

Spis treści

Wstęp 9

1. Koncepcja szeregów czasowych 15

Szeregi czasowe w różnych dziedzinach - krótka historia 15

Szeregi czasowe w medycynie 16

Przewidywanie pogody 20

Prognozy rozwoju gospodarczego 21

Astronomia 23

Początki analizy szeregów czasowych 24

Metody statystyczne w analizie szeregów czasowych 25

Uczenie maszynowe w analizie szeregów czasowych 26

Zobacz też 27

2. Pozyskiwanie i przetwarzanie szeregów czasowych 29

Gdzie można znaleźć szeregi czasowe? 30

Gotowe zestawy danych 30

Odnajdywanie szeregów czasowych 36

Konstruowanie szeregu czasowego na podstawie danych tabelarycznych 37

Przygotowanie danych - instrukcja krok po kroku 38

Konstruowanie szeregu czasowego na podstawie zebranych danych 44

Problemy związane ze znacznikami czasu 46

Czego dotyczy dany znacznik? 46

Praca z danymi pozbawionymi dokumentacji 48

Co to jest znacząca skala czasu? 50

Oczyszczanie danych 50

Brakujące dane 51

Zmiana częstotliwości próbkowania 60

Wyładzanie danych 63

Wahania sezonowe 68

Strefy czasowe 71

Zapobieganie zjawisku lookahead 74

Zobacz też 76

3. Metody eksplorowania danych czasowych 79

Metody ogólnego przeznaczenia 79

Wykresy liniowe 80

Histogramy 82

Wykresy punktowe 84

Metody przeznaczone do eksploracji szeregów czasowych 86

O stacjonarności słów kilka 86

Stosowanie okien czasowych 90

Związki pomiędzy wartościami w szeregu 95

Korelacje pozorne 105

Przegląd użytecznych metod wizualizacji 107

Wizualizacje w jednym wymiarze 107

Wizualizacje w dwóch wymiarach 108

Wizualizacje w trzech wymiarach 114

Zobacz też 117

4. Symulacje szeregów czasowych 119

Czym wyróżniają się symulacje szeregów czasowych? 120

Symulacje kontra prognozy 120

Symulacje w implementacjach 121

Przykład 1. - zrób to sam 121

Przykład 2. - tworzenie świata symulacji, który sam sobą steruje 126

Przykład 3. - symulacja zjawiska fizycznego 132

Uwagi końcowe 137

Symulacje z wykorzystaniem metod statystycznych 138

Symulacje z wykorzystaniem uczenia głębokiego 138

Zobacz też 138

5. Przechowywanie danych czasowych 141

Definiowanie wymagań 143

Dane rzeczywiste a dane przechowywane 144

Bazy danych 146

SQL kontra NoSQL 147

Przegląd popularnych rozwiązań bazodanowych dla szeregów czasowych 149

Przechowywanie danych w plikach 153

NumPy 154

Pandas 155

Odpowiedniki w środowisku R 155

Xarray 156

Zobacz też 157

6. Modele statystyczne 159

Dlaczego nie należy korzystać z regresji liniowej? 159

Metody statystyczne dla szeregów czasowych 161

Modele autoregresyjne 161

Modele ze średnią ruchomą 174

Zintegrowane modele autoregresyjne średniej ruchomej 178

Model wektorowej autoregresji 187

Inne modele 191

Zalety i wady modeli statystycznych 192

Zobacz też 193

7. Modele zmiennych stanu 195

Wady i zalety modeli zmiennych stanu 196

Filtr Kalmana 197

Model 197

Implementacja 199

Ukryte modele Markowa 203

Sposób działania modelu 204

Dopasowywanie modelu 205

Implementacja dopasowania modelu 208

Bayesowskie strukturalne szeregi czasowe (BSTS) 213

Implementacja 214

Zobacz też 218

8. Generowanie i selekcja cech 221

Przykład wprowadzający 222

Ogólne uwagi dotyczące cech 223

Natura danego szeregu 223

Wiedza dziedzinowa 224

Parametry zewnętrzne 225

Przegląd miejsc, w których można szukać inspiracji dotyczących wyboru cech 225

Biblioteki dostępne na licencji open source 226

Przykłady cech powiązanych z konkretnymi dziedzinami 230

Jak dokonać selekcji cech po ich wygenerowaniu? 233

Podsumowanie i wnioski 236

Zobacz też 236

9. Uczenie maszynowe w analizie szeregów czasowych 239

Klasyfikacja szeregów czasowych 240

Generowanie i selekcja cech 240

Drzewa decyzyjne 243

Klasteryzacja 250

Generowanie cech 251

Metryki uwzględniające zmianę czasu 258

Klasteryzacja w kodzie 262

Zobacz też 264

10. Uczenie głębokie 267

Geneza uczenia głębokiego 269

Implementacja sieci neuronowej 271

Dane, symbole, operacje, warstwy i grafy 272

Budowa potoku uczenia 275

Spojrzenie na zestaw danych 275

Elementy potoku uczenia 278

Jednokierunkowe sieci neuronowe 293

Prosty przykład 293

Wykorzystanie modelu uwagi do uczynienia jednokierunkowych sieci bardziej świadomymi czasu 296

Konwolucyjne sieci neuronowe 298

Prosty model sieci konwolucyjnej 300

Alternatywne modele konwolucyjne 302

Rekurencyjne sieci neuronowe 304

Kontynuacja przykładu z zapotrzebowaniem na prąd 307

Autoenkoder 308

Połączenie architektur 309

Podsumowanie 313

Zobacz też 314

11. Pomiar błędów 317

Podstawy: jak przetestować prognozę? 318

Weryfikacja historyczna a kwestie związane z konkretnym modelem 320

Kiedy prognoza jest wystarczająco dobra? 321

Szacowanie niepewności modelu w oparciu o symulację 323

Prognozowanie na wiele kroków naprzód 326

Bezpośrednie dopasowanie do danego horyzontu 326

Podjęcie rekurencyjne do odległych horyzontów czasowych 326

Uczenie wielozadaniowe w kontekście szeregów czasowych 327

Pułapki walidacji 327

Zobacz też 328

12. Kwestie wydajnościowe w dopasowywaniu i wdrażaniu modeli 331

Praca z narzędziami przeznaczonymi do bardziej ogólnych przypadków użycia 332

Modele zbudowane z myślą o danych przekrojowych nie "współdzielą" danych pomiędzy próbkami 332

Modele, które nie wspierają wcześniejszego obliczania, tworzą niepotrzebne opóźnienia pomiędzy pomiarem a prognozowaniem 334

Wady i zalety formatów zapisu danych 334

Przechowuj dane w formacie binarnym 335

Przetwarzaj dane w sposób umożliwiający "przesuwanie się" po nich 335

Modyfikacje analizy dla zwiększenia jej wydajności 336

Wykorzystanie wszystkich danych to niekoniecznie najlepszy pomysł 336

Złożone modele nie zawsze sprawdzają się znacznie lepiej 337

Krótki przegląd innych wysokowydajnych narzędzi 338

Zobacz też 338

13. Zastosowania w obszarze opieki zdrowotnej 341

Przewidywanie grypy 341

Studium przypadku grypy w jednym obszarze metropolitalnym 341

Jak obecnie wygląda prognozowanie grypy? 354

Przewidywanie stężenia cukru we krwi 356

Eksploatacja danych i ich oczyszczanie 357

Generowanie cech 361

Dopasowanie modelu 366

Zobacz też 371

14. Zastosowania w obszarze finansów 373

Pozyskiwanie i eksploatacja danych finansowych 374

Wstępne przetwarzanie danych do uczenia głębokiego 380

Dodawanie interesujących nas wielkości do surowych danych 380

Skalowanie interesujących nas wielkości bez wprowadzania zjawiska lookahead 381

Formatowanie danych do sieci neuronowej 383

Budowanie i uczenie rekurencyjnej sieci neuronowej 386

Zobacz też 392

15. Szeregi czasowe w danych rządowych 393

Pozyskiwanie danych rządowych 394

Eksploatacja dużych zbiorów danych czasowych 395

Zwiększenie częstotliwości próbkowania i agregowanie danych podczas iteracji 399

Sortowanie danych 399

Statystyczna analiza szeregów czasowych "w locie" 403

Pozostałe pytania 412

Dalsze możliwości poprawy 413

Zobacz też 413

16. Pakiety przeznaczone do pracy z szeregami czasowymi 415

Prognozowanie na dużą skalę 415

Wewnętrzne narzędzia Google'a do przemysłowego prognozowania 416

Otwartoźródłowy pakiet Prophet od Facebooka 418

Wykrywanie anomalii 422

Otwartoźródłowy pakiet AnomalyDetection od Twittera 422

Inne pakiety stworzone z myślą o szeregach czasowych 425

Zobacz też 426

17. Prognozy o prognozowaniu 427

Prognozowanie jako usługa 427

Uczenie głębokie zwiększa możliwości probabilistyczne 428

Wzrost znaczenia uczenia maszynowego kosztem statystyki 429

Wzrost popularności metod łączących podejście statystyczne i uczenie maszynowe 429

Więcej prognoz dotyczących życia codziennego 430