

■ 目 次

■ シリーズの刊行にあたって	iii
■ まえがき	iv

Chapter 1	第 1 章 ウェブと機械学習	1
	1.1 はじめに	1
	1.2 ウェブから社会を観測する	2
	1.3 データ分析の段階	5
	1.4 ウェブデータの特徴	6
	1.5 ウェブデータの種類	9
	1.6 ウェブデータで機械学習を行う際の課題	11
	1.6.1 データの量に関する課題	12
	1.6.2 データの質に関する課題	14
	1.6.3 学習データの多様性に関する課題	17
1.6.4 プライバシーに関する課題	17	

Chapter 2	第 2 章 バースト検出	19
	2.1 はじめに	19
	2.2 移動平均線収束拡散法	21
	2.3 ポアソン過程	27
	2.4 Kleinberg のバースト検出	30
	2.5 まとめ	34

Chapter 3	第 3 章 評判分類の学習	35
	3.1 評判分類	35
	3.2 素性抽出	36
	3.3 素性選択	39

3.4	素性の値	42
3.5	評判分類器の学習	44
3.6	ロジスティック回帰による評判分類器学習	46
3.6.1	ロジスティック回帰モデル	46
3.6.2	確率的勾配法	52
3.6.3	バイアス項	55
3.6.4	過学習と正則化	56
3.7	多値評判分類学習	63
3.8	評判分類の評価	64
3.8.1	二値評判分類の評価	65
3.8.2	多値評判分類の評価	65
3.9	評判情報辞書	68
3.10	評判分類における分野適応	70
3.11	構造対応学習	72
3.11.1	ピボット選択	73
3.11.2	ピボット予測	75
3.11.3	適用先ドメインの評判分類	76
3.11.4	ドメイン適応に適したドメインの選び方	79
3.12	まとめ	81

第 4 章	意味表現の学習	83
4.1	意味表現	83
4.2	分布的意味表現	84
4.3	分散的意味表現	89
4.4	連続単語袋詰めモデル	93
4.5	連続スキップグラムモデル	97
4.5.1	モデルの構成	97
4.5.2	連続単語袋詰めモデルと連続スキップグラムモデルの最適化	99
4.5.3	負例サンプリング	102
4.5.4	階層型ソフトマックスによる近似計算	106
4.6	大域ベクトル予測モデル	109

4.7	意味表現の評価	113
4.7.1	意味的類似性予測タスク	114
4.7.2	関係類似性予測タスク	117
4.8	単語の意味表現ベクトルの可視化	119
4.9	分散の意味表現における行列分解	120
4.10	まとめ	123

Chapter 5

第 5 章	グラフデータの機械学習	125
5.1	リンク構造に基づくデータマイニング	125
5.2	グラフの定義	126
5.3	ページランク	128
5.4	パーソナライズド・ページランク	131
5.5	ラベル拡散法	132
5.6	チェイランク	133
5.7	ページランクの応用例：スパムページの検出	134
5.8	HITS	135
5.9	シムランク	136
5.10	まとめ	137

Chapter 6

第 6 章	順序学習	139
6.1	検索エンジンと順序学習	139
6.2	静的順序と動的順序	141
6.3	順序学習のための素性	142
6.4	順序学習手法の分類	145
6.5	点順序学習手法	146
6.6	対順序学習手法	150
6.7	リスト順序学習手法	152
6.8	まとめ	155

付録 A	157
A.1 スカラー値をベクトルで微分	157
A.2 内積を片方のベクトルで微分	157
A.3 ℓ_2 ノルムの二乗をベクトルで微分	158
A.4 行列の特異値分解による行列近似	158
A.5 ソフトマックス関数	162
■ 参考文献	165
■ 索引	173