

目 次

第 1 章	近年の犯罪と計量的文体分析への期待	1
1.1	サイバー犯罪の出現と犯罪捜査	1
1.2	「パソコン遠隔操作事件」解決に向けた新たなツール	4
1.2.1	事件の概要と流れ	4
1.2.2	犯人性立証の可能性と誤認逮捕の防止	6
1.3	印字文書の分析例	7
1.4	遺体無き殺人，電子メールを用いた被害者なりすまし	9
1.5	遺書の分析	10
第 2 章	文体分析に関連する学問分野	13
2.1	法言語学	13
2.1.1	法言語学とは	13
2.2	計量国語学と計量言語学，ならびにコーパス言語学	16
2.3	計量文献学	17
2.3.1	計量文献学とは	17
2.3.2	諸外国における研究史	17
2.3.3	日本における研究史	19
2.4	計量文体学	20
2.5	社会言語学	21
2.5.1	社会言語学とは	21
2.5.2	性別と言語に関する研究	22
2.5.3	年齢と言語に関する研究	23
第 3 章	テキストマイニング概要	25
3.1	テキストマイニングとは	25

3.1.1	データサイエンスとテキストマイニング	25
3.1.2	テキストマイニングの応用研究	27
3.1.3	テキストマイニングの分析の流れ	29
3.2	テキストの加工作業とソフトウェア	32
3.2.1	形態素解析	32
3.2.2	構文解析	35
3.2.3	テキストマイニングのためのソフトウェア	36
3.3	文体的特徴	36
3.3.1	文字単位	37
3.3.2	形態素単位	41
3.3.3	構文単位	48
3.3.4	文体的特徴における経年変化	49
3.4	多変量データ解析	50
3.4.1	主成分分析	51
3.4.2	対応分析	52
3.4.3	多次元尺度法	55
3.4.4	階層的クラスター分析	59
3.4.5	データセットのテキストと変数の数	64
3.4.6	多変量データ解析を用いた日本語研究の変遷	64
3.5	機械学習	65
3.5.1	サポートベクターマシン	65
3.5.2	決定木	69
3.5.3	ランダムフォレスト	71
3.5.4	ナイーブベイズ	74
3.5.5	その他の機械学習（ニューラルネットワーク、k 最近傍法）	77
3.6	推定成績や識別力に関する評価方法や指標	78
3.6.1	交差検証法	78
3.6.2	分類器の性能評価（正解率、再現率、適合率、F 値）	80
3.6.3	感度と特異度、ROC 分析に基づく AUC	82
3.6.4	効果量	86

3.6.5	ランダムフォレストによる評価指標	87
第 4 章	法科学における文体分析の概要	89
4.1	諸外国における事例紹介	89
4.1.1	米国の爆弾魔ユナボマーの犯行声明文	89
4.1.2	米国の「パトリシア・ハースト誘拐事件」	91
4.1.3	英国における遺体無き殺人事件	93
4.2	日本における事例：保険金詐取目的の殺人事件	94
4.3	計量的文体分析の種別	97
4.3.1	著者照合	98
4.3.2	著者同定	98
4.3.3	著者プロファイリング	98
第 5 章	著者識別（著者照合、著者同定）	99
5.1	従来の筆跡鑑定	99
5.1.1	筆跡鑑定とは	99
5.1.2	個人差と個人内恒常性	101
5.1.3	作為筆跡（模倣と韜晦）	102
5.2	機械学習による著者識別	103
5.2.1	機械学習を用いた著者識別研究	103
5.2.2	機械学習による方法の問題点	105
5.3	多変量データ解析による著者識別	106
5.3.1	多変量データ解析による著者識別研究	106
5.3.2	「パソコン遠隔操作事件」の犯人性立証への計量的文体分 析の試み（調査研究 1）	107
5.3.3	実際の事件で用いられた文章に関する著者識別の検討（調 査研究 2）	111
5.3.4	多変量データ解析の著者識別に関する考察ならびに問題 点とその方策	120
5.3.5	分析結果に対するスコアリングルールの導入	122

5.3.6	スコアリングルールの検証（その1）およびテキスト数や文字数の影響（調査研究3）	125
5.3.7	スコアリングルールの検証（その2）（調査研究4）	133
5.3.8	著者識別手法の標準化と正確性の検証（調査研究5）	137
5.4	尤度比による著者識別	142
5.5	作為的に自己の文章表現を変えた文章の分析	144
5.5.1	模倣文章：「グリコ・森永事件」を模倣した「黒子のバスケ脅迫事件」の文章の判別（調査研究6）	144
5.5.2	韜晦文章：自己の文章表現を隠蔽した文章	151
第6章	著者プロファイリング	155
6.1	犯罪者プロファイリングとは	155
6.2	著者プロファイリング研究概要	156
6.2.1	性別推定研究	156
6.2.2	年齢層推定研究	158
6.2.3	その他の著者特徴の推定研究	160
6.3	性別の推定	161
6.3.1	機械学習による性別推定の試み（調査研究7）	161
6.3.2	性別を偽装した文章の文体的特徴の変化（調査研究8）	165
6.4	年齢層の推定	173
6.4.1	機械学習による年齢層推定の試み（調査研究9）	173
第7章	テキストマイニングを応用した犯行動機分析	179
7.1	犯罪者プロファイリングにおける動機分析	179
7.2	殺人事件	179
7.2.1	殺人事件の動機研究	179
7.2.2	殺人事件の動機分類（調査研究10）	181
7.3	放火事件	183
7.3.1	放火事件の動機研究	183
7.3.2	単一放火の動機分類（調査研究11）	185

7.3.3	連続放火の動機分類（調査研究 12）	187
7.3.4	総合考察	190
第 8 章	今後を見据えて	193
8.1	日本の法科学における新たな分析手法としての確立に向けて	193
8.2	科学鑑定としての評価基準	195
8.3	鑑定結論に至る根拠ならびに鑑定結論の表現方法	197
8.3.1	鑑定結論に至る根拠	197
8.3.2	鑑定結論の表現方法	198
8.4	今後の実務上ならびに研究における課題	199
	引用文献	203
	付 録	217
	索 引	219