

## ブログからの料理の失敗事例の検出

辻田 美穂<sup>†</sup> 土居 洋子<sup>†</sup> 難波 英嗣<sup>†</sup> 竹澤 寿幸<sup>†</sup> 角谷 和俊<sup>‡</sup>

<sup>†</sup> 広島市立大学情報科学部 〒731-3194 広島市安佐南区大塚東 3-4-1

<sup>‡</sup> 兵庫県立大学環境人間学部 〒670-0092 兵庫県姫路市新在家本町 1-1-12

E-mail: <sup>†</sup> {nanba, doi, tsujita, takezawa}@ls.info.hiroshima-cu.ac.jp, <sup>‡</sup> sumiya@shse.u-hyogo.ac.jp

あらまし ブログ集合の中から料理の失敗事例について述べたエンTRIESを自動検出する手法を提案する。まず、「[料理名] AND レシピ AND 失敗」などのクエリを用いて、ブログエンTRIESを収集する。次に、特許から収集した手掛かり語を素性とした機械学習に基づく手法により、料理失敗ブログを検出する。提案手法の有効性を確認するため、1,046 件のブログエンTRIESを用いて実験を行った。実験の結果、精度 0.527、再現率 0.100 が得られた。

キーワード レシピ, 料理失敗, ブログ, テキスト分類

## Detecting Cooking Blog Entries Mentioned about Failed Meals

Miho TSUJITA<sup>†</sup> Yoko DOI<sup>†</sup> Hidetsugu NANBA<sup>†</sup>

Toshiyuki TAKEZAWA<sup>†</sup>, and Kazutoshi SUMIYA<sup>‡</sup>

<sup>†</sup> Faculty of Information Sciences, Hiroshima City University  
3-4-1 Ozukahigashi, Asaminamiku, Hiroshima 731-3194 Japan

<sup>‡</sup> School of Human Science and Environment, University of Hyogo  
1-1-12 Shinzaike-honcho, Himeji, Hyogo 670-0092 Japan

E-mail: <sup>†</sup> {nanba, doi, tsujita, takezawa}@ls.info.hiroshima-cu.ac.jp, <sup>‡</sup> sumiya@shse.u-hyogo.ac.jp

**Abstract** We propose a method that detects blog entries mentioned about failed meals. First, we collected blog entries using queries, such as “[dish name] AND recipe AND failure.” Second, we identified blog entries that mention about failed meals based on machine learning with several cue phrases collected from patents as features. To confirm the effectiveness of our method, we conducted an examination using 1,046 blog entries. From the experimental results, we obtained precision of 0.527 and recall of 0.100.

**Keyword** Recipe, Failed Meal, Blog, Text Classification, Patent

### 1. はじめに

人間にとって生活するための基盤となるもの、それは衣食住である。その中でも、食に関しては生きていく上で最も必要であり、人間を含めたすべての生き物にとっては切っても切り離せない存在である。その食を豊かに、そして鮮やかに演出するために用いられる料理レシピとは、人々の経験や工夫の蓄積を経て生まれたものであり、その知識は共有され、継承される。料理レシピはインターネットの普及により、レシピサイトを通して多くの人々が手軽に検索・閲覧できるものとな

った。しかし、料理初心者や時間のない人にとって料理レシピのみを参考に、限られた時間で満足する料理を作ることは容易ではない。また、失敗によって生じる材料の買い足し、外食への変更といったコストの問題も挙げられる。

そこで本研究では、情報共有の手段であり、インターネット上で気軽に投稿・閲覧できるブログに着目する。ブログ上では、個人の意見や感想が自由に記述されており、レシピサイトなどには記載されていない情報が豊富にある。その一つに、失敗談がある。ブログには実際に経験した失敗、

注意点が述べられている。失敗内容を収集、分析し、主な失敗事例を料理レシピに提示することにより、料理未経験者には事前に確認、注意しながら調理でき、失敗を減らす手助けを行うことができる。本研究の最終的な目標は、失敗ブログを自動的に抽出し、失敗箇所を料理レシピに提示することで、料理初心者の支援をすることである。そのための第一歩として、事前に定めた 35 種類の料理を対象に、「料理名 レシピ 失敗」などのクエリを用いて収集したブログが失敗事例について述べたものであるかどうかの自動判定を行う。

本論文の構成は以下のとおりである。2 節では、関連研究について述べる。3 節では、料理失敗ブログを自動判定するために、必要なブログの収集や提案手法、提案手法用いるデータについて述べる。4 節では、実験について述べ、その結果を報告する。5 節で本稿をまとめる。

## 2. 関連研究

志土地ら[1]は、ユーザの調理レベルを考慮し、熟練度に応じた料理レシピの作成手法を提案している。志土地らの手法では、料理初心者に着目し、イメージしにくい料理動作に画像や映像、補足説明を付加することでわかりやすい料理レシピの作成を行っている。本研究では、実際の調理経験に基づいて料理レシピの作成を行い、料理初心者を支援する点で異なる。

橘ら[2]は、レシピの料理名にある「簡単」や「子どもが喜ぶ」や「ヘルシー」といった修飾表現に着目し、それらの修飾表現の根拠(ネーミングコンセプト)をレシピから抽出する手法を提案している。橘らの手法では、ネーミングコンセプトを抽出する際、レシピの材料と調理器具に着目して典型的な要素との差異を抽出することでレシピの特徴を把握し、ユーザに提示しているが、本研究では、個々の料理の失敗しやすい箇所、特徴を把握し、ユーザに提示する点で異なる。

山田ら[3]は、手持ちの材料のみを用いて調理できる料理レシピが多く検索できるシステムを開発する研究を行っている。料理レシピ内の特徴的な単語や文章から材料名などを切り出し、手持ちの材料群と照らし合わせることで、その料理レシピが手持ちの材料のみを用いて調理可能であるかを判別し、ユーザへ返すシステムを構築している。本研究で構築を目指している失敗事例データベースと合わせることで発展的な調理者支援システム開発を行うことができる可能性がある。

料理レシピ全体を理解することを目的としている森らの研究[4]がある。この研究では、調理手順が一目でわかるフローグラフに着目し、そのフローグラフの定義と、それに基づくコーパス作成の手法を提案している。森らは、材料などの固有表現認識に機械学習を用いている。機械学習を用いる点で本研究と類似しているが、機械学習を行う対象がブログである点で異なる。

本研究に関連する研究として、Hamon ら[5]の料理レシピ内の単語の自動検出がある。Hamon らは、フランス語の料理レシピ内の単語を検出するために、ルールベースと機械学習、2 つのシステムを使用し組み合わせる手法を提案している。Hamon らは、フランス語の料理レシピを対象としているが、本研究では、日本語の料理レシピを対象としている点で異なる。

## 3. 料理失敗ブログの自動判定

本節では、料理失敗事例について述べられている料理失敗ブログの収集手法について説明を行う。料理失敗ブログの収集手法は、以下の 3 つのステップに分かれている。

- (1) 料理失敗候補ブログの収集
- (2) 人手による料理失敗ブログの判定基準
- (3) 料理分野の特許による素性の収集

(1)については 3.1 節、(2)については 3.2 節、(3)については 3.3 節で、それぞれ説明を行う。

### 3.1. 料理失敗候補ブログの収集

料理失敗ブログの収集のためには、料理の失敗に関する情報が記載されている可能性の高い、料理失敗候補ブログを収集する必要がある。そこで本研究では以下に示す 5 種類クエリを使用し、Yahoo!検索(ブログ)で検索を行い、ブログの収集を行った。その結果、1,046 件のブログが収集された。

- [料理名] AND レシピ AND 失敗
- [料理名] AND 料理 AND 反省点
- [料理名] AND レシピ AND 反省点
- [料理名] AND 調理ミス
- [料理名] AND 失敗作

### 対象料理の設定

図 1 に対象料理の一覧を示す。対象料理の選定は、和食/洋食/中華、あるいはメインディッシュ/サイドディッシュ/デザートなど、35 種類が特定のカテゴリの料理に偏らないよう考慮してある[6]。

001 酢豚	019 梅酒
002 角煮	020 パンパンジー
003 ミートソースパスタ	021 高野豆腐
004 豚肉のしょうが焼き	022 エビフライ
005 パンプディング	023 大学いも
006 イチゴジャム	024 麻婆豆腐
007 パパロア	025 だし巻き卵
008 みたらし団子	026 おはぎ
009 おからハンバーグ	027 ゴーヤチャンプルー
010 パンプキンスープ	028 水餃子
011 カレーピラフ	029 スパニッシュオムレツ
012 カステラ	030 カスタードプリン
013 あんぱん	031 唐揚げ
014 かぼちゃコロッケ	032 カルボナーラ
015 サバの味噌煮	033 親子丼
016 カキフライ	034 茶碗蒸し
017 切干大根	035 エビマヨ
018 いなりずし	

図 1: 複数レシピ要約に用いた対象料理一覧

### 3.2. 人手による料理失敗ブログの判定基準

本研究での料理失敗ブログとは、ブログ著者が自ら対象料理を調理し、その過程や完成品において失敗だと投稿したブログと定義する。3.1 節で収集した料理失敗候補ブログには、料理に関連のないものや、自ら調理したブログではなくお店での行動、味の感想が失敗であるというもの、また対象料理を調理したが失敗したのは別の料理であるというものなどが含まれる。より正確に、料理失敗事例を抽出する必要があるため、料理失敗候補ブログに対して、人手で料理失敗ブログの判定を行う。人手で料理失敗ブログと判定した例を図 2 に示す。

<dish>酢豚</dish>  
 でも、これから暑くなるし、健康のため酢豚を作りました♪  
 . . . . . (略) . . . . .  
 今回の反省点。  
 片栗粉の量が多すぎて、あんがちょっと固めになっちゃったので気をつけよう。  
 パインは今回入れなかったけど、やっぱり入れたほうが美味しいので、次回は入れよう。

図 2: 料理失敗ブログと判定した例

### 3.3. 料理分野の特許明細書からの手掛かり語の収集

料理失敗ブログの自動判定を行うには、失敗事例と判定するための手掛かり語を数多く収集する必要がある。もし、料理の失敗事例について書かれた大量のブログエントリがあれば、情報利得などの手法を用いて効率的に手掛かり語を収集することができる。しかし、残念ながら我々の研究目的に合致するデータは見当たらなかった。そ

こで、効率的に手掛かり語を収集する別の方法として特許に着目した。特許明細書には、何らかの課題と、それを解決するための発明が記載されている。料理分野に限定すれば、この課題として、料理の失敗に関する表現も数多く出現すると考えられる。したがって、料理分野の特許明細書から、以下の 3 種類の手法を用いて手掛かり語の収集を行う。

- (1) 「A~」における料理関連用語の抽出
- (2) 「~B」における料理関連用語の抽出
- (3) 「~などの失敗」における料理関連用語の抽出[5]

なお、料理分野の特許明細書として、国際特許分類(IPC)のサブクラスレベルで A23L(食品、食料品)、A47J(台所用具)、H05B(電気加熱)が筆頭 IPC(ひとつの特許明細書に付与される複数の分類コードの中で一番重要なもの)として付与されたものを 1993~2012 年の公開特許公報データベースから計 91,736 件抽出し、それらから手掛かり語を収集する。

#### (1) 「A~」における料理関連用語の抽出

手掛かり語を収集する第一の方法として、「[名詞句]+の+[名詞句]」という定型表現に着目した。例えば「A の理由」や「A の原因」という定型表現にマッチする A を収集すると、「焦げ」や「半生」など、失敗に関連する表現が効率的に収集できると考えられる。そこで、まず「失敗の B」という表現を用いて、B に該当する表現(例えば、「理由」や「原因」)を収集し、次にこれらの B に関する表現をもとに「A の B」で A に該当する用語を収集することで、失敗に関する手掛かり語を収集した。収集した用語には様々な失敗の種類があり、失敗要因によって分類することが可能である。そのため、失敗の種類をカテゴリと定義し、手掛かり語を「焦げ」や「割れ」などの 6 つのカテゴリに人手で分類して収集する。実際に収集した 56 語の手掛かり語を分類した結果を表 2 に示す。本研究では、カテゴリに分類した手掛かり語を素性として用いる。

表 2: 料理関連用語「A～」から  
収集した手掛かり語(56 語)

カテゴリ	手掛かり語
焦げ	焦げ, 焦げ付き, 焦げ臭, コゲ, 焼きむら, 黒いシミ, 焦げカス
割れ	割れ, 偏り, キズ, 縦割れ, 胴割れ, 破れ
生煮え	生煮え, 生焼け, 半焼け, 半焼き, 予熱不足, 半調理, だろだろ, 半生
味	食感, 味付け, 食味, 風味, 味覚, 味質, 味, 風味品質, 味全体, 味わい, 旨味
失敗	失敗, 調理後, 悪化, 誤認識, 欠落, 誤操作, 脱落, 誤認, 食感不良, 低下, 劣化, 食欲低下, 故障, 効率低下, 食欲低下
原因	原因, 主原因, 要因, 特徴, 理由, 改善, 主要因, 重点, 要素

表 3: 料理関連用語「～B」から  
収集した手掛かり語(71 語)

カテゴリ	手掛かり語
焦げ	焦げ, 焦げ付き, 上下ムラ, 水分不足, 焦げ発生, コゲ, 硬め, 焦げ色, 焦げ付き防止, 変色, 焦げ付き部, 焼け焦げ, 焼けすぎ, 焼きむら, 焼きムラ, 焼け具合, 焦げ過ぎ, 焼き具合
割れ	形崩れ, 割れ, 縦割れ, 割れ防止, 割れ現象, 胴割れ, 型崩れ, 切り出し
生煮え	生煮え, 火力不足
味	味付け, 食感, 味, 風味, 味覚, 味等, 香味
失敗	失敗, 調理後, 悪化, 失敗防止, 出来栄え, 出来上がり, 終了, 低下, 変化, 半減, 硬化現象, 減少, 状態, 終点, 内容, 劣化
原因	原因, 主原因, 要因, 発生, 問題, 危険
注意	違和感, 注意, 不満, 不注意, 不安感
不足	不足, 欠乏, 過不足部分, 不足気味, 過不足, 損失量
苦手	苦手, 下手, 初心者

## (2) 「～B」における料理関連用語の抽出

手掛かり語を収集する第二の方法として、第一の手法と同様に「[名詞句]+の+[名詞句]」という定型表現を用いた。ただし、第二の方法では、「AのB」でBに該当する用語に着目する。Aとして、例えば「サバ」や「大根」などの料理の材料名や料理名を入力すれば、該当するBには「生焼け」や「生煮え」といった表現が出力される。これらを人手で選定し、さらに9つのカテゴリに分類する。実際に収集した71語の手掛かり語を分類した結果を表3に示す。

## (3) 「～などの失敗」における料理関連用語の抽出

手掛かり語を収集する第三の方法として、定型表現「A(など|等)のB」に着目する。一般に、この定型表現は用語の上位、下位関係の抽出に利用される[7]が、本稿では、料理の失敗に関する手掛かり語の収集に利用する。具体的には「A(など|等)の失敗」という表現のAに該当する用語を収集する。収集した用語を「焦げ」や「割れ」などの7つのカテゴリに人手で分類して収集する。実際に収集した23語の手掛かり語を分類した結果を表4に示す。

表 4: 料理関連用語「～などの失敗」から  
収集した手掛かり語(23 語)

カテゴリ	手掛かり語
焦げ	焦げ, 焦げ付く, ふきこぼれる, 飛散, 硬い
割れ	割れ, 崩れ
生煮え	生煮え
味	味, 濃い, 食感
失敗	失敗, 悪化, 低下, 調理失敗, 調理ミス, ミス, 悪い, 過ぎ
原因	原因, 要因
注意	注意, 違和感

## 3.4. 料理失敗ブログの自動判定

### 単語の出現頻度を用いた手法 (提案手法 1)

3.1 節で述べた手法により収集した料理失敗候補ブログ 1,046 件に対して形態素解析を行い、品詞が動詞、名詞、形容詞である単語を用いる。なお、数字、英字以外の単語であり、単語の長さが1文字以上である単語を対象とする。人手により

料理失敗ブログと判定されたブログを対象に、出現頻度が閾値以上の単語を素性として機械学習を用いる。閾値は、単語の出現回数が 50 から 100 までを 10 刻みにより、最も結果の良かった閾値を用いる。

#### 料理分野の特許を用いた手法（提案手法 2）

3.3 節の(1)(2)(3)で述べた料理分野の特許明細書から収集した手掛かり語(表 2、3、4)の各エントリ中における出現の有無を素性とし、機械学習を用いて料理失敗ブログの判定を行う。

### 4. 実験

#### 4.1. 実験方法

本研究で行った実験について説明する。

##### データセット

実験用データには、3.1節で収集したYahooブログ1,046件に対し、人手で料理失敗ブログかどうかの判定を行った結果を用いる。人手で料理失敗ブログと判定した結果を表5に示す。

表5: 料理失敗ブログの人手での判定結果

料理失敗ブログ(件)	その他(件)	合計(件)
388	658	1,046

##### ベースライン手法

提案手法の有効性を確認するため、3.1節で収集した料理失敗候補ブログ1,046件のブログのデータに対して形態素解析を行い、そこで得られた全単語を素性として、料理失敗ブログの判定を行った場合をベースライン手法とした。なお、形態素解析にはMeCabを用いた。

##### 機械学習と評価尺度

料理失敗ブログの判定の機械学習にはTinySVMを用いた。1次の線形カーネルを使用し、2分割交差検定を行った。評価尺度として、精度・再現率を用いた。

#### 4.2. 実験結果と考察

本節では、4 節で述べたベースライン手法と提案手法の精度・再現率を表 6 に示す。提案手法 1 では精度が 0.508、再現率が 0.506 となりベースラインを上回る結果となった。また、提案手法 2 においては、ベースラインと比較し精度が 0.527 と全手法の中で最大となる結果を得ることが出来た。よって提案手法の有効性を示すことができたといえる。

表 6: 料理失敗ブログの自動判定結果

	評価尺度	
	精度	再現率
ベースライン	0.493	0.321
提案手法 1	0.508	<b>0.506</b>
提案手法 2	<b>0.527</b>	0.100

提案手法 2 では、精度に比べ、再現率が大きく下回る結果となった。ここでは、再現率低下の原因について考察を行う。再現率の低下の原因は、主に手掛かり語の不足によるものであった。本研究では、料理失敗ブログの判定を、3.3 節で示した手掛かり語を用いて機械学習により行った。使用した手掛かり語は、単語「失敗」に関連する用語を特許から収集した。しかし、使用した手掛かり語があまり出現しないという傾向があった。その理由として、失敗の原因は多種多様であること、また、失敗の状況を写真や顔文字などの方法で記述されていることが挙げられる。この場合、本研究で使用した、料理失敗関連の手掛かり語のみでは不十分であると考えられる。この問題を解決するためには、特許から「失敗」だけでなく「反省」や「改善」といった関連用語を抽出し、手掛かり語を追加することが考えられる。また、収集した手掛かり語の分析を行い、単語の重み付けを行うことも考えられる。

次に、人手では料理失敗ブログではないと判定したが、システムでは料理失敗ブログと判定された例を示す。図 3 に示すブログを人手で料理失敗ブログでないと判定した理由として、対象料理である「酢豚」に関する失敗事例が述べられていないことが挙げられる。記事の本文中には、対象料理である「酢豚」や「失敗」という単語が含まれている。しかし、実際に失敗したのは「酢豚」ではなく、「卵トマト炒め」という料理であることが分かる。これは、対象料理を調理し、その過程や完成品において失敗したという前述の定義には当てはまらない。今回は対象とする料理を事前に定めており、今後、料理の失敗事例を提示するうえで、図 3 のようなブログは必要性がないため、料理失敗ブログでないと判定した。この問題を解決するためには、対象となる料理名と手掛かり語の距離に着目することが考えられる。

<dish>酢豚</dish>

.....(略).....

12 日夜

居酒屋さん 酢豚で一杯

13 日昼

トマトが激安やったんで、大量に買って  
しまった(笑)

考えた末、卵トマト炒めにした

湯剥きしたら、失敗(笑) 水分が出ても一た(笑)  
へんに手間かけても一た ((T\_T))

図 3: 料理失敗ブログでないと判定したブログ例

## 5. まとめ

本研究では、料理初心者のために、料理レシピに記載されていないブログ上の料理経験者の失敗事例を料理レシピに対応付けることを目指し、その第一歩として、特許明細書から得られた手掛かり語を用い、失敗事例が述べられている料理失敗ブログかどうかを自動判定する手法を提案した。具体的な収集手順として、一般的なブログから「料理名 レシピ 反省点」を基に定めた5つを検索クエリとして、料理失敗候補ブログの収集を行い、得られた料理失敗候補ブログを対象に機械学習を用いて料理失敗ブログの自動判定を行った。

評価実験では、形態素解析で得られた全単語を素性として料理失敗ブログの判定をした場合をベースライン手法として提案手法との比較実験を行った。精度は提案手法1が0.508、提案手法2が0.527と共にベースライン手法を上回った。また再現率も提案手法1が0.506とベースラインを上回る結果となった。以上より、本研究で提案した手法の有効性を示すことが出来た。

今後の課題としては、料理レシピの自動作成により使いやすく信憑性の高い調理失敗情報提示システムの構築を目指したいと考えている。そのためにブログから多数の調理者が失敗しやすい、危険と判断した箇所の抽出を行い、失敗確率の高い情報をより効果的に提示していく必要がある。また、本研究では料理失敗候補ブログ1,046件を対象に実験を行ったが、機械学習を行うにはデータが不足している。そのため、データ数を増やし、料理失敗に関連する楽天レシピやYahoo!知恵袋、Twitter、ニュースなど、様々なコンテンツに対しても自動収集を行っていく。そして、より便利で信憑性の高い調理失敗情報提示システムの構築を行うことが考えられる。

## 参考文献

- [1] 志土地由香, 出口大輔, 高橋友和, 井手一郎, 中村裕一, 村瀬洋 “料理レシピをわかりやすくするための理解困難な表現の補足” 電子情報通信学会技術研究報告, MVE2009-145, pp. 95-100, (2010)
- [2] 橋明穂, 若宮翔子, 難波英嗣, 角谷和俊 “料理名の修飾表現の関係性に基づくレシピのネーミングコンセプト抽出” 電子情報通信学会技術研究報告, Vol.113, No.214, DE2013-36, pp. 19-24, (2013)
- [3] 山田勇, 兪明連, 志田晃一郎, 横山孝典 “手持ちの材料だけを使って調理できる料理レシピを Web から検索するシステム” 情報処理学会第70回全国大会, 3T-9, pp. 1-705-1-706, (2008)
- [4] 森信介, 山肩洋子, 笹田鉄郎, 前田浩邦 “レシピテキストのためのフローグラフの定義” 情報処理学会自然言語処理研究報告, Vol.2013-NL-214, No.13, pp. 1-7, (2013)
- [5] Hamon, T. and Grabar, N. “Extraction of Ingredient Names from Recipes by Combining Linguistic Annotations and CRF Selection” Proceedings of the 5th International Workshop on Multimedia for Cooking & Eating Activities, pp. 63-68, (2013)
- [6] 難波英嗣, 土居洋子, 辻田美穂, 竹澤寿幸, 角谷和俊 “複数料理レシピの自動要約” 電子情報通信学会技術研究報告, Vol.113, No.338, NLC2013-41, pp. 39-44, (2013)
- [7] Nanba, H., Mayumi, S., and Takezawa, T. “Automatic Construction of a Bilingual Thesaurus using Citation Analysis” In Proceedings of the 4th International CIKM Workshop on Patent Information Retrieval (PaIR'11), pp. 25-30, (2011)