

# 単語のベクトル表現を用いたレシピ推薦システム

## Recipe recommendation system using word vector representation

矢野 達也, 林 豊洋, 大橋 健

Tatsuya YANO, Toyohiro HAYASHI, Takeshi OHASI

九州工業大学

Kyushu Institute of Technology

### Abstract

本研究では、レシピ単語のベクトル表現を用いた柔軟なレシピの推薦を目的とした。Word2Vec を用いてレシピ単語のベクトル表現を学習した。レシピ単語のベクトル演算を用いることで従来のキーワード検索型のレシピ推薦よりも柔軟なレシピ推薦が可能となった。

## 1 はじめに

近年、スマートフォンの普及に伴い、どこでも気軽にレシピの投稿、閲覧が可能であるユーザー投稿型のレシピサイトが幅広く利用されている。ユーザー投稿型レシピサイトではユーザーが投稿した数百万を超えるレシピが掲載されており、調理する際、豊富なレシピデータからユーザーの所望するレシピを効率よく推薦できるシステムが求められる。現在のレシピ推薦手法としてはキーワード検索が主に用いられている。キーワード検索とは、ユーザーが入力したキーワードがレシピ内に含まれるレシピを推薦する手法である。料理名やカテゴリ、材料などをキーワードとして入力し、マッチングしたレシピから人気度順でユーザーに推薦する。しかし、キーワード検索では、レシピ推薦にユーザーの嗜好を考慮しづらい点や、たとえユーザーが好むレシピであってもキーワードが含まれていないレシピは推薦されない。これらの点からキーワード検索によるレシピ推薦では、豊富なレシピデータからユーザーの所望するレシピを推薦するのは困難である。

本研究では、従来のキーワード検索より柔軟なレシピ推薦を可能とする手法として、レシピ単語のベクトル表現を用いたレシピ推薦手法を提案する。レシピ掲載サイトから取得したレシピ情報を元に文章を作成し、Word2Vec 用いてレシピ単語のベクトル表現を学習する。学習したレシピ単語のベクトル表現でベクトル演算を用いてレシピの推薦を行い、その評価をする。

## 2 Word2Vec

Word2Vec とは Tomas Mikolov らによって提唱された、単語のベクトル表現の獲得手法である [1]。学習コーパスとして文章を与えることで、文章中の単語の共起関係に基づき、予め設定していた次元数のベクトル表現 (分散表現) を学習する。この単語のベクトル表現により、「King」-「Man」+「Woman」=「Queen」のような単語間での演算が可能となる。通常、学習コーパスには数十万から数百万の語彙数を持つ比較的大規模なテキストデータを使用する。文章データに含まれる語彙数が多いほど幅広い表

現可能なベクトル空間を学習できる反面、特定の分野に関しては、疎なベクトル空間になる場合が考えられる。本研究では、レシピ推薦に適した密なベクトル空間を学習するため、Word2Vec の学習コーパスは、レシピサイトから収集した情報を元に作成する。

## 3 レシピ単語ベクトル空間の構築

### 3.1 レシピ情報の収集

レシピ情報の収集には、レシピ掲載サイト楽天レシピ\*1から収集する。楽天レシピは現在約 160 万件のレシピを掲載しているユーザー投稿型レシピサイトの 1 つである。楽天レシピに掲載されている 1397 の料理カテゴリから各カテゴリの人気レシピ上位 3 件を取得し、重複を除いた 3070 品のレシピを収集した。1 つのレシピに対し、材料情報、カテゴリ情報、特徴情報を収集する。楽天レシピの掲載レシピ「豆腐ハンバーグ」を例に、材料情報、カテゴリ情報、特徴情報の取得方法を順に 3.1.2 節、3.1.3 節、3.1.4 節に示す。またレシピ情報の収集の際に行った、表記のゆれ対策について 3.1.1 節に示す。

#### 3.1.1 表記のゆれ

表記のゆれとは同義の単語に対し異なる文字表記が付されることである。特にユーザー投稿型のレシピサイトの材料表記には、投稿の書式が統一されていないため表記のゆれが多く存在している。楽天レシピにおける主な材料表記のゆれのパターン例を Figure1 に示す。

Figure1.材料表記のゆれのパターン例

表記パターン	ゆれの内容
[こしょう],[胡椒]	書式のゆれ
[こしょう],[*こしょう]	記号文字によるゆれ
[こしょう],[塩・こしょう]	付属表現によるゆれ (内包)
[バター],[バター (サラダ油でも可)]	付属表現によるゆれ (代用)

同じ材料を使用するレシピは類似するレシピとして学習する必要がある。Word2Vec の学習では同義の単語であっても表記のゆれによって別の単語として学習してしまう。そのため、学習前に表記のゆれは正規化する必要がある。本研究では、レシピの材料情報を取得する際、表記のゆれの対策として以下の手順をとった。

\*1 楽天レシピ 「<https://recipe.rakuten.co.jp/>」

1. 不要な記号文字の除去
2. 表記単語を分かち書きし、名詞単語のみ取得
3. 取得した単語をカタカナに変換

分かち書きは、文章を単語単位に分解する手法である。分かち書きにはオープンソースの形態素解析エンジン MeCab<sup>2</sup> を使用した。材料に対して分かち書きを施すことで「塩・こしょう」のような複合表記された材料に対しても、「塩」「こしょう」と個別の材料として扱うことができる他、「木綿豆腐」、「絹ごし豆腐」などの材料も共通する「豆腐」の材料を取り出すことができる。分かち書きで得た名詞単語はカタカナに変換する。書式を統一することで「こしょう」、「胡椒」のようなゆれに対応する。

### 3.1.2 材料情報の取得

レシピページに記載されている材料から材料情報を取得する。「豆腐ハンバーグ」に記載されている材料一覧を下に示す。

材料【木綿豆腐、豚ひき肉、玉ねぎ（正味）、塩・こしょう、サラダ油、…】

始めに材料一覧から調味料のみを取り出し、食材と調味料に分別する。調味料の判別には事前に用意した調味料辞書を使用する。ここで、食材と調味料を分けた理由は、材料文章の作成時に食材単語と調味料単語を区別する為である。最後に、表記のゆれに対応するため、3.1.1 節で述べた処理を行う。処理後得た結果は以下の様になる。

#### 材料情報

食材【モメン、トウフ、ブタ、ヒキニク、…】  
調味料【シオ、コショウ、サラダユ、…】

取得した材料、調味料単語を「豆腐ハンバーグ」の材料情報とする。

### 3.1.3 カテゴリ情報の取得

カテゴリ情報とは、ユーザーがレシピを投稿する際に登録するレシピのカテゴリである。それぞれのレシピには肉、魚、豆腐などの主要材料や和食、中華料理、スイーツなどのジャンルや様式がカテゴリとして登録されている。このカテゴリをレシピのカテゴリ情報として取得する。「豆腐ハンバーグ」のレシピに登録されているカテゴリを以下に示す。

カテゴリ【肉、挽き肉】

またレシピ名にも「本格中華の…」、「お家で簡単イタリアン…」の様に、重要なカテゴリが含まれていることが多い。そのためレシピ名を分かち書きした各単語もカテゴリ情報として使用する。「豆腐ハンバーグ」のレシピ名を分かち書きした結果を以下に示す。

レシピ名の分かちがき【豆腐、ハンバーグ】

書式の統一のためカテゴリとレシピ名を分かち書きした単語をカタカナに変換し、「豆腐ハンバーグ」のカテゴリ情報とする。

#### 豆腐ハンバーグカテゴリ情報

【ニク、ヒキニク、トウフ、ハンバーグ】

### 3.1.4 特徴情報の取得

料理の特徴情報とは、料理がもつ、「しゃきしゃき」や「しっとり」などの食感覚、「甘い」や「辛い」などの味覚、「ヘルシー」や「がっつり」などの性質を指す。この特徴情報を収集には楽天レシピのつくったよレポート（以下つくレポ）を利用する。つくレポとは、投稿レシピを参考に料理を作ったユーザーがその感想を投稿したものである。「豆腐ハンバーグ」には119件のつくレポが投稿されていた。つくレポから料理の特徴単語を抜き出し、料理の特徴情報とする。特徴単語の抜き出しには事前に用意した特徴単語辞書を使用する。特徴単語辞書は食感覚辞書、味覚辞書、性質辞書の3つを用意した。食感覚辞書には早川らの論文「日本語テクスチャー用語の収集」[2]付表されているテクスチャー用語一覧から445単語を登録した。味覚辞書はよく使用される一般的な味覚表現76単語を登録した。性質辞書は、つくレポ内でよく用いられる性質表現15単語を登録した。食感覚辞書、味覚辞書、性質辞書の一部を Figure 2 に示す。

Figure 2 特徴単語辞書の構成

食感覚辞書	味覚辞書	性質辞書
脂っこい	甘い	がっつり
ジューシー	辛い	たっぷり
カリカリ	甘辛い	あつあつ
しゃきしゃき	渋い	ヘルシー
しっとり	酸っぱい	ビター

特徴単語辞書を用いた特徴情報の取得手順を以下に示す。

1. レシピに投稿されているつくレポを上限160件まで収集し一つのテキストにまとめる
2. 1で得たテキストから特徴単語辞書を用いて特徴単語を取得する
3. 取得した特徴単語をカタカナに変換し、その取得数と共に記録する

1~3の手順で得た特徴情報を Figure 3 に示す。これを「豆腐ハンバーグ」の特徴単語とする。

Figure 3. 豆腐ハンバーグの特徴情報取得結果

特徴単語	取得数
ヘルシー	17
フワフワ	8
フンワリ	3
フックラ	3
アッサリ	3
サッパリ	2
アマイ	1

<sup>2</sup> MeCab 「<http://taku910.github.io/mecab/>」

### 3.2 学習コーパスの作成

収集した各レシピの材料情報、カテゴリ情報、特徴情報と文章作成用の4つのテンプレートを用いて学習コーパスを作成する。文章作成用のテンプレートを以下に示す。

- ・材料文テンプレート  
【レシピ名】 食材【食材 A】【食材 B】【食材 C】  
【レシピ名】 調味料【調味料 A】【調味料 B】【調味料 C】
- ・カテゴリ文テンプレート  
【レシピ名】 カテゴリ【カテゴリ A】【カテゴリ B】【カテゴリ C】
- ・特徴文テンプレート  
【レシピ名】 特徴【特徴 A】【特徴 B】【特徴 C】

テンプレートは共通して3つのレシピ情報単語を並べる構成にしている。これは、単語の共起関係に基づく Word2Vec の学習方法を考慮して、様々な単語の配置パターンを学習させるためである。

#### 3.2.1 材料文章の作成

3.1.2 節の手順で取得した材料情報を材料文テンプレートに当てはめて材料文章を作成する。材料文は、調味料と調味料以外の食材を分けて作成するため、2種類のテンプレートを用いる。食材文は【レシピ名】にレシピ名を、【食材 A】【食材 B】【食材 C】に材料情報から材料単語を3つランダムに取り出し当てはめる。調味料文も同様に作成する。食材文を30文、調味料文を20文作成し合わせて50文を材料文章とする。ここで、食材文を調味料文より多く作成するのは、材料の中で食材が調味料より料理への影響が高いという考えのもと、食材の学習に重みをおくためである。「豆腐ハンバーグ」の材料文作成例を以下に示す

- ・食材文  
豆腐ハンバーグ 食材 トウフ ヒキニク ブタ  
豆腐ハンバーグ 食材 モメン ブタ タマネギ  
⋮

- ・調味料文  
豆腐ハンバーグ 調味料 コシヨウ シオ サラダユ  
豆腐ハンバーグ 調味料 シオ コシヨウ サラダユ  
⋮

#### 3.2.2 カテゴリ文章の作成

3.1.3 節の手順で取得したカテゴリ情報をカテゴリ文テンプレートに当てはめてカテゴリ文章を作成する。カテゴリ文は【レシピ名】にレシピ名を、【カテゴリ A】【カテゴリ B】【カテゴリ C】にカテゴリ情報からカテゴリ単語を3つランダムに取り出し当てはめる。カテゴリ文は50文作成する。「豆腐ハンバーグ」のカテゴリ文作成例を以下に示す。

- 豆腐ハンバーグ カテゴリ ヒキニク トウフ ニク
- 豆腐ハンバーグ カテゴリ ハンバーグ ヒキニク トウフ
- 豆腐ハンバーグ カテゴリ ヒキニク ハンバーグ ニク  
⋮

#### 3.2.3 特徴文章の作成

3.1.4 節の手順で取得した特徴情報を特徴文テンプレートに当てはめて特徴文章を作成する。特徴文は【レシピ名】にレシピ名を、【特徴 A】【特徴 B】【特徴 C】の順に特徴情報から特徴単語の取得数を重みとして順に3つランダムに取り出して当てはめる。特徴単語の取得数を重みとすることで、取得数の多い特徴単語を多く学習できるようにしている。特徴文は50文作成する。「豆腐ハンバーグ」の特徴文作成例を以下に示す。

- 豆腐ハンバーグ 特徴 ヘルシー フワフワ フックラ
- 豆腐ハンバーグ 特徴 ヘルシー アッサリ フワフワ
- 豆腐ハンバーグ 特徴 フワフワ フンワリ ヘルシー  
⋮

作成した全レシピの材料文章、カテゴリ文章、特徴文章を1つのテキストにまとめて Word2Vec の学習コーパスとして使用する。

### 3.3 レシピ単語ベクトル空間の構築

Word2Vec を用いてレシピ単語のベクトル空間を構築する。Word2Vec は通常、数十から数百万の語彙数に対して100から300次元のベクトル空間を構築する。本研究では語彙数が7,411個の比較的小さいコーパスを学習するため構築するベクトル空間の次元数は30に設定した。

### 4 レシピ単語ベクトル空間の評価

レシピ単語ベクトル空間の性能評価を、可視化を用いて行う。30次元のレシピ単語ベクトル空間に主成分分析を用いて2次元に圧縮した。2次元上にレシピ名をプロットし、その位置関係を評価する。一部のレシピ名をプロットした結果を Table 1 に示す。また、図上の円はプロットされたレシピ名を主観評価によってグループ分けしたものである。

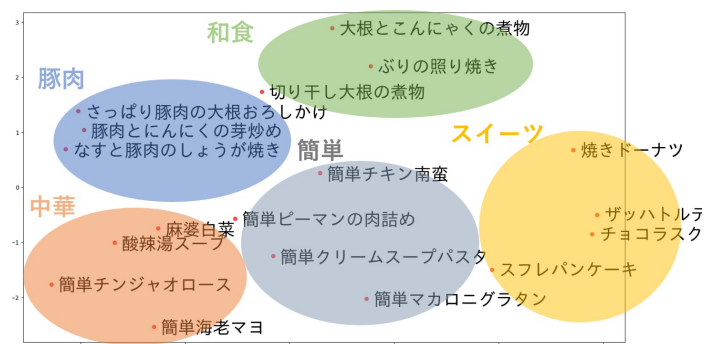


Table 1. レシピ単語ベクトル空間の2次元可視化

豚肉といった材料、中華、和食、スイーツといったカテゴリ、簡単といった特徴ごとにレシピ単語の関係性を表現できるベクトル空間の構築が確認できる。また、スイーツ内で同じチョコを使った「チ

「ヨコラスク」「ザッハトルテ」に近い位置にプロットされていることから、分離されたグループ内でも材料、カテゴリ、特徴情報に従って分布していることがわかる。

### 5 レシピ単語のベクトル演算を用いたレシピ推薦

レシピ単語のベクトル演算を用いてレシピ推薦を行う。材料、カテゴリ、特徴単語を入力単語として、入力単語の単語ベクトルと類似度の高いレシピの単語ベクトルを推薦レシピとして取得する。類似度の計算にはコサイン類似度を用いる。また、推薦されたレシピが妥当であるか評価を行う。学習に使用したコーパスの材料、カテゴリ、特徴単語はカタカナに統一しているため、類似度演算時に入力する単語はカタカナに変換している。

#### 5.1 材料単語ベクトルを用いた類似度演算

材料単語ベクトルを入力とし、類似度順に5件の推薦レシピを取得する。また推薦レシピの材料に入力した材料単語が含まれているかを確認する。含まれている場合は○、含まれていない場合は×とした。「鶏肉」、「じゃがいも」、「卵」の単語ベクトルを入力した際のレシピ推薦結果をそれぞれ Figure 4,5,6 に示す。

Figure 4. 「鶏肉」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
鶏胸肉の粗挽き塩コショウ串	0.733	○
鶏肉とペンネのアラビアータ	0.674	○
鶏もも肉のバルサミコ酢のソテー	0.665	○
大分名物☆鶏むね肉でサクサクとり天	0.659	○
ボリュームのある一品 鶏むね肉で卵なしカツレツ	0.657	○

Figure 5. 「じゃがいも」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
じゃがいもと人参のグラッセ	0.758	○
残りカレーのコロッケ	0.684	○
じゃがいもとカレー粉で🍀カレーポテトコロッケ♥	0.666	○
家でジャガイモの長期保存法・甘味が増します	0.662	○
新じゃが芋の明太子マヨネーズ焼き	0.656	○

Figure 6. 「卵」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
3分で簡単！中身はふわふわ半熟チーズオムレツ	0.555	○
でか！！スコッチエッグ	0.549	○
めっちゃウマ天津炒飯	0.528	○
スパムおにぎらず	0.523	○
レンジでとろーり温泉たまご	0.510	○

「鶏肉」、「じゃがいも」、「卵」のレシピ推薦結果では、推薦されたレシピ全てに入力した材料単語が含まれており、かつ推薦されたレシピは入力した材料単語を主要な材料とするものであった。次に「白菜」、「砂糖」を入力した際の推薦レシピ結果をそれぞれ Figure 7,8 に示す。

Figure 7. 「白菜」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
食べだしたら止まらない！やみつき白菜	0.751	○
あったか豚粕汁☆冬のご馳走	0.697	×
ポカポカ温まる！野菜たっぷりみぞれ鍋	0.683	×
豚肉と白菜の味噌鍋♪	0.665	○
白菜のクリーム煮	0.660	○

Figure 8. 「砂糖」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
ホッと☆チリペッパーベトナムカフェオレ♪	0.447	×
トルコのパン★ゴマ付きシミット	0.424	×
むかごのあまから炒め	0.415	×
赤なまこのさばき方	0.414	×
牛乳で☆キャラメルクリームor☆生キャラメル？	0.401	×

「白菜」のレシピ推薦結果では、5件中2件に「白菜」が材料に含まれないレシピが推薦され、「砂糖」におけるレシピ推薦結果では、推薦された5件全て「砂糖」が材料に含まれないレシピが推薦された。

Word2Vecでは文脈の共起関係に基づき、注目している単語の周辺単語と同じような周辺単語を持つ単語を類似性の高い単語として学習する。学習コーパスは、レシピに含まれる材料を様々な組み合わせで配置し材料のパターンを学習させている。そのため、「白菜」のように使用される料理が限定的で、周辺の材料単語のパターンが少ない場合や、反対に「砂糖」のようにどの料理にでも使用され、周辺の材料単語のパターンが多すぎる場合にはうまくレシピ推薦ができない。しかし、推薦されるレシピは入力した材料単語は含まれていない場合でも、入力単語を材料として使えそうなレシピである。そのため、検索をしたユーザーに向けてアレンジや創作レシピの提案などが期待できる。

#### 5.2 カテゴリ単語ベクトルを用いたレシピ推薦

カテゴリ単語ベクトルを入力とし、類似度順に上位5件の推薦レシピを取得する。また推薦レシピのカテゴリに入力したカテゴリ単語が含まれているかを確認する。「肉料理」、「中華料理」、「カレー」、「から揚げ」の単語ベクトルを入力した際のレシピ推薦結果をそれぞれ Figure 9,10,11,12 に示す。

Figure 9. 「肉料理」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
シークカバブ	0.682	○
ダッチオープンで豚ブロック燻製っ♪	0.674	○
みんな大好き！マヨ照り焼きチキン	0.671	○
ササミじゃないよ！鶏胸肉の粗挽き塩コショウ串	0.667	○
【低糖質】チキン×野菜のガリバタ醤油	0.661	○

Figure 10. 「中華料理」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
5分で作れる簡単エビチリ☆	0.731	○
プロ直伝！海老チリ	0.730	○
プロ直伝！麻婆茄子	0.710	○
旨辛★担々麺	0.709	○
簡単に美味しい♪焼きビーフン	0.708	○

Figure 11 「カレー」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
ドライカレー	0.808	○
しらたき de カレーヌードル	0.782	○
インド人コックさん直伝カレー粉で簡単キーマカレー	0.765	○
簡単おいしい☆キーマカレー	0.762	○
子供が喜ぶカレー	0.757	○

Figure 12. 「から揚げ」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
居酒屋の味！簡単タコのから揚げ	0.800	○
骨まで食べれる♪ カレイの唐揚げ	0.752	○
簡単♪カラッと揚げるししゃも唐揚げ	0.735	○
糖質制限★芝海老のからあげ	0.734	○
絶品☆カリカリ蛸の唐揚げ	0.733	○

「肉料理」、「中華料理」、「和風」、「カレー」、「から揚げ」のレシピ推薦結果では、推薦されたレシピ全てに入力したカテゴリが登録されているレシピであった。次に、「スイーツ」のレシピ推薦結果を Figure 13 に示す。

Figure 13. 「スイーツ」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
秋のスイーツ♪簡単スイートポテト	0.772	○
焼きバーム*ナッツ&チョコ生クリーム☆お家スイーツ	0.765	○
Kirch-Eistörtchen	0.700	×
おこしブランラズベリーアイス	0.690	×
とっても簡単♪チョコトルテ Schokotorte	0.678	×

「スイーツ」のレシピ推薦結果では、上位 5 件全て、スイーツとして妥当なレシピが推薦できているが、下位 3 件のレシピは、「スイーツ」のカテゴリが登録されていないレシピである。これは、レシピ内に明示的に「スイーツ」という単語が含まれていなくとも、スイーツによく含まれる「砂糖」や「チョコレート」の様な材料単語や「甘い」の様な特徴単語の組み合わせから「スイーツ」の要素を学習できているからである。従来のキーワード検索では、キーワードが含まれるレシピしか推薦できないのに対し、本手法ではキーワードが含まれないレシピでも柔軟に推薦できることが明らかとなった。

### 5.3 特徴単語ベクトルを用いたレシピ推薦

特徴単語ベクトルを入力とし、類似度順に上位 5 件の推薦レシピを取得する。また推薦レシピに入力した特徴単語が含まれているかを確認する。「辛い」、「シャキシャキ」、「しっとり」、「ヘルシー」、「濃厚」を入力した際のレシピ推薦結果をそれぞれ Figure 14,15,16,17,18 に示す。

Figure 14. 「辛い」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
お箸が進む〜！あとを引く辛さが旨いトッポギ☆	0.503	○
天ぶら屋さんの天井のタレ	0.502	○
ご飯すすむ♪高菜のピリ辛油炒め	0.487	○
甘辛コチュジャンで韓国の鶏唐揚げヤンニョムチキン	0.486	○
もやしたつぷり☆豚キムチ	0.457	○

Figure 15. 「シャキシャキ」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
簡単☆黒瀬スパイス de 無限ピーマン♪	0.612	○
シャキシャキ長いもの豚バラ巻き	0.599	○
シャキシャキ☆エシャレットの肉巻き	0.596	○
ちゃちゃっと美味しい♪豆苗炒め	0.581	○
サッパリヘルシー！長芋ともやしのバタポン炒め	0.579	○

Figure 16. 「しっとり」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
ふんわりしっとり本場のスコーン	0.611	○
5分でできるむちむちリンゴ蒸しパン	0.591	○
塩麴でしっとり”鶏むね肉のレンジで鶏ハム”	0.563	○
★簡単！HMでチーズケーキ風【ヨーグルトケーキ】	0.561	○
ホットケーキミックスで超簡単りんごケーキ	0.553	○

Figure 17. 「ヘルシー」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
ダイエット応援！おからのチキンナゲット	0.503	○
とろ〜りヘルシー♪エリンギの豆腐グラタン	0.490	○
厚揚げのそぼろあんかけ	0.483	○
簡単！節約！ボリューム！厚揚げの酢豚風♪	0.478	○
バジルとチーズのヘルシーおから豆腐コロッケ♪	0.476	○

Figure 18 「濃厚」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
一晚なんて待てないっ！超時短で水切りヨーグルト☆	0.664	○
本格濃厚♪簡単にできる半生ガトーショコラ	0.662	○
フライパンでほったらかし〜簡単濃厚チョコプリン♪	0.641	○
かぼちゃのスープ♪	0.625	○
混ぜるだけ！濃厚パンナコッタ	0.622	○

5件の特徴単語を入力として、推薦レシピを取得した結果、上位5件は全て入力特徴単語が含まれるレシピを推薦できた。しかし、材料単語でのレシピ推薦同様に、「甘い」など多くのレシピに使用される特徴単語においては、有意な結果は得られなかった。

#### 5.4 合成単語ベクトルを用いた類似度演算

材料、カテゴリ、特徴単語の複数の入力単語ベクトルを合成したベクトルを入力ベクトルとして類似度順に上位5件の推薦レシピを取得する。また合成ベクトルに用いた単語が全て推薦されたレシピに含まれている場合は○、1つでも含まれていない場合は×とした。

「キムチ」+「鍋」、「イチゴ」+「スイーツ」、「卵」+「中華」+「スープ」のレシピ推薦結果を順に Figure 19,20,21 に示す。

Figure 19 「キムチ」+「鍋」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
もやしたっぷり☆豚キムチ	0.755	○
簡単♪キムチ鍋	0.755	○
お家で簡単♪チゲ鍋♪キムチ鍋	0.730	○
さば味噌煮缶のキムチ豆腐	0.723	×
豚肉、もやし、キムチのクッパ♪	0.705	×

Figure 20. 「イチゴ」+「スイーツ」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
オレンジと苺の柑橘クリームフルーツサンド	0.844	○
甘酸っぱい♪いちご&ヨーグルトクリームサンド	0.838	○
簡単☆おうちにあるものでパフェ♪	0.826	○
パーティーにぴったり！簡単パフェ♪	0.792	○
市販のロールケーキで☆デコケーキ	0.784	○

Figure 21. 「卵」+「中華」+「スープ」のレシピ推薦結果

推薦レシピ	コサイン類似度	含まれるか
お店のあの味！絶品！中華卵スープ♪	0.810	○
とろとろ〜モロヘイヤの中華卵スープ	0.789	○
とろとろ白菜の中華卵スープ	0.783	○
モロヘイヤのとろとろスープ	0.772	○
簡単なのに絶品！干しエビの中華スープ	0.748	×

複数の入力単語ベクトルを合成し、推薦レシピを取得した結果、それぞれの入力単語の特性を加味したレシピの推薦が確認できた。推薦されたレシピには推薦レシピとして適切でないレシピも含まれている。これは、合成する単語ベクトルが多くなるにつれて推薦の条件が厳しくなる点や、本研究で扱うレシピが3070品と少数なためである。学習するレシピ数を増やすことで合成単語ベクトルに対して適切なレシピ推薦率の向上が期待できる。

#### 5.5 ユーザーの調理履歴に用いたレシピ推薦

ユーザーの調理履歴に基づいたレシピを推薦できれば、より効率的なレシピ推薦が可能となる。例えば、とあるユーザーの調理履歴にヘルシーなレシピが多い場合、そのユーザーはヘルシー志向の持ち主であると判断できる。そこでレシピ推薦の際に、ヘルシーなレシピを優先的に推薦することで、ユーザーの嗜好を考慮した効率的なレシピ推薦ができる。本節では、ユーザーの調理履歴A、調理履歴Bが取得済みであると仮定し、ユーザーの料理履歴合成ベクトルと入力したレシピ情報の合成ベクトル



を入力としてレシピ推薦を行う。

調理履歴 A「ダイエット応援！おからのチキンナゲット」+調理履歴 B「エリンギのヘルシー豆腐グラタン」+「コロッケ」を入力とした際のレシピ推薦結果を Figure 22 に示す。

Figure 22. 調理履歴+「コロッケ」のレシピ推薦結果

推薦レシピ	コサイン類似度
揚げないヘルシー焼きコロッケ	0.807
バジルとチーズのヘルシーおから豆腐コロッケ	0.804
揚げない*かぼちゃコロッケ	0.802
揚げないヘルシーなカボチャのコロッケ	0.793
おからコロッケ	0.769

調理履歴から、ヘルシー志向をもつユーザーだと仮定できる。結果よりユーザー嗜好を反映したヘルシーかつ「コロッケ」の料理が推薦できた。

調理履歴 A「韓国味噌汁テンジャンチゲ」+調理履歴 B「韓国冷麺ネンミョン」+入力単語「鍋」のレシピ推薦結果を Figure 23 に示す

Figure 23. 調理履歴+「鍋」のレシピ推薦結果

推薦レシピ	コサイン類似度
簡単キムチ鍋	0.807
豆腐と長ネギのチゲ風スープ	0.804
豚バラ肉とのカムジャンタン風鍋	0.802
韓国テンジャングク	0.793
サムゲタン風スープ	0.769

調理履歴から、韓国料理を好むユーザーだと仮定できる。結果よりユーザー嗜好を反映した韓国料理かつ「鍋」のレシピが推薦できた。

調理履歴 A「お弁当にも☆スパゲティーバター醤油」+調理履歴 B「子供が喜ぶハッシュドビーフ」+入力単語「カレー」のレシピ推薦結果を Figure 24 に示す。

Figure 24. 調理履歴+「カレー」のレシピ推薦結果

推薦レシピ	コサイン類似度
子供が喜ぶカレー	0.882
子供が喜ぶウインナーとジャガイモのチーズ焼き	0.836
市販のルーで簡単カレーチャーハン	0.822
フライパンで簡単カレーピラフ	0.818
大人も子供も大好き我が家のポークケチャップ	0.807

調理履歴から、子供に料理を作るユーザーだと仮定できる。結果より最も類似度が高いレシピにユーザー嗜好を反映した子供料理かつ、「カレー」のレシピが推薦できた。

## 6. おわりに

本研究ではレシピ単語のベクトル空間を用いた、キーワード検索より柔軟なレシピの推薦を目的とした。レシピサイトから収集したレシピ情報をもとに Word2Vec の学習コーパスを作成し、レシピ単語ベクトル空間を構築した。構築したレシピ単語ベクトル空間にベクトル演算を用いてレシピ推薦を行った。結果、直接的にキーワードが含まれないレシピにおいても推薦できる点や、ユーザーの調理履歴を元にレシピを推薦するなど、キーワード検索では実現できない機能を実現できた。しかし、どのレシピにも含まれる単語や限定的なレシピにしか含まれない単語を使つてのレシピ推薦は上手くできなかった点において、検討が必要であり今後の課題としたい。

## 7. 参考文献

- [1] Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean "Efficient Estimation of Word Representations in Vector Space," ICLR, 12pages, (2013)
- [2] 早川文代, 奥井加奈, 阿久澤さゆり, 齋藤昌義, 西成勝好, 山野善正, 神山かおる:日本語テキスト用語の収集, 日本食品化学学会誌, 53m 327-336(2006)
- [3] 矢野達也, 林豊洋, 大橋健:食品単語のベクトル空間の構築と評価 (第 2 報), 人口知能学会研究会, (2017)
- [4] 田中規介, 林豊洋, 大橋健:家具 EC サイトにおける商品画像を用いたおすすめ機能の提案 (2017)
- [5] 原島純:レシピサービスと情報処理, 人工知能学会誌, vol.34 3-8(2019)
- [6] 西尾泰和, word2vec による自然言語処理, オライリー・ジャパン, (2014)