. Formuła alternatywna	225
Implementacja X-Learner	226
Podsumowanie	230
Bibliografia	231

ROZDZIAŁ 10

Wnioskowanie związków przyczynowych i uczenie maszynowe — zaawansowane estymatory, eksperymenty, oceny i nie tylko

Metody DR. Spróbujmy uzyskać więcej!	234
Czy potrzebujemy czegoś więcej?	234
Podwójnie wzmocniony nie oznacza niezniszczalny... ..	236
...ale pozwala wiele zyskać	236
Sekretny, podwójnie mocny sos	236
Estymator DR a założenia	238
DR-Learner. Przechodzenie nad przepaścią	238
Modele DR-Learner — opcje dodatkowe	242
Ukierunkowany estymator maksymalnego prawdopodobieństwa	242
Jeśli uczenie maszynowe jest fajne, to co powiesz na podwójne uczenie maszynowe?	246
Dlaczego DML i co jest w nim podwójnego?	246
Implementacja DML za pomocą bibliotek DoWhy i EconML	249
Dostrajanie hiperparametrów za pomocą bibliotek DoWhy i EconML ...	252
Czy DML jest srebrną kulą?	256
Techniki DR a DML	258
Co z tego będę miał?	259

Lasy przyczynowe i nie tylko	260
Drzewa przyczynowe	260
Przepętnięcia lasów	260
Zalety lasów przyczynowych	260
Lasy przyczynowe z wykorzystaniem bibliotek DoWhy i EconML	261
Niejednorodne efekty interwencji z danymi eksperymentalnymi — odyseja upliftingu	263
Dane	263
Wybór frameworka	268
Nie znamy połowy tej historii	268
Wyzwanie Kevina	269
Otwarcie skrzynki z narzędziami	270
Modele uplift a wydajność	274
Inne wskaźniki dla wyników ciągłych z wieloma interwencjami	279
Przedziały ufności	279
Zwycięski wynik w wyzwaniu Kevina	280
Kiedy należy stosować estymatory CATE dla danych eksperymentalnych?	280
Wybór modelu. Uproszczony przewodnik	281
Dodatek. Objaśnienia kontrfaktyczne	283
Zła wola czy nieodpowiednia technologia?	283
Podsumowanie	284
Bibliografia	285

ROZDZIAŁ 11

Wnioskowanie związków przyczynowych i uczenie maszynowe — uczenie głębokie, przetwarzanie języka naturalnego i inne techniki 288

Wykorzystanie technik uczenia głębokiego do wyznaczania heterogenicznych efektów interwencji	289
Wskaźniki CATE sięgają głębiej	289
SNet	291
Transformatory i wnioskowanie związków przyczynowych	299
Teoria znaczenia w pięciu akapitach	299
Co zrobić, by komputery rozumiały język naturalny?	300
Od filozofii do kodu Pythona	301
Modele LLM a przyczynowość	301
Trzy scenariusze	303
CausalBert	306
Przyczynowość i szeregi czasowe, czyli kiedy ekonometryk przechodzi na Bayesa	312
Metody quasi-eksperymentalne	312
Przejście Twittera i wzorce googlowania	313

Logika syntetycznych kontroli	313
Wizualne wprowadzenie do logiki kontroli syntetycznej	314
Na początek dane	317
Kontrola syntetyczna w kodzie	317
Wyzwania	322
Podsumowanie	323
Bibliografia	324

CZĘŚĆ 3. Odkrywanie związków przyczynowych

ROZDZIAŁ 12

Czy można prosić o graf przyczynowy?	329
Źródła wiedzy przyczynowej	329
Zalew informacji	330
Siła zaskoczenia	330
Spostrzeżenia naukowe	331
Logika nauki	331
Hipotezy są gatunkiem	332
Jedna logika, wiele dróg	333
Eksperymenty kontrolowane	333
Randomizowane badania kontrolowane	334
Od eksperymentów do grafów	334
Symulacje	335
Osobiste doświadczenia i wiedza dziedzinowa	335
Osobiste doświadczenia	336
Wiedza dziedzinowa	337
Uczenie się struktury przyczynowej	337
Podsumowanie	338
Bibliografia	339

ROZDZIAŁ 13

Odkrywanie związków przyczynowych i uczenie maszynowe — od założeń do zastosowań	340
Odkrywanie związków przyczynowych — przypomnienie informacji o założeniach	341
Przygotowania	341
Należy zawsze dążyć do zapewnienia wierności... ..	341
...ale czasami to jest trudne	341
Minimalizm jest cnotą	342
Cztery (i pół) rodziny	342
Cztery strumienie	343

Wprowadzenie do pakietu gCastle	344
Witaj, gCastle!	344
Dane syntetyczne w gCastle	345
Dopasowywanie pierwszego modelu odkrywania związków przyczynowych	348
Wizualizacja modelu	349
Wskaźniki oceny modelu	350
Odkrywanie związków przyczynowych oparte na ograniczeniach	353
Ograniczenia i niezależność	353
Wykorzystanie struktury niezależności w celu odtworzenia grafu	354
Algorytm PC — ukryte wyzwania	357
Algorytm PC dla danych kategoryjnych	358
Odkrywanie związków przyczynowych na podstawie punktacji	359
Tabula rasa — zaczynamy od nowa	359
GES — punktacja	359
Algorytm GES w bibliotece gCastle	360
Funkcyjne odkrywanie związków przyczynowych	361
Błogosławieństwa asymetrii	361
Model ANM	362
Ocena niezależności	365
Czas na LiNGAM	366
Odkrywanie związków przyczynowych oparte na gradientach	372
Czym jest ten gradient?	372
Proszę nie ronić łez!	374
Nie płacz, GOLEM!	374
Porównanie	375
Kodowanie wiedzy eksperckiej	377
Czym jest wiedza ekspercka?	378
Wiedza ekspercka w bibliotece gCastle	378
Podsumowanie	379
Bibliografia	379

ROZDZIAŁ 14

Odkrywanie związków przyczynowych i uczenie maszynowe — zaawansowane uczenie głębokie i nie tylko

382

Zaawansowane odkrywanie związków przyczynowych za pomocą uczenia głębokiego	383
Od modeli generatywnych do przyczynowości	384
Spójrz wstecz, aby dowiedzieć się, kim jesteś	384
Elementy składowe frameworka DECI	385
Implementacja DECI	386
DECI to rozwiązanie kompleksowe	398

Odkrywanie związków przyczynowych w wypadku występowania ukrytych zakłóceń	398
Algorytm FCI	398
Inne podejścia do danych z zakłóceniami	403
Dodatek. Nie tylko obserwacje	404
ENCO	404
ABCI	404
Odkrywanie związków przyczynowych — praktyczne zastosowania, wyzwania i otwarte problemy	405
Podsumowanie	406
Bibliografia	407

ROZDZIAŁ 15

Epilog	409
Czego nauczyłeś się z tej książki?	409
Pięć kroków do jak najlepszego wykorzystania projektów przyczynowych	410
Zadaj pytanie	410
Zdobądź wiedzę ekspercką	411
Wygeneruj hipotetyczny graf (grafy)	412
Sprawdź identyfikowalność	412
Dokonaj falsyfikacji hipotez	413
Przyczynowość w biznesie	414
Jak eksperci analizy przyczynowej przechodzą od wizji do implementacji?	414
Przyszłość przyczynowego uczenia maszynowego	416
Gdzie jesteśmy dziś i dokąd zmierzamy?	417
Wskaźniki przyczynowe	418
Fuzja danych przyczynowych	418
Agenty interwencji	419
Uczenie się struktury przyczynowej	419
Uczenie się przez naśladowanie	420
Studiowanie przyczynowości	421
Pozostańmy w kontakcie	422
Podsumowanie	422
Bibliografia	423
Skorowidz	425