

概要

本研究では、観光開発の支援技術として、観光地におけるポジティブな体験談をブログから収集し、分類する手法を提案する。ブログからポジティブな体験談を表す文を判定するために、情緒推定を行う。得られた体験談は、用言の意味的な属性、および、観光地に関する特徴的な名詞に基づいて分類する。用言の意味的な属性は、日本語語彙大系 [7] の用言意味属性を用いる。観光地における名詞の特徴度は、Okapi BM25 のスコアを用いる。

本手法を用いて、開発目標地に類似した観光地を調査することで、開発目標地に欠けている観光のポイントを発見することをねらう。開発目標地を「山陰海岸」、類似する観光地を「江ノ島海岸」、「若狭湾」、および、「三陸海岸」に設定し、類似する観光地からポジティブな体験談を収集、分類した。分類結果を人手分析することで、開発目標地に適用可能であると考えられるポイントを取得することができた。

さらに、体験談を分類した結果が、人手による分析の支援となっているか評価するため、名詞の特徴度と観光地開発のヒントとなる情報の出現数との関係を調査した。その結果、特徴度が高い情報ほど、新しい観光地開発案につながりやすいことを確認した。

目次

第1章	はじめに	1
第2章	関連研究	2
2.1	Web データを対象とした情報収集に関する研究	2
2.2	本手法で用いる技術に関する研究	3
2.2.1	抽出で用いる技術	3
2.2.2	分類で用いる技術	3
第3章	ブログからの長所情報の抽出	5
3.1	長所情報の抽出対象	5
3.2	長所情報の抽出方法	5
3.3	長所情報の例	6
3.4	長所情報の抽出結果	6
第4章	長所情報の分類	8
4.1	用言による分類	8
4.1.1	用言による分類手法	8
4.1.2	用言による分類結果	9
4.2	名詞による分類	13
4.2.1	名詞の特徴度の算出	13
4.2.2	名詞の特徴度の算出結果	13
4.3	用言および名詞による分類結果	15
第5章	開発目標地への適用可能性の分析	19
5.1	分析方法	19
5.2	分析結果	19

第6章	分類の評価	21
6.1	評価手順	21
6.2	ヒントの付与	21
6.3	ヒントの付与結果	22
6.4	比較する分類手法	22
6.5	特徴度ごとのヒント数集計結果	23
6.6	考察	29
第7章	今後の課題	30
第8章	おわりに	31

目 次

3.1	長所情報の一例	6
4.1	NY:5 (属性) の例 1	11
4.2	NY:5 (属性) の例 2	11
4.3	NY:18 (物理的移動) の例 1	11
4.4	NY:18 (物理的移動) の例 2	12
4.5	NY:23 (身体動作) の例 1	12
4.6	NY:23 (身体動作) の例 2	12
4.7	江ノ島海岸における長所情報の具体例 (一部)	15
4.8	若狭湾における長所情報の具体例 (一部)	16
4.9	三陸海岸における長所情報の具体例 (一部)	17
6.1	ヒントの付与例	22
6.2	長所情報に対する特徴度の付与例	23
6.3	分類手法 A における名詞の特徴度とヒントの出現数の関係 (江ノ島海岸)	24
6.4	分類手法 B における名詞の特徴度とヒントの出現数の関係 (江ノ島海岸)	24
6.5	分類手法 A における名詞の特徴度とヒントの出現数の関係 (若狭湾) . .	26
6.6	分類手法 B における名詞の特徴度とヒントの出現数の関係 (若狭湾) . .	26
6.7	分類手法 A における名詞の特徴度とヒントの出現数の関係 (三陸海岸) .	28
6.8	分類手法 B における名詞の特徴度とヒントの出現数の関係 (三陸海岸) .	28
7.1	ネガティブな体験談の例	30

表 目 次

3.1	取得したブログ記事および抽出した長所情報	7
4.1	日本語語彙大系の用言意味属性の番号と名称	9
4.2	各観光地の長所情報における用言意味属性の出現頻度	10
4.3	取得した特徴語数	13
4.4	各観光地の長所情報における名詞とその特徴度（上位 30 件）	14
6.1	ヒントの付与結果	22
6.2	名詞の特徴度におけるヒントの出現数（江ノ島海岸）	25
6.3	名詞の特徴度におけるヒントの出現数（若狭湾）	27
6.4	名詞の特徴度におけるヒントの出現数（三陸海岸）	29

第1章 はじめに

本研究では、観光地開発の支援技術として、観光地におけるポジティブな体験談をブログから収集、かつ、分類を行う手法を提案する。一般に、飲食店やホテルという施設に対しては、予約サイトや口コミ情報を集約したサイトがあるため、分析しやすい[1]。これに対して、鳥取砂丘や山陰海岸という非施設に対しては、予約サイトなどがなく、口コミ情報は集約されにくい。この場合、従来では、ブログから情報を収集する方法がある[2]。情報収集の際、形容詞や品詞の並びに注目した観光の評判情報を収集する手法がある[3]。ここで、観光地開発支援という立場では、評判情報を収集するだけでなく、関連する体験談を収集することも要求される。さらに、分析者の負担を軽減するため、分析対象の絞り込みも必要となる。

そこで、本研究では、開発目標地と類似する観光地におけるポジティブな体験談をブログ記事から収集し、体験談を種類別に分類する。例えば「山陰海岸」を開発目標地とした場合、開発目標地以外の海岸におけるポジティブな体験談をブログ記事から収集、および、分類する。取得した分類結果は、観光開発案を得るために人手で分析する。さらに、体験談を分類することで、人手分析の支援になっているかを確認するため、名詞の特徴度と観光開発案に結び付く情報の出現数との関係を調査する。

本論文の構成は以下の通りである。第2章において、関連研究について述べる。第3章でブログからポジティブな体験談を抽出する方法を示し、第4章で体験談を分類する方法を示す。第5章では、分類結果を人手で分析することで、開発目標地に対する観光のポイントの適用可能性を考察する。第6章において、分類の効果を確認する。第7章にて、本研究の今後の課題を述べ、第8章で研究をまとめる。

第2章 関連研究

2.1 Web データを対象とした情報収集に関する研究

Web 上の口コミ情報の重要性に着目し、評価表現と評価理由の因果関係を利用した、高野らによる評判情報の自動抽出がある [1]。まず、高野らは、Web 上におけるホテルに関する掲示板の書き込み情報から、評価理由を伴う評価表現（例：[A] なので [B]; A: 評価理由, B: 評価表現）を抽出した。次に、抽出した評価理由を種として、評価表現を取得した。このように、評価理由と評価表現を交互に取得することで、評判情報を自動的に抽出した。この手法を用いて、定性的な評価を行ったところ、評判対象ごとの固有な評判情報が取得された。

石野らは、旅行に関するブログエントリーから自動的にリンクを取得し、観光情報に関するリンク集を構築する手法を提案している [2]。まず、複数のエントリーに渡り旅行記事を書いているブログ著者がいることを考慮し、取得目標となるエントリーとその前後エントリーに着目し、機械学習 (CRF) を用いて旅行ブログエントリーを取得した。次に、取得したエントリーから、観光情報に関するリンクを自動取得した。ここで、ブログ著者がエントリーにて紹介した観光に関するリンクに対して、旅行者が行った施設や宿泊地といった、リンクの種類（文献 [2] では、リンクタイプと呼ぶ）を設定している。リンクタイプの判定には、リンク情報が述べられている部分の各単語と、リンクの手がかりとなる語を素性とした機械学習を用いている。以上の手法を評価するため、人手で付与したリンクタイプを正解とし、機械が出力したリンクタイプの精度と再現率を算出したところ、精度が7割以上、再現率が6割以上となり、文献 [2] の手法の有効性は示された。

観光評判情報の抽出において、山田の、形容詞や品詞の並びに注目した抽出がある [3]。山田は、評判情報に形容詞がよく使用されていることに注目して、形容詞を主軸とした品詞の組合せを作成した。作成した組合せを用いて、ブログ記事に対してフィルタリングを行うことで評判情報を取得した。文献 [3] の手法の評価実験として、フィルタリングによる評判情報の取得結果が評判情報であるかどうかを人手で評価したところ、約7割が評判情報であると判断された。

2.2 本手法で用いる技術に関する研究

2.2.1 抽出で用いる技術

本研究では、ポジティブな体験談の抽出に情緒推定を用いる。そこで、本節において、従来の情緒推定技術について紹介する。

田中らによる情緒生起情報を付与したパターン辞書がある [4]。田中らは、日本語語彙大系 [7] の結合価パターン 14,819 個に対して、7,320 個の情緒属性（「前提条件」、「情緒主」、「情緒対象」、「原因」、「情緒名」）を付与した。ここで、情緒名は、8 つの基本情緒（「喜び」、「悲しみ」、「好ましい」、「嫌だ」、「驚き」、「期待」、「恐れ」、「怒り」）としている。これにより、情緒生起原因による情緒推定の可能性を示した。

黒住らによって、情緒状態と直結した情緒反応を表す用言の結合価パターン辞書が作成されている [5]。黒住らの辞書は、日本語語彙大系 [7] における「感情動作」および「感情状態」カテゴリに属する結合価パターン 1,615 件の内、情緒を明示するもの 1,030 件と情緒による反応のもの 100 件に情緒属性（「情緒主」、「情緒対象」、「情緒名」）の付与を行っている。情緒名に関しては、8 つの情緒に「なし」を加えた 9 種類としている。

徳久は、言語表現から話者の情緒を機械的に推定する手法を提案した [6]。情緒推定手法については、情緒原因、情緒状態、および、情緒表出についての表現をパターン辞書で解析している。さらに、機械に対する入力文とパターンの照合の際に複数のパターンが適合する曖昧性の問題に対して、ヒューリスティクスを用いることで解消している。

2.2.2 分類で用いる技術

本研究では、用言による分類と名詞による分類を行う。そこで、本節では、分類に用いる技術について紹介する。

用言の意味解析

日本語語彙大系 [7] では、日英機械翻訳の意味解析のために、用言 6,000 語に対して、14,700 件の日英対訳パターンが付与されている。文献 [7] は、意味体系、単語体系、および、構文体系から構成される。意味体系には、2,830 個の名詞意味属性と、36 個の用言意味属性の体系図が収録されている。単語体系においては、一般名詞と固有名詞を合わせた 30 万語に対して、単語表記、読み、品詞分類、および、名詞意味属性が付与されて

いる．構文体系には，14,700 の日英文型パターン対があり，格要素の名詞制約が記されている．

単語（名詞）の特徴度

文書検索における関数に，Okapi BM25 がある．Okapi BM25 は，複数のキーワードから成る検索クエリ Q に対して，文書のスコアを出す関数である [8]．文書 D とキーワード q によるスコア $score(D, q)$ は，次の式で求める．

$$score(D, Q) = \sum_{q \in Q} s(D, q) \quad (2.1)$$

$$s(D, q) = IDF(q) \cdot \frac{f(q, D) \cdot (k + 1)}{f(q, D) + k \cdot (1 - b + b \cdot \frac{|D|}{avgdl})} \quad (2.2)$$

$$IDF(q) = \log \frac{N - n(q) + 0.5}{n(q) + 0.5} \quad (2.3)$$

ここで， $f(q, D)$ は D における q の出現頻度， $n(q)$ は q を含む文書数， $|D|$ は文書 D の長さ， $avgdl$ は全文書における平均長， N は全文書数である． k と b は，定数である．

第3章 ブログからの長所情報の抽出

本研究では，観光地におけるポジティブな体験談を長所情報と呼ぶ．長所情報は，ポジティブであると判断できる1文（以下，主文と呼ぶ）とその前後1文ずつを合わせた3文で構成する．本章では，長所情報に関して，抽出対象となるブログ，抽出方法，および，抽出結果について述べる．

3.1 長所情報の抽出対象

長所情報の抽出対象は，Yahoo!ブログにおける「トップ>カテゴリトップ>趣味とスポーツ>レジャー>旅行」カテゴリ（2010年7月16日現在）に登録されたブログ記事とする．本研究では，開発目標地を「山陰海岸」とし，類似する観光地を，同じ海岸である地域の「江ノ島海岸」，「若狭湾」，および「三陸海岸」とする．類似する観光地には，開発目標地における開発案につながる情報が多いと予想した．

そこで，上記カテゴリに登録されたブログ記事において「江ノ島海岸」，「若狭湾」，または「三陸海岸」と記述されたブログ記事を抽出対象とする．

3.2 長所情報の抽出方法

長所情報の抽出には，情緒推定を用いる．本研究で用いる情緒推定は，情緒原因，情緒状態，および，情緒表出についての表現をパターン辞書で解析する手法である [4],[5],[6]．実際には，文献 [6] の機械による情緒推定を使用する．機械による情緒推定では，喜び，好ましい，期待，悲しみ，恐れ，怒り，嫌だ，驚き，および，情緒なしの9分類が出力される．本研究では，ポジティブな体験談を抽出するため，上記の9分類を5分類に変換する．5分類を以下に示す．

《P》 喜び，好ましい，期待などの情緒

《N》 悲しみ，恐れなどの情緒

《A》 怒り，嫌だなどの情緒

《S》 驚きなどの情緒

《なし》 情緒なし

ブログ記事を文単位に分割し，それぞれ情緒推定を行うことで，ブログ著者の情緒を推定する．推定結果が《P》となった文を主文と判定する．主文とその前後文を合わせて長所情報として取得する．なお，第4章で示す長所情報の分類において，主文をキーとして使用する．主文の前後文は，第5章での長所情報の人手分析の際に，文脈の理解を補助するものとなる．

3.3 長所情報の例

取得した長所情報の例を図3.1に示す．文ID「E00126」が主文である．前後文を取得することで，ブログ著者が「しらす料理を食べた店名」を得ることができた．つまり，長所情報として3文ずつ取得することで，ポジティブな情報を多く取得することができる．

E00125	いつの間にか由比から湘南になりサクラエビから しらすに入れ替わってしまいました(汗
<u>E00126</u>	ま、どっちも海にいるものだしどっちも生で食べた り、かき揚げにしたりするところはよく似ているから 今回はしらす料理のほうをどーんと堪能する ことにしました。
E00127	事前にぴの字さんが調査して、連れて行ってくれたの は茅ヶ崎にある【快飛】(かっとび)というお店。

図 3.1: 長所情報の一例

3.4 長所情報の抽出結果

長所情報の抽出結果を表3.1に示す．なお，長所文数とは，取得した全ての長所情報の構成文から重複した文を除いた文数のことである．表3.1の通り，ブログ記事の総文数12,047文から，長所情報として見るべき文数を6,348文まで絞り込むことができた．

表 3.1: 取得したブログ記事および抽出した長所情報

観光地名	ブログ記事数 [件]	総文数 [文]	長所情報 [件]	長所文数 [文]
江ノ島海岸	21	662	199	357
若狭湾	180	5,028	1,199	2,660
三陸海岸	243	6,357	1,496	3,331
(合計)	444	12,047	2,894	6,348

第4章 長所情報の分類

長所情報は、第5章において人手で分析するので、人手分析の支援のために、事象の種類、すなわち用言の意味属性で分類する。用言の意味属性による分類後、観光地における特徴的な話題を見るために、観光地の特徴を表す名詞による分類を行う。

4.1 用言による分類

本研究では、主文における用言の意味属性を用いて、長所情報を分類する。用言の意味属性による分類を行うことで、観光地の長所情報を動作や状態の種類別に大別できると考えられる。本節では、用言による分類方法と結果について述べる。

4.1.1 用言による分類手法

用言による分類において、日本語語彙大系 [7] における用言意味属性 36 分類 (表 4.1) を用いる。

本研究では、主文から用言意味属性を取得して、長所情報を分類する。1つの用言に複数の用言意味属性が対応しているため、用言パターンの照合によって、曖昧性を減らす。主文に複数の用言がある場合は、複数の分類先を出力することを認める。なお、用言が全く無い文は、NY:0 と表記する。

例えば、「ようやく長い上りの階段が終わって着いたところは江ノ島神社。」という主文があった場合、「終わる」の「NY:36 (終了)」, および「着く」の「NY:18 (物理的移動)」の用言意味属性にそれぞれ分類される。以上のように、全ての主文を、用言意味属性 NY:1~36, および、用言が全く無い NY:0 へ分類する。

表 4.1: 日本語語彙大系の用言意味属性の番号と名称

番号	属性名	番号	属性名	番号	属性名
NY:1	事象	NY:13	心的状態	NY:25	結合動作
NY:2	状態	NY:14	身体状態	NY:26	生成
NY:3	抽象的關係	NY:15	自然現象	NY:27	消滅・破壊
NY:4	存在	NY:16	行動	NY:28	精神的行動
NY:5	属性	NY:17	物理的行動	NY:29	精神的移動
NY:6	所有	NY:18	物理的移動	NY:30	知覚動作
NY:7	相對關係	NY:19	所有的移動	NY:31	感情動作
NY:8	因果關係	NY:20	属性变化	NY:32	思考動作
NY:9	精神的關係	NY:21	身体变化	NY:33	使役
NY:10	知覚状態	NY:22	結果	NY:34	可能
NY:11	感情状態	NY:23	身体動作	NY:35	開始
NY:12	思考状態	NY:24	利用	NY:36	終了

4.1.2 用言による分類結果

「江ノ島海岸」, 「若狭湾」, および「三陸海岸」のブログ記事における長所情報の分類結果を表 4.2 に示す。表 4.2 では, 主要な体験談から長所情報を分析するために, 分類数における降順で長所情報をソートする。なお, 分類数が 0 のものは省略している。

表 4.2 より, 次のことがいえる。

表 4.2: 各観光地の長所情報における用言意味属性の出現頻度

(a) 江ノ島海岸		(b) 若狭湾		(c) 三陸海岸	
NY	分類数	NY	分類数	NY	分類数
NY: 5	50	NY: 5	463	NY: 5	548
NY:23	34	NY:18	305	NY:18	414
NY:18	32	NY:23	252	NY:23	312
NY:29	29	NY:29	205	NY:32	259
NY: 0	23	NY:32	190	NY:20	255
NY:31	21	NY:20	185	NY:29	254
NY:20	18	NY:31	125	NY:31	142
NY:32	17	NY:19	87	NY: 0	91
NY:19	12	NY: 0	77	NY:19	83
NY: 7	4	NY:30	65	NY:30	80
NY: 2	4	NY:26	37	NY:22	48
NY:16	4	NY: 2	35	NY: 7	47
NY:22	3	NY:11	33	NY:25	45
NY:11	3	NY: 7	32	NY:16	36
NY:35	2	NY:25	32	NY:26	35
NY:15	2	NY:22	29	NY: 2	30
NY:36	1	NY:35	23	NY:11	24
NY:30	1	NY:16	20	NY:21	19
NY:26	1	NY:21	11	NY:35	15
NY:25	1	NY: 6	9	NY: 6	12
NY:21	1	NY:15	6	NY:15	6
NY:12	1	NY:36	5	NY: 8	5
NY:10	1	NY: 8	1	NY:36	4
		NY:27	1	NY:10	2
		NY:12	1	NY:27	1
		NY:10	1		

「江ノ島海岸」、「若狭湾」、および「三陸海岸」の出現頻度を見ると、上位3件が同じ分類となっている。まず、NY:5（属性）の実際の例を図4.1、4.2に示す。

-
- E00480 すばな通りにて。
E00481 江ノ島駅から江ノ島までの約500メートルが「すばな通り」と呼ばれる商店街です。
E00482 歴史を感じる旅館や商店が多く立ち並んでいます。
-

図 4.1: NY:5（属性）の例 1

-
- W00410 敦賀といえばキレイな砂浜の風景を雑誌で見たことがあり、行ってみる事にしました。
W00411 敦賀半島にある「水晶浜」は雑誌で見たとおり白い砂と透き通った海の色で私達を迎えてくれました。
W00412 ウチの子は海に来たのは初めてで、砂浜の感触や打ち寄せる波に興味深々の様子です。
-

図 4.2: NY:5（属性）の例 2

NY:5（属性）には、形容詞の用言が分類されるため、観光地の事象に対するブログ著者の印象が書かれている文が分類されると予想できる。図4.1では江ノ島の商店街にある“歴史を感じる”建物について、図4.2では敦賀半島にある“綺麗な”水晶浜について書かれている。以上より、NY:5には、観光地の事象（建造物や自然物など）に対するポジティブな印象を示す文が分類されている可能性がある。

次に、NY:18（物理的移動）の例を図4.3、4.4に示す。

-
- E00182 あなた方は私のことには構わずに先に行って…」と迫真の演技をしながら進むのが鉄則。
E00183 ようやく長い上りの階段が終わって着いたところは江ノ島神社。
E00184 ここで無事の到着のお礼を報告。
-

図 4.3: NY:18（物理的移動）の例 1

S05211	同じ事をしている方が2, 3名、やっぱり鉄道オタクなのでしょう か。
S05212	ここから再びバスに乗り込み北山崎に向かいます。
S05213	北山崎は『海のアλπス』とも言われる絶景ポイント。

図 4.4: NY:18 (物理的移動) の例 2

NY:18 (物理的移動) には、観光地における移動先や移動方法が書かれている文が分類されることを期待できる。実際に例を見ると、図 4.3 では、移動先が江ノ島神社であること、図 4.4 では、北山崎までの道中のこと書かれている。以上より、NY:18 には移動に関する文が分類されている可能性がある。

最後に、NY:23 (身体動作) の例を図 4.5, 4.6 に示す。

E00127	事前にぴの字さんが調査して、連れて行ってくれたのは茅ヶ崎にある【快飛】(かっとび) というお店。
E00128	水産会社直営とあって、どこより新鮮なしらすを食べさせてくれるところです。
E00129	お昼時なので、お店に到着したときにはすでに順番待ちの行列。

図 4.5: NY:23 (身体動作) の例 1

S00282	岩手県花巻で立ち寄った食事処では、「鹿踊り(ししおどり)」の祭り衣装でのお迎えです。
S00283	仙台空港で、いつものように仙台名物「牛タン」を食べて、大阪へ帰りました。
S00284	東北の日本海側から津軽海峡側、太平洋側と海岸線を、天候に恵まれ巡った「東北の旅」もこれでおしまいです。

図 4.6: NY:23 (身体動作) の例 2

NY:23 (身体動作) には、観光地における行動(例えば、「食べる」)について書かれている文が分類されることを期待する。実際に例を見ると、図 4.5 では、しらすを食べられる店について、図 4.6 では、牛タンを食べたことについて書かれている。以上より、NY:23 には観光地においての行動「食べる」ことに関する文が分類されていることが確認できる。

以上より、用言による分類によって、事象の種類が類似した文ごとに、ある程度集めることができたと考える。

4.2 名詞による分類

ブログ記事において、観光地の特徴的な名詞に着目することで、観光地の主要な場面を取得できると仮定した。そこで、名詞の特徴度を測るために、Okapi BM25 を用いる。本節では、名詞の特徴度の算出方法と結果について述べる。

4.2.1 名詞の特徴度の算出

第 2.2.2 節で述べた式 (2.2) に注目する。式 (2.2) の値は、キーワード q が複数の文書で共通して現れる場合、値は低く、1つの文書にだけ現れる場合、値は高くなる。そこで、本研究では、名詞の特徴度を式 (2.2) を用いて算出する。具体的には、地域 r の長所情報の集合を 1つの文書 D_r とみなし、キーワード q について、各観光地の名詞の特徴度を $s(D_r, q)$ で求めることにする。なお、定数は $k = 2.0, b = 0.75$ とした。

4.2.2 名詞の特徴度の算出結果

各観光地の長所情報の主文から名詞を抽出し、特徴度を求めた。各観光地ごとの取得できた特徴語数を表 4.3 に示す。

観光地名	特徴語数 [個]
江ノ島海岸	447
若狭湾	2374
三陸海岸	2601
(合計)	5422

具体的な特徴語と特徴度を表 4.4 に示す。「江ノ島海岸」の長所情報には「江ノ島」や「しらす」、若狭湾については「敦賀」や「鯖」という名詞が上位に位置しており、各観光地ごとの特徴的な名詞が上位に現れていることが確認できた。

表 4.4: 各観光地の長所情報における名詞とその特徴度（上位 30 件）

(a) 江ノ島海岸		(b) 若狭湾		(c) 三陸海岸	
語	特徴度	語	特徴度	語	特徴度
江ノ島	1.469	敦賀	1.415	東北	1.416
坊主	1.296	三方	1.411	仙台	1.340
ロマンスカー	1.233	鯖	1.387	盛岡	1.321
しらす	1.233	レインボー	1.276	宮古	1.311
島牧	1.233	福井	1.276	北山崎	1.286
湘南	1.233	松原	1.257	釜石	1.271
相棒	1.233	越前	1.257	気仙沼	1.255
片瀬	1.233	家族	1.209	義経	1.236
江	1.233	舞鶴	1.178	岩手	1.236
k a g o	1.123	名古屋	1.140	浄土	1.215
-----	1.123	山頂	1.140	榎本	1.190
練り香	1.123	久々子湖	1.094	巢	1.190
極楽寺	1.123	日向湖	1.094	久慈	1.162
寿都	1.123	東尋坊	1.094	ウミネコ	1.128
喧嘩	1.123	水月湖	1.094	アルプス	1.128
百人	1.123	リフト	1.094	日本一	1.128
かま	1.123	富山	1.094	石巻	1.128
平塚	1.123	菅湖	1.094	本州	1.128
賀老	1.123	サバ	1.094	青森	1.128
はぁ	1.123	27	1.094	地底	1.128
針	1.123	直接	1.094	陸中海岸	1.087
ウォード・	0.887	京	1.094	ブログ	1.087
ファーマーズマーケット		ウインズ	1.034	鍾乳洞	1.087
ウォード・	0.887	1934	1.034	水族館	1.087
ウェア・ハウス		汽水湖	1.034	南部	1.087
バーベキュー	0.887	北陸道	1.034	ご飯	1.087
サラリーマン	0.887	福知山	1.034	八戸	1.087
ウォーキング	0.887	てます	1.034	タクシー	1.036
片瀬江ノ島	0.887	五木	1.034	秋田	1.036
サクラエビ	0.887	早朝	1.034	市場	1.036
されてみる	0.887				
めっちゃ	0.887				

4.3 用言および名詞による分類結果

各観光地の長所情報を、用言意味属性で大別した後、各属性の中において名詞の特徴度で並び替えた。具体例の一部を図4.7, 4.8, 4.9に示す。

分類1：NY:5 (属性)；江ノ島 [1.469]		
長所情報1	E00193	日常では見られない風景を見たとき心の中に旅の感動がどかーん !! と押し寄せてきます。
	E00194	ああ、とっても楽しい江ノ島体験でした。
	E00195	ぴの字さん&ろの字さん、楽しい時間をありがとう
長所情報2	E00223	その後展望台をあとにして、島のさらに奥の方辺りを軽〜く探索。。
	E00224	ひっそりとした風情ある小路やなんかをあちこち見てまわったあと、歩いてきた順路を戻り楽しかった江ノ島散策を終えることにする。。
	E00225	以下、其の参へ続く...
分類2：NY:23 (身体動作)；しらす [1.233]		
長所情報3	E00127	事前にぴの字さんが調査して、連れて行ってくれたのは茅ヶ崎にある【快飛】(かっとび)というお店。
	E00128	水産会社直営とあって、どこより新鮮なしらすを食べさせてくれるところです。
	E00129	お昼時なので、お店に到着したときにはすでに順番待ちの行列。
分類3：NY:18 (物理的移動)；江 [1.233]		
長所情報4	E00563	そして鎌倉の街を歩きまわって、鎌倉コロッケ食べて、まだちょっと時間あるな。。
	E00564	江ノ電に乗って海見に行こ !!
	E00565	由比ヶ浜駅下車後、浜へ !!!

図 4.7: 江ノ島海岸における長所情報の具体例 (一部)

図4.7における、長所情報1,2は、NY:5(属性)に属し、名詞「江ノ島(特徴度1.469)」を文中に含む分類1、長所情報3は、NY:23(身体動作)に属し、名詞「しらす(特徴度1.233)」を含む分類2、長所情報4は、NY:18(物理的移動)に属し、名詞「江(特徴度1.233)」を含む分類3である。分類1を見ると、どちらの例も江ノ島の風景について述べられている。一方、分類2を見ると、しらすを食べられる店について述べている。分類3では、江ノ電に乗って移動することが記述されている。

分類 1 : NY:5 (属性); 久々子湖 [1.094]		
長所情報 1	W01394	対して他の湖は、地層の変化によってできた天然湖である。
	W01395	即ち、五つの湖は海水・汽水・淡水とそれぞれに違った水質や水深を持ち、また同じ汽水湖でも日本海に直接つながっている久々子湖と奥にある菅湖や三方湖、中間の水月湖ではそれぞれ海水と淡水の割合が違っている。
	W01396	そのため梅丈岳（三方五湖レインボーライン展望台）から見える景色は、五つの湖がそれぞれに違った青色をして見えるという。
長所情報 2	W04717	湖の周囲には梅畑が広がる景勝地。
	W04718	五つの湖は淡水・海水・汽水とそれぞれに違った性質を持ち、また同じ汽水湖でも日本海に直接つながっている久々子湖と一番奥にある菅湖、中間の水月湖ではそれぞれ海水と淡水の比率が違っている。
	W04719	そのため梅丈岳（三方五湖レインボーライン展望台）から見える景色は、五つの湖がそれぞれに違った青色をして見えるのである。
分類 2 : NY:18 (物理的移動); 敦賀 [1.415]		
長所情報 3	W00400	またこの辺りの山の斜面には冬になると水仙が咲き乱れ、美しい風景となるようです。
	W00401	越前を後にしてもと来た道を戻り、敦賀を目指します。
	W00402	この道中の国道 305 号は日本海を眺めながら走れる風光明媚なドライブコースで、GWということもあってバイクツーリングの集団と何度もすれ違いました。
分類 3 : NY:23 (身体動作); 敦賀 [1.415]		
長所情報 4	W00470	若狭三方五湖周辺の梅の里会館で「梅干し」を買い美浜へしこ館で「鯖へしこ」と「イワシのへしこ」を買いお土産を買い終えて敦賀に向かいました
	W00471	敦賀に着く頃には、お昼となり昼食を何処で食べようかなと
	W00472	向かった先は、敦賀麺の名店【中華そば一力】

図 4.8: 若狭湾における長所情報の具体例（一部）

図 4.8 における、長所情報 1,2 は、NY:5 (属性) に属し、名詞「久々子湖 (特徴度 1.094)」を文中に含む分類 1、長所情報 3 は、NY:18 (物理的移動) に属し、名詞「敦賀 (特徴度 1.415)」を含む分類 2、長所情報 4 は、NY:23 (身体動作) に属し、名詞「敦賀 (特徴度 1.415)」を含む分類 3 である。分類 1 を見ると、共通して、5 つの湖の色の違いについて述べられている。一方、分類 2 を見ると、敦賀への道中について述べている。分類 3 では、敦賀での昼食のことが記述されている。

分類 1 : NY:5 (属性) ; 北山崎 [1.286]		
長所情報 1	S02204	2 日目の朝ホテルの窓を開けると、昨夜までしとしと降っていた雨は嘘のように朝日が輝いていた。
	S02205	バスは北山崎展望台へ、ここは海のアルプスと称される険しい断崖が見られました。
	S02206	展望台から断崖の下の方へ下りる階段もあり、是非下りて見たかったのだが時間と、数百段の階段には自信が無くて諦めた。
長所情報 2	S02364	(^^ ♪
	S02365	北山崎は、海のアルプスとも称される三陸海岸北部を代表する観光スポット。
	S02366	200メートルの切り立った断崖が、8キロに渡り連なっています。
分類 2 : NY:18 (物理的移動) ; 東北 [1.416]		
長所情報 3	S00337	(2008年) 犬っこ・アメッコ・冬の東北の旅
	S00338	雪見列車に行こうと思立ち冬の東北へ。
	S00339	東北の冬の祭り、湯沢の犬っこ祭り、大館のアメッコ祭り
分類 3 : NY:23 (身体動作) ; 宮古 [1.311]		
長所情報 4	S05106	今日のルートは盛岡から国道106号を經由し、宮古から三陸海岸沿いを走りました。
	S05107	いつもコメントいただいている「黒りんご」サンお薦めの宮古駅前にある「魚彩亭すみよし」で「三陸丼」を昼食に取りました。
	S05108	さすがに産地モノなので、活きがよく旨かったですね～

図 4.9: 三陸海岸における長所情報の具体例 (一部)

図 4.9 における、長所情報 1,2 は、NY:5 (属性) に属し、名詞「北山崎 (特徴度 1.286)」を文中に含む分類 1、長所情報 3 は、NY:18 (物理的移動) に属し、名詞「東北 (特徴度 1.416)」を含む分類 2、長所情報 4 は、NY:23 (身体動作) に属し、名詞「宮古 (特徴度 1.311)」を含む分類 3 である。分類 1 を見ると、どちらの例も共通して、北山崎では「海のアルプス」といわれる海岸を見ることができると記されている。一方、分類 2 を見ると、東北の雪見列車について述べている。分類 3 では、産地物を使った「三陸丼」を食べたと記されている。

以上の結果より，同じ分類内では，長所情報が共通の話題であるため，少数の文をななめ読みするだけで観光地の情報を把握できる．一方，分類先が異なると長所情報の種類が違うことが確認できる．ゆえに，本分類によって，長所情報を種類別に分析でき，わずかな文を読むだけで，ある程度の観光地情報を把握できると考えられる．

第5章 開発目標地への適用可能性の分析

観光地開発案を取得するため、分類結果を人手で分析する。開発案を取得した後、開発目標地に対する適用可能性を考察する。本章では、開発目標地を「山陰海岸」とし、類似する観光地を「江ノ島海岸」、「若狭湾」、および「三陸海岸」として分析する。

5.1 分析方法

各観光地ごとに、第4章で述べた分類結果を見て、人手分析を行う。分析方法は以下の方法が考えられる。

1. 各観光地において分類別に分析
2. 用言意味属性に注目して各観光地ごとに比較分析

上記1は、分類別に分析することで、旅行者の行動の種類、および、事象ごとに分析できる。例えば、「江ノ島海岸」のブログ記事において、NY:18（物理的移動）に大別され、特徴語「江ノ島 [特徴度 1.469]」を含む分類先を見ると、江ノ島への移動方法や移動中の出来事について分析できると考えられる。これにより、移動方法についての観光のポイントが得られると考える。

上記2は、旅行者の行動の種類ごとに分析し、観光地ごとに比較することで、類似する観光地において主要なことがそうでないかを確認することができると考えられる。

5.2 分析結果

上記1による分析結果

「江ノ島海岸」におけるNY:18（物理的移動）に注目して分析する。E00013には「ロマンスカー」が、E00564には「江ノ電」が記述されている。これより、「江ノ島海岸」においては、ご当地の乗物を旅行者が利用していることが読み取れる。

そこで、「山陰海岸」においても、ご当地の乗物を用意することで、旅行者が注目する移動手段を作ることができると思う。また、既にご当地の乗物がある場合は、効果的な宣伝方法をとることで、旅行者に利用してもらえることが考えられる。

実際に、「山陰海岸」の鳥取砂丘エリアについて、観光情報誌まっふるマガジン鳥取大山・三朝温泉（2009年5月発行）で、ご当地の移動手段を調査した。鳥取砂丘のページでは、鳥取砂丘の概要や土産についての紹介しかなく、ご当地の移動手段については触れられていなかった。これらより、新しい開発案が一つできたと考えられる。

上記2による分析結果

NY:23（身体動作）に注目して分析する。ブログ著者が食べたと考えられる、図4.8における「若狭湾」の「敦賀麺」（W00472）と、図4.9における「三陸海岸」の「三陸丼」（S05107）に注目する。ともに、料理名に土地名を付けた商品であることがわかる。S05108を見ると、産地物なのでよかったと書かれている。以上より、観光地特有の産地物を利用し、一目で産地物を利用したものであるとわかる商品に、旅行者は惹かれる可能性がある。

そこで、「山陰海岸」において、産地物を利用した「山陰丼」や「山陰麺」を作ることによって、旅行者が注目する商品を作ることができると思う。

実際に「山陰丼」でWeb検索を行ったところ、個人的に山陰の魚介類を利用して、自宅で丼をつくっているブログ著者はいたが、鳥取県の観光地における飲食店で「山陰丼」という名称で売りだしている店は無かった。また、鳥取市観光協会のHP(<http://www.torican.jp/>)においても、「山陰丼」に関する記述は存在しなかった。これらより、開発目標地に対する一つの開発案ができたと考えられる。

第6章 分類の評価

人手分析のために開発目標地と類似する観光地に関するブログ文を分類した結果に、観光地開発案に結び付くと考えられる情報（以下、ヒントと呼ぶ）が多く含まれることが望ましい。そこで、名詞の特徴度と、ヒントの出現数との関係を調査する。

6.1 評価手順

まず、取得した類似観光地に関する全てのブログ文 12,047 文（表 3.1）に対して、それぞれヒントの有り・無しを人手で付与する。次に、本手法における名詞の特徴度（第 4.2 節）の効果を評価するため、(分類手法 A) 長所情報のみを用いた特徴度および(分類手法 B) ブログ全文を用いた特徴度それぞれと、ヒントの出現数との関係を比較する。

ヒント、および、分類手法の詳細について、それぞれ第 6.2, 6.3 節でそれぞれ述べる。

6.2 ヒントの付与

取得した類似観光地に関する全てのブログ文 12,047 (表 3.1) 文に対して、それぞれヒントの有り・無しを人手で付与する。例えば、「さらにゲットしたウニをご飯にかけウニ丼も頂きました。」という文であった場合、ブログ著者は食事をした事実しか書いていないためヒントにはならない。これは、開発目標地において「ウニ」が取得できない場合、開発目標地へ適用できる情報にならないためである。

一方、「ほんのりと苦味があり、酢としょうがをあわせて頂いたら、さっぱりと美味しかったです。」という文の場合、美味しかった食べ方とともに食事の感想を書いている。この場合、『食事を紹介する際には、ただ紹介するだけでなく、お薦めの食事方法とともに紹介する』という観光地開発案につながる。よって、ヒントであると判断する。

付与の際、名詞の特徴度とヒントの関係性を隠すため、付与者には特徴度の情報を与えず、文 ID と文のみを見せる。本研究では、筆者がヒントの付与を行う。ヒントの付与例を図 6.1 に示す。

ID	文	ヒント
...
E00062	弁財天もお参り。	無し
E00063	みんな幸せでありますように。	無し
...
E00069	ほんのりと苦味があり、酢としょうがをあわせて頂いたら、さっぱりと美味しかったです。	有り
...
E00128	水産会社直営とあって、どこより新鮮なしらすを食べさせてくれるところです。	有り
E00129	お昼時なので、お店に到着したときにはすでに順番待ちの行列。	無し
E00130	でもそれほど長く待たずに席に着くことが出来ました。	無し
...

図 6.1: ヒントの付与例

6.3 ヒントの付与結果

ヒントの付与結果を表 6.1 に示す。

表 6.1: ヒントの付与結果

観光地名	総文数 [文]	ヒント無し [文]	ヒント有り [文]
江ノ島海岸	662	417	245
若狭湾	5,028	4,203	825
三陸海岸	6,357	5,900	457
(合計)	12,047	10,520	1,527

6.4 比較する分類手法

本手法における名詞の特徴度を評価するため、以下の2つの手順の結果を比較する。分類手法 A は長所情報のみを利用した名詞の特徴度を、分類手法 B はブログ全文を利用した名詞の特徴度を用いている。

分類手法 A

1. 長所情報に特徴度を付与

2. 特徴度ごとにヒント数を集計
3. 集計結果をヒストグラム化

分類手法 B

1. 類似する観光地 3 つのブログ全文をそれぞれ文集合 D_1, D_2, D_3 として第 4.2.1 節の手法で特徴語と特徴度を取得
2. 各観光地のブログ全文を 3 文ごとに分割
3. 3 文単位に分割した情報に 1. で取得した特徴度を付与
4. 特徴度ごとにヒント数を集計
5. 集計結果をヒストグラム化

長所情報に特徴度を付与する作業について説明する．長所情報を構成する 3 文に含まれる特徴語の内，最も高い特徴度の値をその情報の特徴度とする．図 6.2 に示したの例の場合，長所情報は複数の特徴語を含むが，最も特徴度の高い特徴語「しらす」の特徴度「1.233」を長所情報 E00127～E00129 の特徴度とする．分類手法 B における 3. の作業も同様に行う．

E00127	事前にぴの字さんが調査して、連れて行ってくれたのは茅ヶ崎にある【快飛】(かっとび)というお店。
E00128	水産会社直営とあって、どこより新鮮な <u>しらす</u> を食べさせてくれるところですよ。
E00129	お昼時なので、お店に到着したときにはすでに順番待ちの行列。

図 6.2: 長所情報に対する特徴度の付与例

特徴度付与後，特徴度ごとにヒントの出現数を集計し，分類手法 A と分類手法 B を比較する．

6.5 特徴度ごとのヒント数集計結果

まず「江ノ島海岸」のブログ記事における集計結果のヒストグラムをそれぞれ図 6.3, 6.4 に示す．なお，長所情報を構成する 3 文の内，1 文でもヒント有りが付与されている場合，その長所情報はヒント有りとして集計している．

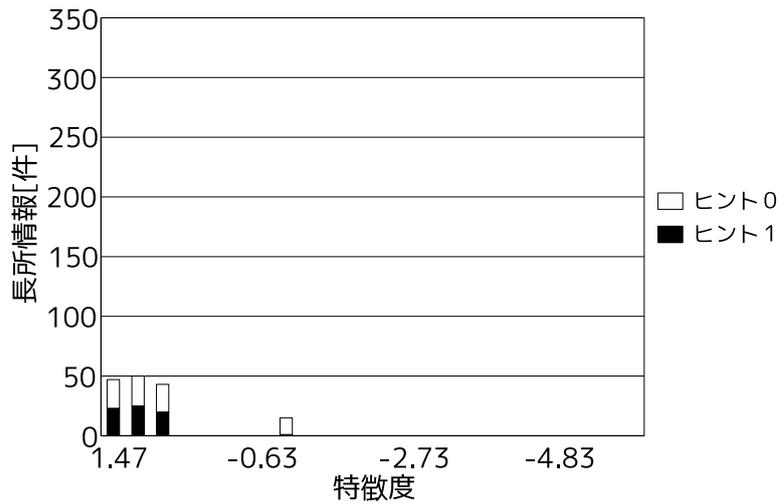


図 6.3: 分類手法 A における名詞の特徴度とヒントの出現数の関係 (江ノ島海岸)

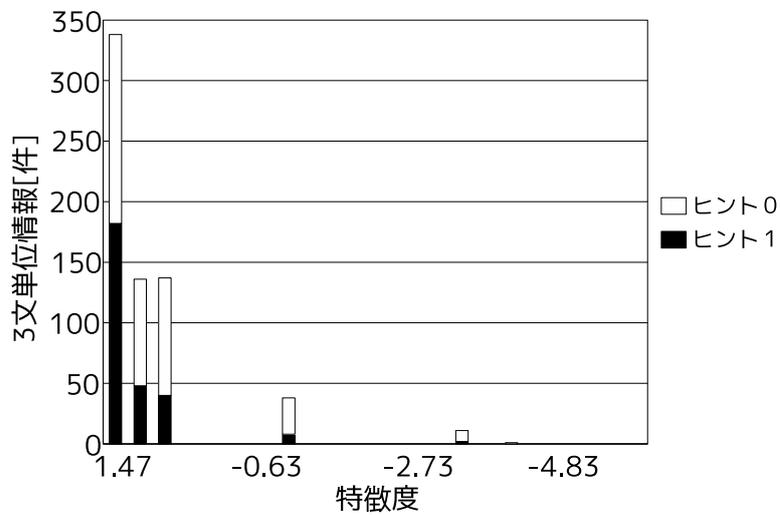


図 6.4: 分類手法 B における名詞の特徴度とヒントの出現数の関係 (江ノ島海岸)

図 6.3,6.4 では、以下の条件でヒストグラムを作成した。

- 特徴度の範囲：1.645 ~ -6.055
- 特徴度区間幅：0.35
- 分割数：21

図 6.3,6.4 を見ると、ともに特徴度が高い区間に集中してヒントが出現していることが

わかる。また、特徴度が負の範囲において、図 6.4 の方が、ヒント有りの出現数が多いことがわかる。ここで、各特徴度区間における具体的なヒントの出現数を表 6.2 に示す。ヒント有の列における括弧内の数値は、同じ特徴度区間内におけるヒント有の割合を表す。

表 6.2: 名詞の特徴度におけるヒントの出現数 (江ノ島海岸)

特徴度区間	(a) 分類手法 A		(b) 分類手法 B	
	ヒント有 [個](%)	ヒント無 [個]	ヒント有 [個](%)	ヒント無 [個]
1.645 ~ 1.295	23(49%)	24	182(54%)	156
1.295 ~ 0.945	25(50%)	25	48(35%)	88
0.945 ~ 0.595	20(47%)	23	40(29%)	97
0.595 ~ 0.245	0(0%)	0	0(0%)	0
...
-0.455 ~ -0.805	0(0%)	0	0(0%)	0
-0.805 ~ -1.155	1(7%)	14	8(21%)	30
-1.155 ~ -1.505	0(0%)	0	0(0%)	0
...
-2.905 ~ -3.255	0(0%)	0	0(0%)	0
-3.255 ~ -3.605	0(0%)	0	2(18%)	9
-3.605 ~ -3.955	0(0%)	0	0(0%)	0
-3.955 ~ -4.305	0(0%)	0	0(0%)	1
-4.305 ~ -4.655	0(0%)	0	0(0%)	0
...
-5.705 ~ -6.055	0(0%)	0	0(0%)	0

特徴度 1.645 ~ 0.595 の区間に集中してヒントが出現していることがわかる。また、分類手法 A では特徴度 -0.805 ~ -1.155 にヒント有が 1 個だけ出現していることに対して、分類手法 B では特徴度 -0.805 ~ -1.155 に 8 個、特徴度 -3.255 ~ -3.605 に 2 個出現している。

これにより、本手法における名詞の特徴度を用いた分類手法 A の方が、特徴度が高い範囲に集中してヒントが出現していることがわかる。

次に、「若狭湾」のブログ記事における集計結果のヒストグラムをそれぞれ図 6.5, 6.6 に示す。

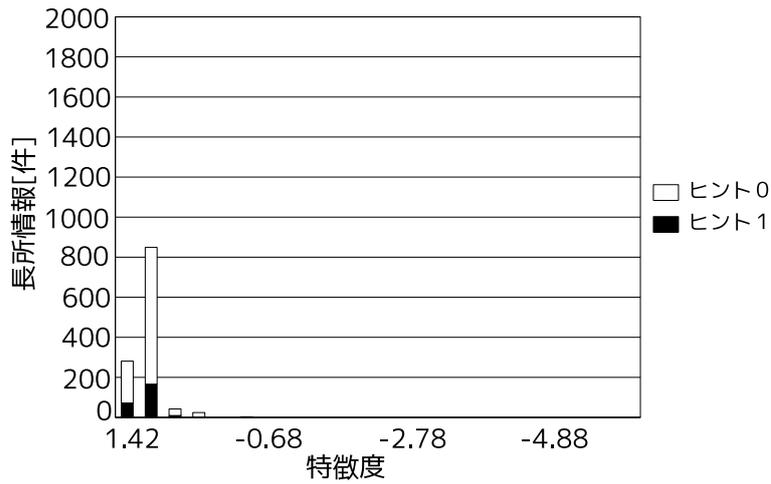


図 6.5: 分類手法 A における名詞の特徴度とヒントの出現数の関係 (若狭湾)

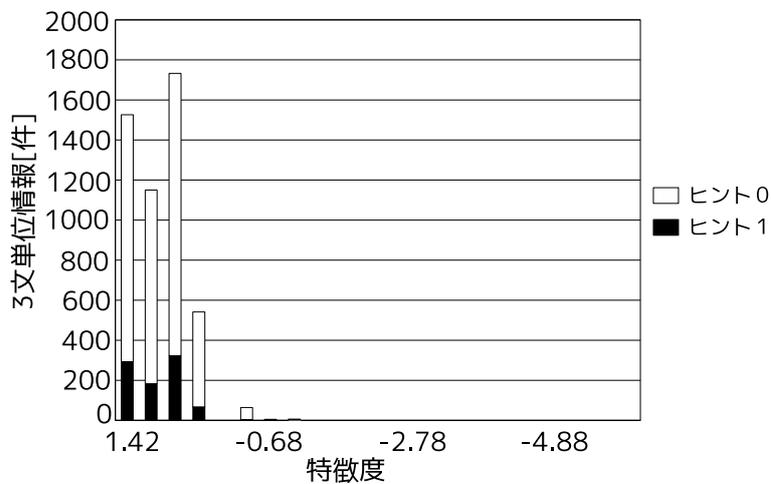


図 6.6: 分類手法 B における名詞の特徴度とヒントの出現数の関係 (若狭湾)

図 6.5,6.6 では, 以下の条件でヒストグラムを作成した.

- 特徴度の範囲 : 1.590 ~ -6.110
- 特徴度区間幅 : 0.35
- 分割数 : 21

図 6.5,6.6 を見ると, 特徴度が高い範囲にヒントの出現数が多いことがわかる. ここで, 各特徴度区間における具体的なヒントの出現数を表 6.3 に示す.

表 6.3: 名詞の特徴度におけるヒントの出現数（若狭湾）

特徴度区間	(a) 分類手法 A		(b) 分類手法 B	
	ヒント有 [個](%)	ヒント無 [個]	ヒント有 [個](%)	ヒント無 [個]
1.590 ~ 1.240	72(26%)	209	293(19%)	1233
1.240 ~ 0.890	166(20%)	683	183(16%)	967
0.890 ~ 0.540	9(21%)	33	323(19%)	1409
0.540 ~ 0.190	2(8%)	22	67(12%)	474
0.190 ~ -0.160	0(0%)	0	0(0%)	0
-0.160 ~ -0.510	0(0%)	3	4(6%)	60
-0.510 ~ -0.860	0(0%)	0	1(20%)	4
-0.860 ~ -1.210	0(0%)	0	1(17%)	5
-1.210 ~ -1.560	0(0%)	0	0(0%)	0
...
-5.760 ~ -6.110	0(0%)	0	0(0%)	0

表 6.3 の分類手法 A において、特徴度 1.240 ~ 0.890 の区間に、突出してヒント数が多い。これは、特徴度 1.590 ~ 1.240 における名詞の数が“7”に対して、特徴度 1.240 ~ 0.890 における名詞の数が“54”と、特徴度区間内の名詞数が多いため、ブログ文数が多くなるからだと考えられる。

また、分類手法 A では特徴度が負の場合はヒント有りが出現していないことに対して、分類手法 B では特徴度が負の場合にもヒント有りが出現している。これにより、江ノ島海岸の場合と同様に、分類手法 A の方が、特徴度が高い範囲に集中してヒントが出現していることがわかる。

次に「三陸海岸」のブログ記事における集計結果のヒストグラムをそれぞれ図 6.7, 6.8 に示す。

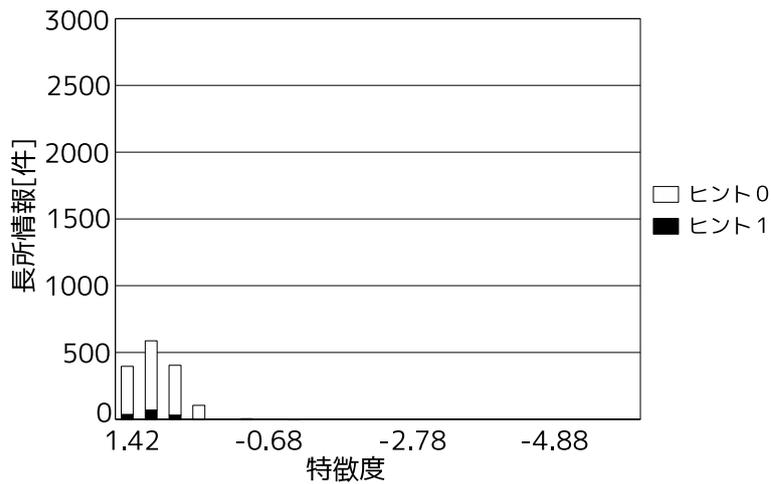


図 6.7: 分類手法 A における名詞の特徴度とヒントの出現数の関係 (三陸海岸)

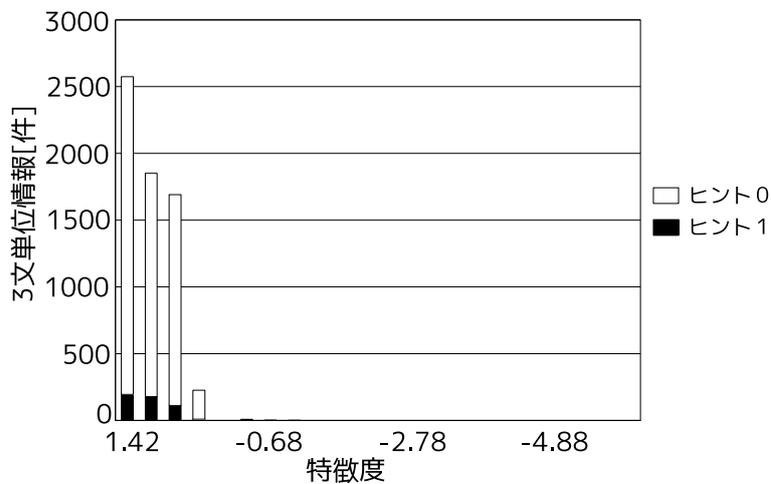


図 6.8: 分類手法 B における名詞の特徴度とヒントの出現数の関係 (三陸海岸)

図 6.7,6.8 では, 以下の条件でヒストグラムを作成した .

- 特徴度の範囲 : 1.591 ~ -6.109
- 特徴度区間幅 : 0.35
- 分割数 : 21

図 6.7,6.8 を見ると, 特徴度が高いところにヒントの出現数が集中していることがわかる . ここで, 各特徴度区間における具体的なヒントの出現数を表 6.4 に示す .

表 6.4: 名詞の特徴度におけるヒントの出現数（三陸海岸）

特徴度区間	(a) 分類手法 A		(b) 分類手法 B	
	ヒント有 [個](%)	ヒント無 [個]	ヒント有 [個](%)	ヒント無 [個]
1.591 ~ 1.241	36(9%)	360	195(8%)	2379
1.241 ~ 0.891	69(12%)	518	178(10%)	1673
0.891 ~ 0.541	31(8%)	373	112(7%)	1579
0.541 ~ 0.191	3(3%)	101	9(4%)	217
0.191 ~ -0.159	0(0%)	0	0(0%)	0
-0.159 ~ -0.509	0(0%)	4	2(25%)	6
-0.509 ~ -0.859	0(0%)	0	0(0%)	2
-0.859 ~ -1.209	0(0%)	1	0(0%)	1
-1.209 ~ -1.559	0(0%)	0	0(0%)	0
...
-5.759 ~ -6.109	0(0%)	0	0(0%)	0

若狭湾の場合と同様に，分類手法 A では特徴度が負の場合ヒント有りは出現していないが，分類手法 B では特徴度が負の場合にヒント有りがわずかながら出現している。

6.6 考察

表 6.2, 6.3, 6.4 の通り，特徴度が下がるほど，ヒント有りの出現数も減っている。ここで，表 6.2 の下線部に注目すると，提案手法を使用した手順 1 の方が，同じ特徴度区間におけるヒント有り出現数の割合が高い。表 6.3, 6.4 においても同様のことがいえる。これらより，特徴度が高い区間において，様々な場面に対する分析を実行できるといえる。

したがって，提案手法を用いることで「江ノ島海岸」における特徴度区間 1.295 ~ 0.595，「若狭湾」における特徴度区間 1.590 ~ 0.540，および「三陸海岸」における特徴度区間 1.591 ~ 0.541 でも，観光開発案につながる様々な情報を高い確率で読むことができるといえる。

以上より，提案手法を用いることで，ヒントを容易に発見することができるため，長所情報に対する人手分析の支援となっていると考えられる。

第7章 今後の課題

本研究では、観光地におけるポジティブな体験談に注目して分析を行った。しかし、ネガティブな体験談については調査していない。観光地開発案は、ポジティブな体験談のみではなく、ネガティブな体験談からも生まれると考えられる。実際の例を、図7.1に示す。なお、情緒推定が《N》となった文は、E00212である。

-
- E00212 意気揚々と登り始めたのはいいけれど、途中からはなんだか結構なつらさを感じてしまう始末ではあったかや...
- E00213 でも時々歩を止めて後ろを振り返ると、朝日に照らされた江ノ島海岸の美しい景色が徐々にスケールを増して眼下に広がってくる。。
-

図 7.1: ネガティブな体験談の例

図7.1のブログ著者は、E00212において、目的地まで徒歩で行き、辛かったというネガティブな感想を述べている。しかし、E00213において、辛い移動途中でも、美しい景色が広がっているというポジティブな感想を述べている。ここから『移動は苦勞するが、ここで楽しめる景色があることを紹介する』という観光地開発案につながる。よって、ネガティブな体験談から、新しい観光地開発案が生まれる可能性がある。

以上より、今後の課題として、ネガティブな体験談から観光地開発案を得ることが挙げられる。

第8章 おわりに

本研究では、観光地開発の支援技術として、長所情報をブログから収集し、分類する手法を提案した。長所情報は、開発目標地が非施設であることを考え、ブログから収集し、情緒推定によって長所の判定を行った。長所情報の分類は、長所情報の主文の用言意味属性で大別し、観光地における特徴的な名詞の特徴度で並べ替えることとした。

「山陰海岸」を開発目標地と設定し、類似する観光地として「江ノ島海岸」「若狭湾」、および「三陸海岸」における長所情報の収集と分類を本手法で行った。その結果、12,047文(444記事)のブログ記事から、2,894件(6,348文)の長所情報を取得した。取得した長所情報を、用言意味属性による分類と特徴的な名詞の特徴度による並べ替えによって分類した。分類結果を人手分析することで、新しい観光地開発案につながる情報を得ることができた。さらに、本手法による分類が人手分析の支援になっているかを確認するため、特徴度とヒントの出現数の関係を調査した。その結果、人手分析の支援となっていることを確認した。

今後の課題は、情緒推定結果が《P》以外になった文を調査対象にして、観光地開発案につながる情報を取得することである。

謝辞

はじめに，学部4回生から大学院1回生に渡って，研究に対する姿勢を教えて頂き，また，数々のご助言を頂きました故 池原悟先生にお礼を申し上げます．

本研究を進めるにあたり，様々なご助言を頂きました村田真樹教授にお礼を申し上げます．本論文をまとめるにあたって，ご指導頂きました川村尚生教授にお礼申し上げます．3年間に渡って，種々のご助言を頂きました村上仁一准教授にお礼を申し上げます．本研究を進める上で，終始に渡って多大なご協力，ご助言を頂きました徳久雅人講師にお礼を申し上げます．

また，参考にさせて頂いた論文の著者の方々，本研究にご協力頂いた計算機工学講座Cの皆様へ深く感謝致します．

参考文献

- [1] 高野敦子, 池奥涉太, 北村泰彦: “因果関係に着目した口コミ情報からの評判情報抽出”, 情報処理学会研究報告. 自然言語処理研究会報告, pp.79-86, 2007.
- [2] 石野亜耶, 小林大祐, 難波英嗣, 竹澤寿幸: “ブログを利用した観光情報リンク集の自動構築”, 言語処理学会第 16 回年次大会発表論文集, pp.246-249, 2010.
- [3] 山田忍: “形容詞を用いた観光評判情報の抽出に関する研究”, 修士学位論文, 高知工科大学大学院フロンティア工学コース, 2007.
- [4] 田中努, 徳久雅人, 村上仁一, 池原悟: “結合価パターンへの情緒生起情報の付与”, 言語処理学会第 10 回年次大会発表論文集, pp.345-348, 2004.
- [5] 黒住亜紀子, 徳久雅人, 村上仁一, 池原悟: “情緒状態と直結した情緒反応を表す用言の結合価パターン辞書の構築”, 電子情報通信学会ソサイエティ大会講演論文集, 基礎・境界, AS-5-2, pp.S-51-52, 2007.
- [6] 徳久雅人: “パターン言語処理に基づく情緒推定に関する研究”, 博士学位論文, 鳥取大学大学院工学研究科, 2008.
- [7] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦: 日本語語彙大系, 岩波書店, 1997.
- [8] Robertson, S. E., Walker, S., Jones, S., Hancock-Beaulieu, M. M., and Gatford, M.: “Okapi at TREC-3”, *Proc. of the 3rd Text REtrieval Conference*, 1994.