

Spis treści

1	Wprowadzenie	7
2	Dane	12
2.1	Natura danych	12
2.1.1	Niepewność danych	15
2.1.2	Ziarna informacji	16
2.2	Odległość i podobieństwo	17
2.2.1	Miary odległości i podobieństwa dla cech ilościowych	18
2.2.2	Miary odległości i podobieństwa dla cech jakościowych, i mieszanych	20
2.2.3	Miary odległości i podobieństwa dla danych tekstowych	21
2.3	Pojęcie wyjątku	25
2.3.1	Przyczyny powstania wyjątków	25
2.3.2	Definicja wyjątku	26
2.3.3	Obszar wyjątkowy	29
2.4	Rodzaje wyjątków	32
2.4.1	Wyjątek punktowy	32
2.4.2	Wyjątek kontekstowy	32
2.4.3	Wyjątek zbiorowy	33
2.5	Podsumowanie	34
3	Metody statystyczne, odległościowe i gęstościowe	36
3.1	Detekcja wyjątków przy użyciu regresji liniowej	36
3.1.1	Rodzaje obserwacji wyjątkowych w regresji	38
3.1.2	Miary oceniające wpływ obserwacji na prostą regresji	41
3.2	Algorytm maksymalizacji wartości oczekiwanej	43
3.3	Metody oparte na prawdopodobieństwie	43
3.3.1	Klasyfikator bayesowski	44
3.3.2	Naiwny klasyfikator bayesowski	46

3.4	Metody oparte na odległości	50
3.4.1	Algorytm k -najbliższych sąsiadów	50
3.4.2	Algorytm zagnieżdżonych pętli	52
3.4.3	Algorytm grawitacyjny	52
3.5	Metoda wektorów nośnych	56
3.6	Metody porównywania kierunków	59
3.7	Metody oparte na gęstości	63
3.7.1	Algorytm lokalnego współczynnika wyjątku	63
3.7.2	Algorytm współczynnika łączenia wyjątków	68
3.7.3	Algorytm przestrzennego grupowania gęstości	70
3.8	Metody oparte na podziale	75
3.8.1	Algorytm k -średnich	75
3.8.2	Ulepszony algorytm genetyczny k -średnich	77
3.8.3	Algorytm losowego przeszukiwania	78
3.9	Metody hierarchiczne	79
3.9.1	Algorytm hierarchicznego zrównoważonego redukowania	79
3.9.2	Aglomeracyjny hierarchiczny algorytm obiektów reprezentatywnych	82
3.10	Podsumowanie	83
4	Zastosowanie podsumowań lingwistycznych	86
4.1	Podstawowe definicje zbiorów rozmytych	87
4.1.1	Przedziałowe zbiory rozmyte	87
4.1.2	Intuicjonistyczne zbiory rozmyte	88
4.2	Kwantyfikatory jako zbiory wartości lingwistycznych	89
4.3	Definicje wyjątku dla kwantyfikatorów lingwistycznych	92
4.4	Podsumowania lingwistyczne	94
4.4.1	Podsumowania lingwistyczne z klasycznymi zbiorami rozmytymi	94
4.4.2	Podsumowania lingwistyczne z przedziałowymi zbiorami rozmytymi	96
4.5	Definicja wyjątku w podsumowaniach lingwistycznych	101
4.5.1	Procedura wykrywania wyjątku przy użyciu podsumowań lingwistycznych	102
4.5.2	Dobór zbioru zmiennych lingwistycznych a wykrywanie wyjątków	103
4.5.3	Miary podsumowań lingwistycznych w świetle wykrywania wyjątków	106
4.6	Podsumowanie	110

5	Przykłady użycia zbiorów przybliżonych	112
5.1	Podstawowe pojęcia teorii zbiorów przybliżonych	112
5.2	Reguły decyzyjne	118
5.3	Indukcja reguł	119
5.4	Definicja wyjątku oparta na zbiorach przybliżonych	121
5.5	Wyjątki reguł decyzyjnych	122
5.6	Podsumowanie	124
6	Wielokryterialne wykrywanie wyjątków	125
6.1	Zagadnienia podstawowe	126
6.1.1	Wielokryterialny wektorowy algorytm genetyczny	129
6.1.2	Wielokryterialny algorytm ewolucyjny	130
6.1.3	Algorytm genetyczny niezdominowanego sortowania	131
6.1.4	Ewolucyjna optymalizacja Pareto	133
6.2	Wielokryterialne wykrywanie wyjątków	134
6.3	Podsumowanie	136
7	Wnioskowanie o istnieniu przypadków wyjątkowych	138
7.1	Definicja przypadku w metodzie CBR	141
7.2	Definicja wyjątku oparta na przypadkach	142
7.3	Klasyfikacja wyjątku oparta na przypadkach	143
7.3.1	Klasyfikator bayesowski oparty na przypadkach w wykrywaniu wyjątków	145
7.3.2	Algorytm k -najbliższych sąsiadów oparty na przykładach w wykrywaniu wyjątków	148
7.3.3	Regresja liniowa oparta na przypadkach w wykrywaniu wyjątków	149
7.4	Podsumowanie	150
8	Wykrywanie wyjątków w strumieniach danych	151
8.1	Definicje strumieni danych	152
8.1.1	Określanie podobieństwa sekwencji strumieni . . .	154
8.1.2	Miara dynamicznego dopasowania czasu	155
8.2	Detekcja wyjątków modelami stochastycznymi	156
8.3	Detekcja wyjątków metodami opartymi na odległości . . .	157
8.3.1	Nieparametryczna metoda Vinturniego	158
8.3.2	Hybrydowa metoda sezonowo ekstremalnych odchyłeń	159
8.3.3	Detekcja wyjątków oparta na regresji	161

8.3.4	Wykrywanie wyjątków w strumieniach opisanych zbiorem rozmytym	162
8.4	Podsumowanie	164
9	Praktyczne przykłady wykrywania wyjątków	165
9.1	Wykrywanie wartości odstających przy użyciu regresji . .	165
9.2	Przykład użycia naiwnego klasyfikatora bayesowskiego . .	166
9.3	Wybrane metody odległościowe oraz gęstościowe	168
9.4	Przykład zastosowania w medycynie	170
9.5	Detekcja wyjątków przy użyciu podsumowań lingwistycznych	176
9.5.1	Wykrywanie wyjątków dla atrybutów numerycznych	177
9.5.2	Detekcja wyjątków dla atrybutów tekstowych . . .	180
9.6	Podejścia wielokryterialne z użyciem NSGA-II oraz PESA-II	182
9.6.1	Przykład nr 1	182
9.6.2	Przykład nr 2 - zastosowanie w medycynie	184
9.7	Zastosowanie metody CBR	187
9.8	Przykłady wykrywania wyjątków w strumieniach danych	189
10	Wnioski końcowe	193
	Bibliografia	196