

# AdPsyche: 広告心理学に基づく選好データセット

三田雅人<sup>1</sup> 村上聡一郎<sup>1</sup> 本多右京<sup>1</sup> 岡達志<sup>2</sup>  
<sup>1</sup> 株式会社サイバーエージェント<sup>2</sup> 慶應義塾大学  
 {mita\_masato,murakami\_soichiro,honda\_ukyo}@cyberagent.co.jp  
 tatsushi.oka@keio.jp

## 概要

広告文生成は、個々のニーズに応じた広告文を自動生成する技術であり、広告文の誘引性を高めることが広告効果を向上させるうえで重要である。しかし、広告の魅力を左右する要因は十分に解明されておらず、広告効果の高い広告を作成する際の方略が立てにくい。本研究では、広告誘引性における影響要因を分析するために広告特有の特性を考慮した選好データセット AdPsyche を構築し、Bradley-Terry-Luce モデルを要因分析に援用することで、心理的特性が広告選好に与える影響を明らかにする。

## 1 はじめに

広告文生成は、商材説明やユーザ信号（検索クエリなど）をもとに、個々のニーズに合った広告文を生成する技術である。従来の広告制作ではコストがかかり、多様なニーズへの対応が困難であったが、ニューラル言語モデルを活用した自動生成技術はこれらの課題解決に寄与している。特に、エンコーダ・デコーダモデルや大規模言語モデル（LLM）の進展により、生成文の品質が著しく向上している [1, 2, 3, 4]。

広告文生成の技術的な課題の一つは、単に流暢で正確な文を生成するだけでなく、ユーザの潜在ニーズを捉え魅力的なテキスト生成に繋げることである。クリック率といった広告効果指標を向上させるには、魅力度（広告誘引性）を考慮することが重要であることが報告されている [4]。しかし、広告の魅力を左右する要因は十分に解明されておらず、特定の広告がユーザにとって魅力的である理由や、その背後にある選好の要因については未解明の部分が多い。

人間の選好に関する研究では、ペアワイズ比較と人間による評価値に基づく選好データセットが広く用いられている。例え

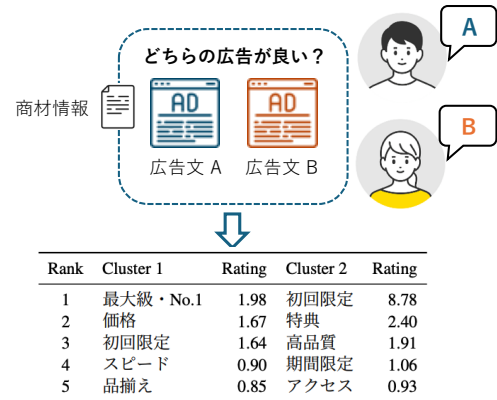


図 1: 広告選好判断と影響要因分析の概要。

ば、選好データセットには、Anthropic が公開している HH-RLHF<sup>1)</sup> や OpenAI が公開している summarize-from-feedback<sup>[5]2)</sup> などが公開されており、モデルに人間の価値観を学習させるために利用されている [6, 7, 8, 9]。しかし、広告の選好は判断者のパーソナリティ特性や商材との組み合わせから生じるバイアスなど広告特有の特性に影響を受けるにも関わらず、既存の選好データセットには広告ドメインに特化したものがないため、広告誘引性の影響要因を分析することが困難である。

そこで本研究では、広告誘引性における影響要因を分析するために、広告特有の特性を考慮した選好データセット AdPsyche を構築する<sup>3)</sup>。AdPsyche は、広告心理学で広く用いられる Big Five モデルをもとに定量化したユーザの心理的特性と訴求タイプ（例: 「価格」、「高品質」）を含む広告に特化した日本語選好データセットである。作成したデータセットを用い、Bradley-Terry-Luce (BTL) モデル [10] に基づき各訴求を要因とみなして影響要因分析をすることで、どの訴求タイプがどのユーザクラスターの広告選好に影響を与えるかを明らかにする (図 1)。

1) <https://huggingface.co/datasets/Anthropic/hh-rlhf>

2) <https://github.com/openai/summarize-from-feedback>

3) 構築したデータセットは今後一般公開する予定である。

## 2 関連研究

### 2.1 人間の選好に関する要因分析

人間の選好判断に関する要因分析を行った研究は限られている。Hu ら [11] は、OpenAI が公開した要約の選好データセット [5] を対象に、GPT-4 を用いて要約から要因候補を抽出し、BTL [10] を適用することで影響要因分析を行った。しかし、この手法を広告選好に単に適用するには限界がある。広告には訴求表現（例：「品質」や「価格」）と呼ばれる選好に影響を与える重要な要因があるが、これらは商材との組み合わせによるバイアスが生じる。例えば、EC 系の広告では価格に関する訴求が響くが、車のような高い商品は品質などの訴求の方が重要視される [12]。Hu らの手法は、主に言語的要因（例：流暢性、文法性など）および内容の正確さに関する要因（例：幻覚を含むかどうか）を独立的に扱うため、こうした言語的要因（訴求表現）と内容に関する要因の組み合わせによる影響を十分に考慮できない。本研究は、Hu らの方法論を拡張し、心理的特性と広告特有の要因を統合した分析基盤を構築し、新たな広告誘引性の影響要因を明らかにする。

### 2.2 Big Five と広告心理学

Big Five (Big 5) は、心理学において広く用いられるパーソナリティ特性モデルであり、外向性、神経症傾向、開放性、協調性、誠実性の 5 因子から構成される [13]。これらの特性は、個人の行動や嗜好を予測するために利用され、マーケティングや広告研究においても重要な指標となっている。広告心理学では、個人のパーソナリティ特性が広告の受容や購買行動に与える影響が注目されている [14]。例えば、外向性の高い人は感情的で自己価値を高める訴求に惹かれやすい一方、誠実性が高い人は信頼性や実用性を重視する傾向があるとされる。また、神経症傾向が高い人は不安や安心感を訴求する広告に、開放性の高い人は革新的で新しいアイデアを訴求する広告に好意的であることが示されている [15]。これらの知見は、ターゲティング広告の最適化において活用されている。本研究では、Big Five モデルを活用し、広告嗜好の要因分析を通じて広告心理学への新たな洞察を提供する。

表 1: Big Five に基づくクラスタリング結果。

ID	神経症傾向	外向性	開放性	協調性	誠実性
C0	0.7	-0.8	-0.6	0.2	-0.1
C1	-0.3	0.4	0.3	0.1	0.1
C2	0.9	-0.8	-1.1	-1.9	-1.5
C3	-1.4	1.4	1.4	0.5	0.8

## 3 AdPsyche

### 3.1 設計方針

広告特有の特性を考慮した選好データセットの構築をするにあたり、次の 2 つの基本要件を定めた：(1) 同一商材に対して異なる訴求をもつ広告文ペアであること；(2) 選好判断にはユーザの心理的特性に関する情報を付与すること。要件 1 は、2.1 節で述べたように、訴求表現と商材との組み合わせによるバイアスを除去するため既定した。また、同一商材に対して異なる訴求表現をもつ広告文でペアワイズ比較することにより、どのような訴求が好まれる・好まれないといった要因を分析するための基盤を提供できる。要件 2 は、広告の魅力度がユーザの性格特性に依存することが報告されている [15] ため、性格特性を考慮した訴求要因分析を可能となるよう設計する。本研究では、ユーザの性格特性の測定をするために Big Five モデルを採用する。

### 3.2 ペアワイズ比較の作成

アノテーション対象のデータとして、CAMERA<sup>3</sup> [16] を用いる。CAMERA<sup>3</sup> は、日本語広告文生成ベンチマーク CAMERA [4] の評価セットをもとに構築されており、ランディングページ (LP) に記載の商材文書 (LP テキスト) に対して訴求タイプごとに人手で作成された多様な広告文が含まれている。商材の多様性の担保と業種 (商材タイプ) ごとの要因分析を可能にするため、CAMERA の評価セットに付与されている人材、EC、金融、教育の各業種から 10 件ずつ広告アセット (LP テキスト、広告文、広告訴求ラベル、の三つ組) をランダムに抽出した。次に、ペアデータを作成するため、同一広告アセット内の広告文の組み合わせでペアを作成する。例えば、広告文が 5 つ付与されている広告アセットの場合は、 ${}^5C_2 = 10$  個の広告ペアデータが作成される。最終的に 343 件のペアデータが作成

表 2: 最も好まれた訴求タイプ上位 10 位 (ユーザクラス別)。

順位	C0		C1		C2		C3	
	訴求タイプ	値	訴求タイプ	値	訴求タイプ	値	訴求タイプ	値
1	初回限定	2.64	初回限定	2.26	初回限定	2.58	初回限定	2.96
2	最大級・No.1	2.12	最大級・No.1	2.22	期間限定	2.56	還元	2.74
3	対象限定	2.08	特典	2.21	対象限定	2.18	特典	2.12
4	期間限定	1.98	期間限定	1.95	その他のオファー	1.76	期間限定	1.86
5	特典	1.86	価格	1.95	価格	1.71	対象限定	1.79
6	無料	1.83	対象限定	1.91	還元	1.46	無料	1.68
7	価格	1.72	その他のオファー	1.78	最大級・No.1	1.36	価格	1.60
8	還元	1.52	還元	1.77	特典	1.28	最大級・No.1	1.47
9	その他のオファー	1.49	無料	1.74	スピード	1.14	その他のオファー	1.26
10	スピード	1.34	スピード	1.50	無料	1.11	スピード	1.19

された<sup>4)</sup>。

### 3.3 アノテーション

クラウドソーシング<sup>5)</sup>を通じて、109 人募った。全体の男女比は 6:4 になり、年齢は 20 代から 60 代まで幅広く、30 代 (31%) と 40 代 (43%) が多数であった。アノテーションでは、心理的特性の測定と広告選好判断という二つのタスクを実施した。心理的特性の測定には、並川ら [17] が作成した日本語版 Big Five 尺度の 29 項目に基づき 5 段階のリッカード尺度 (1: 全く当てはまらない 5: 非常に当てはまる) によるアンケートを使用した。広告選好判断では、LP テキストと 2 つの広告文ペアを提示し、この商材の広告文としてより魅力的な広告文を選択させた。なお、どちらも魅力度は同じで優劣の判断がつかない場合のみ「どちらも魅力度は同じ」を選択可能とした。さらに、順番効果を緩和するため、参加者を 5 グループに分けて広告文の提示順序をランダム化し、またチェック設問を用意したうえで正解率が低い判断者はアノテーションに含めないなど品質担保に努めた。

### 3.4 分析

事前アンケートにて収集した Big Five の評価値をもとに判断者の心理的特性を分析する。具体的には、各判断者の 5 因子 (2.2 節参照) ごとの評価値からそれぞれ z 値を算出し、K=4 の K-means [18] クラスタリングを行った<sup>6)</sup>。Big Five に基づくクラスタリング結果を表 1 に示す<sup>7)</sup>。各ユーザクラスは表 1 の 5 因子におけ

る評価値により、概ね以下のように特徴付けられる：

- 感情的にやや不安定で内向的な人々 (C0)
- バランスの取れた性格の人々 (C1)
- 感情的に不安定で新しい経験や計画性が低い人々 (C2)
- 外向的で協調性や誠実性も高いポジティブな人々 (C3)

## 4 広告誘引性の影響要因分析

### 4.1 問題設定

本研究で用いるデータセットは、 $N$  個の広告文ペア  $\mathcal{D} = \{(S_1^{(n)}, S_2^{(n)})\}_{n=1}^N$  で構成される。各ペアは同一の商材に対して異なる訴求タイプに沿って作成された広告文  $S_1^{(n)}$  と  $S_2^{(n)}$  を含み、これに対する人間の評価として、優れた広告文  $\hat{S}^{(n)}$  が指定される。広告を多面的に評価するため、以下の  $M$  個の訴求タイプに関連する要因の集合  $A = \{a_i\}_{i=1}^M$  を導入する。要因間の優劣を評価するために、Hu ら [11] に従い、BTL モデルを使用する。このモデルでは、勝敗の確率が要因間の相対的な「強さ」の差に依存すると仮定する。具体的には、広告文ペアにおいて  $S_1^{(n)} \succ S_2^{(n)}$  (すなわち  $S_1^{(n)}$  が選好される) 場合、次の関係が成り立つとみなす：

$$a_1^{(n)} \succ a_2^{(n)}$$

ここで、本設定では各広告文につき 1 つの要因が関連付けられているため、 $S_1^{(n)}$  にのみ  $a_1^{(n)}$  が存在し、 $S_2^{(n)}$  にのみ  $a_2^{(n)}$  が存在する。本研究の目的は、人間の評価に影響を与える主要な要因を特定することである。このため、BTL モデルを用いて各要因の相対的な重要度  $\{p_i\}_{i=1}^M$  を推

4) 作成したペアデータの例を付録 A に示す。

5) <https://www.lancers.jp/mypage?>

6) クラスタ数はエルボー法を用いて決定した。

7) PCA によるクラスターの可視化を付録 B に示す。



表 3: 最も好まれた訴求タイプ上位 10 位 (業種別).

順位	人材		金融		EC		教育	
	訴求タイプ	値	訴求タイプ	値	訴求タイプ	値	訴求タイプ	値
1	価格	2.84	初回限定	4.75	特典	3.10	その他のオファー	2.36
2	初回限定	2.45	その他のオファー	2.62	価格	2.67	期間限定	2.00
3	最大級	1.62	期間限定	2.12	無料	2.31	無料	1.94
4	無料	1.46	対象限定	2.09	還元	1.89	特典	1.90
5	品揃え	1.36	特典	1.95	初回限定	1.72	その他の特徴	1.86
6	スピード	1.24	最大級	1.86	その他のオファー	1.61	悩み解決	1.70
7	実績	1.22	価格	1.72	期間限定	1.37	高品質	1.70
8	特徴	1.02	利便性	1.54	悩み解決	1.17	実績	1.50
9	利便性	0.75	無料	1.40	最大級	1.09	対象限定	1.33
10	悩み解決	0.69	高品質	1.06	利便性	1.05	利便性	1.25

定する. この推定は以下の反復式を用いて行われる:

$$p'_i = \frac{W_i}{\sum_{j \neq i} \frac{w_{ij} + w_{ji}}{p_i + p_j}}, \quad p_i = \frac{p'_i}{\sum_{j=1}^M p'_j}$$

## 4.2 心理的特性に基づく選好要因

表 2 は, BTL モデルによる推定値に基づいたユーザクラスごとに好まれた訴求タイプを示している. ここで BTL モデルによって算出された各推定値はその要因が他の要因と比較した際の相対的な強さを示している. 例えば, 「初回限定」の推定値が 4.75, 「スピード」の推定値が 2.62 であれば, 「初回限定」が「スピード」よりも選ばれる確率は次の式で計算できる:

$$P(\text{初回限定が勝つ}) = \frac{4.75}{4.75 + 2.62} \approx 0.644$$

つまり, 約 64.4% の確率で「初回限定」という訴求が「スピード」という訴求に勝つ (好まれやすい) と解釈できる.

全ユーザクラスに共通して「初回限定」への評価が高いことから, 新規性を訴求する施策は全般的に有効であることが示唆される. 一方で, ユーザクラスごとの特性や傾向の違いも確認できる. C0 は「最大級・No.1」や「特典」への評価が高い一方, 「スピード」への関心が低い. 確実性や品質の高さを重視するため, 特典付きで信頼感を強調する訴求が有効である. C1 は「価格」への評価が最も高いため, コストパフォーマンスを重視する傾向があり, 質と価格を両立した訴求が効果的である. C2 は, 「その他のオファー」や「スピード」への評価が高いことから, 多様性や迅速な対応を重視するため, 柔軟で即応性の高い提案が適している. C3 は, 「還元」への評価が最も高いことから実質的な利益

や信頼性を重視するため, 長期的なメリットを強調した施策が効果的であると考えられる.

## 4.3 業種毎の選好要因

表 3 は BTL モデルによる推定値に基づいた業種ごとに好まれた訴求タイプを示している. 人材系の商材では, 「価格」や「初回限定」が高評価を得ており, コスト意識や信頼性が重要視される傾向がある. 一方, 金融系の商材では, 「初回限定」や「その他のオファー」が特に効果的であり, 利用開始時のインセンティブや限定性が鍵となることが示唆された. EC 系の商材では, 「特典」や「価格」が上位を占め, 金銭的利益や付加価値を訴求する広告が有効である. 教育系の商材では, 「その他のオファー」や「期間限定」が高評価を得ており, 希少性や特典の訴求が好まれる. これらの結果から, 業種によって広告で重視される訴求タイプが大きく異なり, 顧客層や商材特性に応じた柔軟な広告戦略が必要であることが示された. 特に, 限定性や価格訴求は複数の業種で高評価を得ており, 汎用的な効果もあると考えられる.

## 5 おわりに

本研究では, 広告選好における心理的特性の影響を明らかにするため, Big Five モデルに基づいた選好データセット AdPsyche を構築し, BTL モデルを用いた要因分析を実施した. その結果, 特に「限定性」や「特典」の訴求が多くユーザクラスで好まれる一方, 心理的特性や業種によって選好が異なることが示された. これらの結果は, ユーザ特性に応じたパーソナライズ広告の重要性が強調すると同時に, ターゲティング広告の最適化に新たな視点を提供した.

## 参考文献

- [1] J. Weston Hughes, Keng-hao Chang, and Ruofei Zhang. Generating better search engine text advertisements with deep reinforcement learning. In **Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, KDD '19, p. 2269–2277. Association for Computing Machinery, 2019.
- [2] Shaunak Mishra, Manisha Verma, Yichao Zhou, Kapil Thadani, and Wei Wang. Learning to create better ads: Generation and ranking approaches for ad creative refinement. CIKM '20, p. 2653–2660. Association for Computing Machinery, 2020.
- [3] Hidetaka Kamigaito, Peinan Zhang, Hiroya Takamura, and Manabu Okumura. An empirical study of generating texts for search engine advertising. In **Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Papers**, pp. 255–262. Association for Computational Linguistics, 2021.
- [4] Masato Mita, Soichiro Murakami, Akihiko Kato, and Peinan Zhang. Striking gold in advertising: Standardization and exploration of ad text generation. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, **Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 955–972, Bangkok, Thailand, August 2024. Association for Computational Linguistics.
- [5] Nisan Stiennon, Long Ouyang, Jeff Wu, Daniel M. Ziegler, Ryan Lowe, Chelsea Voss, Alec Radford, Dario Amodei, and Paul Christiano. Learning to summarize from human feedback. In **Proceedings of the 34th International Conference on Neural Information Processing Systems**, NIPS '20, Red Hook, NY, USA, 2020. Curran Associates Inc.
- [6] Nisan Stiennon, Long Ouyang, Jeffrey Wu, Daniel Ziegler, Ryan Lowe, Chelsea Voss, Alec Radford, Dario Amodei, and Paul F Christiano. Learning to summarize with human feedback. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, **Advances in Neural Information Processing Systems**, Vol. 33, pp. 3008–3021, 2020.
- [7] Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback. **arXiv preprint arXiv:2203.02155**, 2022.
- [8] Jacob Menick, Maja Trebacz, Vladimir Mikulik, John Aslanides, Francis Song, Martin Chadwick, Mia Glaese, Susannah Young, Lucy Campbell-Gillingham, Geoffrey Irving, and Nat McAleese. Teaching language models to support answers with verified quotes. **arXiv preprint 2203.11147**, 2022.
- [9] Rajkumar Ramamurthy, Prithviraj Ammanabrolu, Kianté Brantley, Jack Hessel, Rafet Sifa, Christian Bauckhage, Hannaneh Hajishirzi, and Yejin Choi. Is reinforcement learning (not) for natural language processing: Benchmarks, baselines, and building blocks for natural language policy optimization. **arXiv preprint 2210.01241**, 2023.
- [10] Ralph Allan Bradley and Milton E. Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. **Biometrika**, Vol. 39, No. 3/4, pp. 324–345, 1952.
- [11] Yebowen Hu, Kaiqiang Song, Sangwoo Cho, Xiaoyang Wang, Hassan Foroosh, and Fei Liu. DecipherPref: Analyzing influential factors in human preference judgments via GPT-4. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, **Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing**, pp. 8344–8357, Singapore, December 2023. Association for Computational Linguistics.
- [12] Soichiro Murakami, Peinan Zhang, Sho Hoshino, Hidetaka Kamigaito, Hiroya Takamura, and Manabu Okumura. Aspect-based analysis of advertising appeals for search engine advertising. In **Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Track**, pp. 69–78. Association for Computational Linguistics, 2022.
- [13] Sebastiaan Rothmann and Elize Coetzer. The big five personality dimensions and job performance. **South African Journal of Industrial Psychology**, Vol. 29, pp. 68–74, 10 2003.
- [14] Jennifer L. Aaker. Dimensions of brand personality. **Journal of Marketing Research**, Vol. 34, No. 3, pp. 347–356, 1997.
- [15] 志織原, 奈々子水津, 利紘分部. 購買行動を規定する性格特性の検討. 福岡女学院大学紀要. 人間関係学部編, Vol. 20, pp. 61–64, 2019.
- [16] Go Inoue, Akihiko Kato, Masato Mita, Ukyo Honda, and Peinan Zhang. CAMERA<sup>3</sup>: An evaluation dataset for controllable ad text generation in Japanese. In Nicoletta Calzolari, Min-Yen Kan, Veronique Hoste, Alessandro Lenci, Sakriani Sakti, and Nianwen Xue, editors, **Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)**, pp. 2702–2707, Torino, Italia, May 2024. ELRA and ICCL.
- [17] 並川努, 谷伊織, 脇田貴文, 熊谷龍一, 中根愛, 野口裕之. Big five 尺度短縮版の開発と信頼性と妥当性の検討. 心理学研究, Vol. 83, No. 2, pp. 91–99, 2012.
- [18] J MacQueen. Some methods for classification and analysis of multivariate observations. In **Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability/University of California Press**, 1967.

## A ペアワイズデータセットの例

表 4: 広告ペアワイズ評価データセットの例.

業種	LP テキスト	広告文 A	広告文 B	訴求タイプ A	訴求タイプ B
教育	... スクールの受講実績はスクール業界 No.1! パソコン資格 (MOS) ...	参加者限定選べるプレゼント	北海道から鹿児島まで全国展開	対象限定	アクセス
金融	... 自動車火災保険「おとなの自動車保険」はあなたにあった充実の補償内容がえらべてお得! ...	簡単お見積もり無料	全国 13,000 ヶ所のサービス拠点	無料	アクセス
EC	外反母趾や甲高、幅広の方にも履きやすい靴。柔らかいシープスキン...	ポイント還元	返品・交換送料無料で安心	還元	無料
人材	即戦力 & 若手の人材採用に強い! オープニングスタッフ募集も...	初回限定割引	豊富な求人広告掲載実績	初回限定	その他の実績

## B PCA によるユーザクラスタの可視化

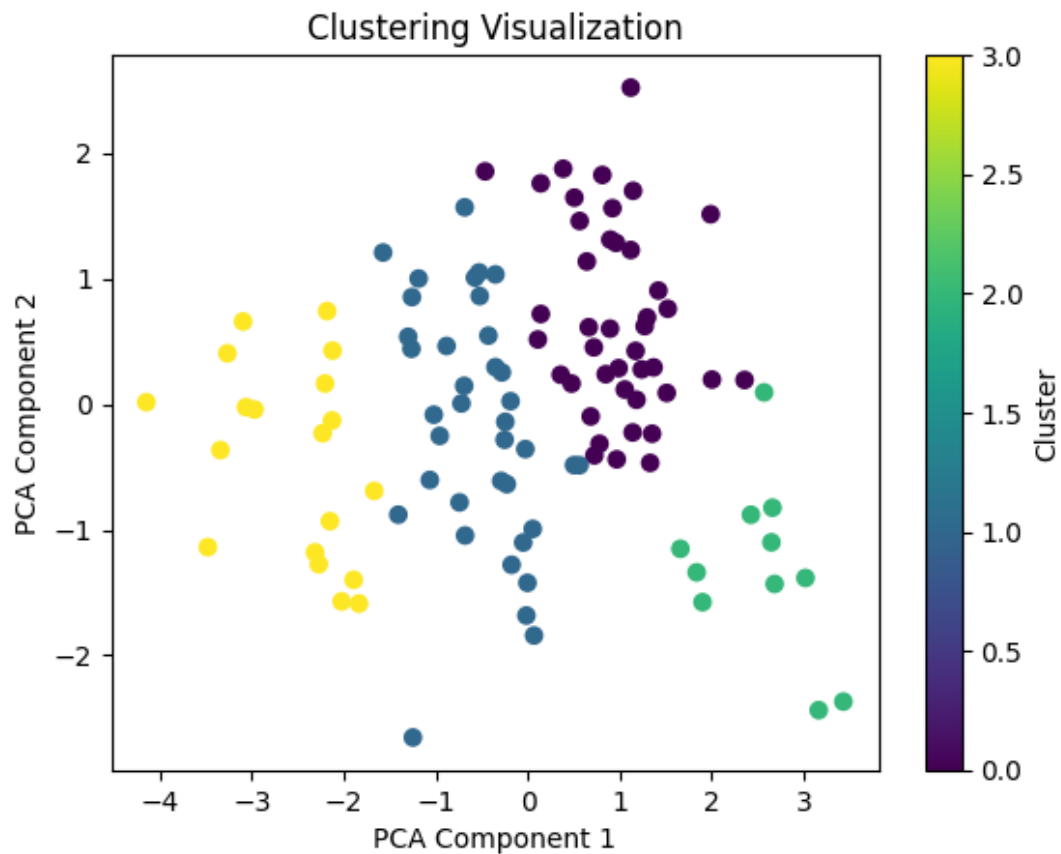


図 2: PCA によるクラスター可視化.