料理技術習得のためのナイーブベイズ分類器を用いた未習得要素を含む レシピ推薦

新宇 北山 大輔

† 工学院大学大学院工学研究科情報学専攻 〒 163-8677 東京都新宿区西新宿 1-24-2 E-mail: †em21029@ns.kogakuin.ac.jp, ††kitayama@cc.kogakuin.ac.jp

あらまし 料理や学習など、以前に難しくて失敗したことでも経験を積んだら成功するようなものがある. そこで本研 究では、料理に焦点を当て、ユーザの料理に対する失敗や成功のログを使って、失敗を成功に変えるために経験する べきレシピを推薦する手法を提案する.具体的には、まず、ユーザが過去に料理したレシピとそれに対する失敗や成 功のラベルを用意する.次に、レシピの材料と調理動作を素性とした、ナイーブベイズ分類器を使って、失敗および 成功となる原因を見つける. 成功要因を多数含むが、失敗要因も含むレシピを推薦することで、失敗要因の練習の心理 的ハードルを下げたレシピの推薦を実現する.

キーワード 料理レシピ,未習得要素抽出,機械学習,ナイーブベイズ,食材,調理動作

1 はじめに

近年、料理レシピの検索において、クックパッド 1や楽天レ シピ²などのレシピ投稿サイトが利用されている. このような サイトはレシピの閲覧だけでなく、ユーザ自身が作成した料理 レシピを投稿することができる. それに伴い、投稿サイトには 数多くの料理レシピが存在している. ユーザが好みのレシピを 選ぶことは可能であるが、しかし、材料や手順の得手不得手な ど様々な理由で、実際の調理では思い通りに調理できないこと がある. そこで、本研究ではこれらの問題を解決するために、 ユーザの調理失敗原因を克服するための練習ができるレシピ推 薦手法を提案する.

具体的には、ユーザの失敗と成功の記録を利用し、そこから 必要な未習得要素を抽出し、その未習得要素を含みかつ成功率 が高いレシピを推薦する. 例えば、ユーザはハンバーグが得意 であり唐揚げが苦手だとする、システムは得意もしくは苦手な レシピより、得手と不得手の要素を抽出する. その結果、得意 な材料が「玉ねぎ」で、不得意な調理動作が「揚げる」であっ た場合、これらの要素に基づいて、「天ぷら」を不得意要素を練 習できるレシピとして推薦する.

本論文の構成は以下のとおりである. 2章では関連研究につ いて述べる.3章では提案手法について述べる.4章では実行 例とその考察について述べる. 最後に5章で本論文のまとめと 今後の課題について述べる.

2 関連研究

2.1 一般的なレシピ推薦

これまで、一般的なレシピ推薦関連の研究は数多くされてい

る. 上田ら[1] はレシピへの得点付けの動作を確認するととも に、評価実験用システムを用いて適合率によって提案手法の有 効性を示す. 高畑ら[2] は料理レシピを推薦する際にはユーザ の好き嫌いといった嗜好を考慮することが重要だと考え, 食材 に対する好き嫌いを考慮した料理レシピ推薦システムの提案と 実装に向けた検討を行い報告した.

橘ら[3] は料理の典型的な手順や材料などの要素とあるレシ ピの要素との差異に着目し、料理名のネーミングコンセプトを 抽出し、それを活用したレシピ推薦システムを提案している. このとき、典型的な要素との差異には料理の特徴であるものと そうでないものが混在しているため、料理名に同一の修飾表現 が付与されているレシピ集合において、典型的な要素との差異 が類似するレシピから共通的な要素をネーミングコンセプトと して抽出した.

一般にレシピ推薦は、ユーザの食の好みに基づいて行われる. 本研究では、食べたいレシピではなく、能力向上のためのレシ ピ推薦に取り組む.

2.2 初心者向けのレシピ検索/生成

本研究では、調理の練習ができるレシピを扱う. 初心者に示 す練習用のレシピとして、標準的なレシピを抽出するアプロー チも有り得るため、それらを紹介する.

難波ら[4]は、複数テキスト要約の技術を用いて、レシピサ イトに投稿された特定の料理に関する複数のレシピから、典型 的な材料と調理手順を出力する手法を提案する. また、その過 程で複数レシピ要約の生成に必要な料理オントロジーの構築を 行った. さらに、要約を行う際に考慮すべき点として、材料や 手順の表記揺れ、手順の省略、多対多の同定の問題といった点 を挙げている. それぞれに関して、レシピ間の類似個所の検出 には、同義語辞書、格フレーム辞書、述語項構造解ツールなど、 様々な言語資源が必要とされる. そこで、統計的言語処理技術 を用いて、レシピデータからこれらのリソースを構築し、複数

¹: https://cookpad.com/

²: https://recipe.rakuten.co.jp/

瀧本ら[5]は、レシピテキスト間の手順から異なる2つの操 作に対して, 出現位置を考慮したペアアライメントを行うこと で操作を対応付け、施設配置問題の考えに基づいて基本操作の 選択を行い、アライメントの結果を利用して操作の順序付けを 行う、という手順で基本手順の抽出を行っている.

Mori ら [6] は、料理レシピを用いて手順を表示したフローグ ラフから手順を示す文を自動生成する手法を提案した. この研 究の目的は、学習コーパスからテンプレートを作成し、入力と してフローグラフを用いて統計的な文を生成することである. 料理レシピ中の固有表現である「食材」や「道具」に対してタ グ付けを行い, 抽出した単語を接点として扱うことで, フロー グラフを作成した.

志土地ら[7] の研究では初心者には理解困難な料理レシピに わかりやすい料理レシピを作成して提供する、そして料理レシ ピ中の初心者にとって分かりにくい箇所を検出する手法を提案 した. 横井ら [8] は食材の組合せを考慮した食材の追加, 削除 の推薦と、ユーザが意図した料理レシピを検索する手法につい て検討するため、食材の組合せの典型度を分析する手法を提案 した. 秋口ら [9] は料理レシピ動画の調理動作の並列性, 動作 の頻度、カメラワークによる編集意図を要素とし抽出すること で、レシピ動画の時間特性に基づいた難易度を判定する手法を 提案した. 呉ら[10] は投稿型レシピサイトに掲載されたレシピ を利用して、調理を行うユーザを支援することを目的として、 料理の手順を表すインフォグラフィックを自動的に生成する手 法を提案した. 古本ら [11] はあるレシピから料理アドバイスを 抽出し、類似するレシピに補完することでより初心者に易しい レシピを提示する手法を提案した.

本研究では、個人の不得意を扱っているが、最終的に提示す るレシピとして、このような標準レシピとの近さを加味するな ど、補完的に利用できると考えている.

2.3 ユーザの能力向上を目的とした推薦

本研究では、ユーザの能力向上を目的としている. 以下に代 表的な研究を紹介し、本研究の位置づけを示す.

蒲原ら[12] はある料理に対して共通して使用される食材に着 目して、その食材群との差分を用いたレパートリを増やすため の料理レシピ検索手法を提案した. 長屋ら [13] はそれぞれのレ シピから好みの調理工程を選択し、調理者自身のアレンシレシ ピを作成できるシステムを提案した. これらの研究は、ユーザ が未知のレシピに挑戦することで能力向上をはかることができ る. 一方で我々は既知の不得意要素に挑戦させることで能力向 上をはかる.

大杉ら [14] は失敗しやすい工程を明示する方法に関する基礎 検討した,料理の失敗に関する定義や原因の分類と半構造的イ ンタビューを利用した. 本研究とは失敗しやすい要素に着目し ている点で共通している. 大杉らの一般的な失敗要因の知見を, 例えば、素性の重みとして使うなどすることが考えられる.



材料(8人分)	
プレーンヨーグルト	4 0 0 g
ホットケーキミックス	1 0 0 g
お好みのジャム 又は砂糖	50~80g
99	2個
サラダ油	大さじ3

図1 料理レシピの構成



図 2 料理レシピの手順

3 提案手法

3.1 料理レシピの構成要素

本研究では、ユーザの得意/不得意の要素をレシピの材料と 調理動作の単位で扱う.一般に、料理レシピは図1と図2のよ うにタイトル,材料,調理手順からなる.

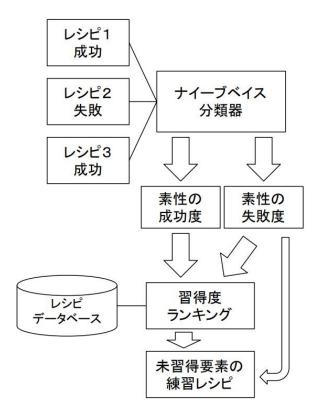


図 3 提案手法の処理の流れ

本稿では、材料として材料名の文字列を用い、調理動作として調理手順中の動詞を用いる。これらを素性とし、レシピ中に含むか含まないかの 1 か 0 として扱ったベクトルで表現する。

素性の抽出方法について述べる。まずは「しょう油」と「しょうゆ」という同じ意味の語であるが、表記が異なるので別の単語となる問題を防止するため、材料については漢字とひらがなを全部カタカナに変換したものを抽出する。調理動作の抽出については大仁田ら [15] の研究で採用された前処理方法を参考する。レシピの手順に対し MeCab [16] により形態素解析を行い、その後、オントロジを用いて各動作を上位の単語に置き換える処理を行なっている。処理の例として、例えば「加える」という動作が存在した際には「入れる」という変換がなされる。そして、オントロジの辞書に登録されてない動詞は「非調理動作」として除外する。本研究のオントロジは、料理オントロジー構築プロジェクトで公開されている土居ら [17] の研究を基にした料理オントロジーを用いている.

3.2 概 要

図3に提案手法の処理の流れを示す.ユーザは、過去の調理に基づき、それぞれのレシピに対し、上手くできた度合を付与しているものとする.付与された度合の高さで、成功、失敗のラベルを与える.次に、過去の調理レシピのベクトルとその成功/失敗のラベルからナイーブベイズ分類器により、失敗と成功の要因を抽出する.最後に、候補レシピの習得度のスコアを算出する.スコアが高く、かつ失敗要因を含むレシピを推薦結果として表示する.

3.3 ナイーブベイズ分類器と習得度の算出

抽出した材料と調理手順については、本実験ではナイーブベイズを用いて分析している。ナイーブベイズとは、とあるデータがどのカテゴリーに属するものなのか判定する、機械学習の手法のひとつである。本実験では特徴量の材料と調理手順と被験者の得意度との関係を比較するために、ナイーブベイズ分類器の MultinomialNB³を採用した。これは特徴ベクトルに多項分布を仮定する場合に使われる。

未習得要素について、本研究ではユーザの料理履歴から抽出された失敗の原因と定義している。本研究の目的はそれを成功に変換することである。そこで、ベクトルの各素性と成功ラベルの関係をナイーブベイズ分類器で分析した結果を成功度とし、ベクトルの各素性と失敗ラベルの関係を分析した結果を失敗度とする。成功度と失敗度の差分を習得度として定義する。習得度が高いほど、その素性が習得済みであることを示す。

3.4 未習得要素を含むレシピ推薦

本研究のレシピの推薦手法について述べる。まず、個々の材料と調理動作に付与された習得度の総和を候補レシピのスコアとし算出する.

$$score(r) = \sum_{e \in E_r} acq(e)$$
 (1)

式 1 中の E_r はレシピr の材料と調理動作の集合である, acq は習得度を返す関数である.

スコアは習得度の総和であるため、この値が高いほど、成功 しやすいと考えられる。成功しやすいレシピの方が未習得要素 を練習しやすいと考える。そのため、ユーザの履歴中の習得度 の下位 10 件の素性を未習得要素とし、未習得要素を含みかつ スコアの高いレシピを推薦する。

4 レシピ推薦の例

4.1 使用データ

本章では、手法の動作例について紹介する。楽天レシピ [18] に含まれるレシピのうち約 400000 件を抽出し、対象とした.

表 1 に入力として用いた 10 種類のレシピの材料と調理動作および成功ラベルを示す。ラベルは 0 と 1 で定義され,0 は失敗,1 は成功であることを意味している。

4.2 素性の習得度の算出

抽出した材料と調理手順を運用して、3.3章で説明したナイーブベイズ分類器の MultinomialNB を利用して 10種類の料理の素性の習得度を算出した。表2の上の部分は習得度上位の素性である。下の部分は習得度下位の素性である。それぞれ10件づつ示している。

今回の入力では得意とする素性は引く, まぶす, 鶏もも肉, なじむ, ゆすぶる, 青ネギ, エリンギ, お好みの具材, おろし

 $^{3:} https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html\\$

表 1 10 種類のレシピ

レシピタイトル	材料	調理動作	
元○長がこっそり教える びっくり○ンキーの ハンバーグ	牛豚合びき肉, 豚ひき肉, 玉ねぎ, パン粉, 卵,塩, 胡椒, 味噌, ナツメグ, 果物の果汁, コーヒーフレッシュ, 水, 醤油, 料理酒, みりん	入れる、混ぜる、混ざる、ねかす、茹でる、する、揉む、さらす、 切る、取る、調える、取り分ける、のばす、作る、とろける、乗せる 押さえる、折る、はみ出す、くっつける、延ばす、熱す、裏返す、 ひっくり返す、焼く、取りかかる、炒める、加える、煮る	
お家にあるもので 【らーめん屋さんの チャーハン】	卵, 青ネギ, ごはん, しょうゆ, だしの素, おろし生姜, 塩, コショウ	する, 入れる, 熱す, 使う, なる, 混ぜる, 炒める, 返す, 焦がす, 混ぜ合う, 取る, 飛ばす, おろす, 出す, 馴染む, ある, かぶせる, 持つ, くるう, 外す, できあがる	
甜麺醤なしで出来る! 簡単 ホイコーロー	豚肉, キャベツ, ピーマン, ねぎ, にんにく, 味噌, しょうが, 砂糖, しょうゆ, 酒, ごま油, 豆板醤	切る, する, 合わせる, 熱す, 入れる, 炒める, 出す, 出る, 焼ける, 加える, 出来上がる	0
卵炒め	卵, 牛乳, トマト, ブロッコリー, 塩コショウ, ごま	むく、食べる、洗う、とる、いれる、溶く、ひく、流す、 炒る、つくる、だす、いためる、とりだす、あわせる、 しおる、調える、ふりかける、もつ、かける	
激うま♪ 豚肉と茄子の味噌炒め	茄子, 豚肉, にんにく, 味噌, みりん, 砂糖, サラダ油	切る, 溶く, 入れる, 熱す, 立つ, 炒める, 変わる, 回る, する	
主人が、いくらでも食べれる と絶賛のナス・ピーマン	ナス, ピーマン, 砂糖, 醤油, ゴマ油, だしの素, 白いりゴマ	取る, 切る, する, さらす, ひく, きる, 入れる, 炒める, 加える, 出来上がる, 振る	
我が家の人気者!! 鶏の唐揚げ	鶏もも肉, しょう油, 塩, 酒, 砂糖, にんにくのすりおろし, こしょう, 片栗粉	切る、入れる、つけ込む、あげる、まぶす、揚げる、出来上がる	
ヨーグルトと HM で 超簡単濃厚チーズケーキ	プレーンヨーグルト, お好みのジャム, 卵, ホットケーキミックス, サラダ油	する, 塗る, ふるう, 敷く, 入れる, 混ぜる, 流し込む, 焼く, もつ, 上がる	
♪♪吉●家より美味しい? ツユだく吉野家風牛丼♪♪	牛バラ肉, 玉ねぎ, だしの素, 水, しょうが, しょうゆ, 酒, 砂糖	する, 入れる, 煮る, おろす, 見る, 足りる, 加える, のせる, 出来上がる, かける, なる	
塩麹で野菜炒め	鶏もも肉, キャベツ, にんじん, エリンギ, ピーマン, ちくわ, しょうが, 塩コショウ, 塩麹	切る, 食べる, 熱す, 引く, 炒める, 付く, いれる, する, 入れる	1

生姜、片栗粉. このうち、苦手な素性は玉ねぎ、焼く、水、煮る、うらがえす、牛バラ肉、牛豚合挽き肉、お好みのジャム、辛子、果物の果汁、であることが分かった. 玉ねぎや煮るに関しては、2つのレシピに出現し、かつどちらも失敗レシピであるため、うまく抽出できているといえる. しかしながら、お好みの具材やお好みのジャムなどはあるレシピにたまたま出現しているだけであり. ユーザが真に苦手としているかにはわからない. そのため、一般的な出現しやすさや、履歴中での出現数によって、確度のようなものを算出する必要があると考える.

4.3 レシピの推薦結果

まず表 2 の習得度下位 10 件のどれかを含むレシピをデータベースから抽出した.次に、それらのレシピのスコアを算出した. 最後にこれらのレシピをスコアでランキングした.

上位 5 件のレシピを表 3 に示す、「玉ねぎ」を練習するレシピとして、「手軽にできるカニコロ」が推薦されている。このレシピでは、「鶏モモ肉」といった履歴中の成功レシピに出現する食材や、「まぶす」と「ゆすぶる」という習得度の高い動作が出現する。このレシピの手順を確認したところ、「玉ねぎ」は、具材として使われており、練習レシピとして適切であると考える。上位 5 件のレシピの中で 4 件の含まれる苦手要素は「焼く」である。「鶏肉のソテー」については、「焼く」は主要な動作であるため、練習用のレシピとして適切であると考える。しかしながら、他のレシピでは、「焼く」は一部の食材にのみ使う動作であるため、練習用としては不適切なレシピと考えられる。そのた

表 2 得意と苦手上位の素性とそれの習得度

得意/苦手	素性名	成功度	失敗度	習得度
14127/111				
得意	引く	-3.998	-5.327	1.329
	まぶす	-4.285	-5.327	1.042
	ニワトリモモニク	-4.285	-5.327	1.042
	なじむ	-4.691	-5.327	0.636
	ゆすぶる	-4.691	-5.327	0.636
	アオネギ	-4.691	-5.327	0.636
	エリンギ	-4.691	-5.327	0.636
	オコノミノグザイ	-4.691	-5.327	0.636
	オロシショウガ	-4.691	-5.327	0.636
	カタクリコ	-4.691	-5.327	0.636
苦手	タマネギ	-5.384	-3.718	-1.666
	焼く	-5.384	-3.941	-1.443
	ミズ	-5.384	-4.229	-1.155
	煮る	-5.384	-4.229	-1.155
	うらがえす	-5.384	-4.635	-0.749
	ウシバラニク	-5.384	-4.635	-0.749
	ウシブタゴウビキニク	-5.384	-4.635	-0.749
	オコノミノジャム	-5.384	-4.635	-0.749
	カラシ	-5.384	-4.635	-0.749
	クダモノノカジュウ	-5.384	-4.635	-0.749

め、主要な食材や動作であるかを判断することで、結果を改善 できると考える.

表 3 推薦された習得度上位のレシピ

習得度	含まれる苦手要素	
6 750	焼く	
0.759		
6.597	玉ねぎ	
6.528	焼く	
6.528	焼く	
6.528	焼く	
	6.759 6.597 6.528 6.528	

5 ま と め

本研究では、料理における失敗箇所を練習できるレシピ推薦を提案した.具体的には、まず、ユーザが過去に料理したレシピとそれに対する失敗や成功のラベルを用意する.次に、レシピの材料と調理動作を素性とした、ナイーブベイズ分類器を使って、失敗および成功となる原因を見つけ、基本的な特徴量を抽出する.最後に成功要因を多数含むが、失敗要因も含む習得度のスコアが高いレシピを推薦する.本稿では、提案手法を動作させ、適切な推薦結果を得ることができることを確認した.

今後の課題についてまとめる.まず,比較手法を用いた評価 実験を行う予定である.比較手法としては,不得意要素を含む 量によるランキングを考えている.

そして、本手法は成功と失敗に対する材料の量の影響を考慮 していない.ナイーブベイズでは、量のような連続値を扱えな いため、扱えるように拡張する必要がある.

本稿の調理動作は単純に調理手順から抽出した単語レベルである.一般に、同じ調理動作でも異なる材料では、その難しさが変わると考える. 今後は調理動作と材料の組み合わを素性として扱う必要があると考える.

失敗と成功の要因抽出の結果について,実行例ではその成功率と失敗率が最も高いいくつかの特徴量を選択した.今後の課題では,失敗と成功の要因抽出を改善し,より正確な抽出とその結果の正確性の検証が必要だと考える.

最後に、不得意を含むレシピの推薦方法について、レシピ中の最低習得度の値の正負を反転して、合計スコアにかけるとういう方法も検討している。また、以上の研究の内容を用いて、実験を行い、システムの有効性を評価する。

謝辞

本研究の一部は、2021 年度科研費基盤研究 (C)(課題番号: 21K12147) によるものです。ここに記して謝意を表すものとします。本研究では、国立情報学研究所の IDR データセット提供サービスにより楽天グループ株式会社から提供を受けた「楽天データセット」(https://rit.rakuten.com/data_release/) を利用した。

文 献

- [1] 上田真由美, 石原和幸, 平野靖, 梶田将司, 間瀬健二. 食材利用履歴 に基づき個人の嗜好を反映するレシピ推薦手法. DBSJ Letters, Vol. 6, No. 4, pp. 29–32, 2008.
- [2] 高畑麻理, 上田真由美, 中島伸介. 食材に対する好き嫌いを考慮

- した料理レシビ推薦手法の提案. データ工学と情報マネジメント に関するフォーラム, E3-5, Feb, 2011.
- [3] 橘明穂, 若宮翔子, 難波英嗣, 角谷和俊. 料理名のネーミングコンセプト抽出によるレシピ推薦システム. ARG Web インテリジェンスとインタラクション研究会, No. 2, pp. 1–2, 2013.
- [4] 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, pp. 39-44, 2013.
- [5] 瀧本洋喜, 笹野遼平, 高村大也, 奥村学. 施設配置問題に基づく同一料理のレシピ集合からの基本手順の抽出. 言語処理学会第 21 回年次大会発表論文集, pp. 1092–1095, 2015.
- [6] Shinsuke Mori, Hirokuni Maeta, Yoko Yamakata, and Tetsuro Sasada. Flow graph corpus from recipe texts. In *LREC*, pp. 2370–2377, 2014.
- [7] 志土地由香,出口大輔,高橋友和,井手一郎,中村裕一,村瀬 洋.料理レシピ中の初心者に理解困難な表現の抽出.信学技報, MVE2009-70, Nov, 2009.
- [8] 横井聡,道満恵介,平山高嗣,井手一郎,出口大輔,村瀬洋.料理レシピにおける食材の組合せの典型度分析.電子情報通信学会技術研究報告 IMQ,イメージ・メディア・クオリティ, Vol. 114, No. 485, pp. 49–54, 2015.
- [9] 秋口いくみ, 王元元, 河合由起子, 角谷和俊. 料理レシピ動画の時間特性抽出による難易度判定. 第 10 回データ工学と情報マネジメントに関するフォーラム (DEIM2018), 2018.
- [10] 呉子, 牛尼剛聡. 料理レシピの理解支援を目的としたインフォグラフィックの自動生成. 第 13 回データ工学と情報マネジメントに関するフォーラム (DEIM2021), 2021.
- [11] 古本健太, 難波英嗣, 角谷和俊. 料理アドバイスを補完したレシピ提示方式. 第 9 回データ工学と情報マネジメントに関するフォーラム (DEIM2017), 2017.
- [12] 蒲原智也,楠和馬,崔赫仁,木下泰輝,塩井隆,波多野賢治ほか. 基準レシピからの差分に基づいた料理レシピ推薦によるレパートリ拡大法の提案.研究報告データベースシステム (DBS), Vol. 2017, No. 22, pp. 1–5, 2017.
- [13] 長屋健, 佐藤哲司. 既存レシピの調理手順を用いたアレンジレシピ作成システムの提案. 第 10 回データ工学と情報マネジメントに関するフォーラム (DEIM2018), 2018.
- [14] 大杉隆文, 永井廉人, 松下光範ほか. 料理レシピ中の失敗しやす い工程の提示方法に関する基礎検討. 研究報告データベースシス テム (DBS), Vol. 2017, No. 20, pp. 1–6, 2017.
- [15] 大仁田龍也, 北山大輔. 手順のレシピ頻度特徴量と lexrank に基づくアレンジ抽出手法. 研究報告データベースシステム (DBS), Vol. 2020, No. 7, pp. 1–6, 2020.
- [16] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. Applying conditional random fields to Japanese morphological analysis. In Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, pp. 230–237, Barcelona, Spain, July 2004. Association for Computational Linguistics.
- [17] 土居洋子, 辻田美穂, 難波英嗣, 竹澤寿幸, 角谷和俊. 料理レシピと特許データベースからの料理オントロジーの構築. 信学技報, MVE2013-68, March, 2014.
- [18] 楽天グループ株式会社. 楽天レシピデータ. 国立情報 学研究所情報学研究データリポジトリ. (データセット). https://doi.org/10.32130/idr.2.4, 2021.