

献立の雰囲気考慮した料理推薦システムの提案

Proposal of dishes recommendation system considering the atmosphere of menu

大野 礼儀^{*1} 福原 知宏^{*2} 山田 剛一^{*1} 増田 英孝^{*1}
 Masanori Ono Tomohiro Fukuhara Koichi Yamada Hidetaka Masuda

^{*1} 東京電機大学
 Tokyo Denki University

^{*2} 産業技術総合研究所
 National Institute of Advanced Industrial
 Science and Technology

With the spread of the Internet, information about recipes and menu is available on the web. When a person plans a menu which is a set of one meal of dishes, he or she has to consider the combination of dishes so that they have a harmony. We propose a menu planning support system considering the atmosphere of dish combination. As a preliminary study, we analyzed combinations of dishes which are on a recipe site. We performed a clustering of recipes, and analyzed relations between recipe clusters within a menu. An overview of the system, and analysis results are described.

1. はじめに

料理レシピサイトと呼ばれるサイトが支持を集めている。cookpad[cookpad 2015]に代表されるこれらのサイトの中には、単品の料理だけでなく、料理を献立単位でまとめて紹介しているものもある。

ユーザが一食分の献立を考える為にこれらのサイトを利用する際、作りたい料理や「中華風」などの献立の雰囲気をもとに献立を組み立てていく。単品料理が紹介されているサイトの場合、献立内の料理の組み合わせをユーザ自身が考える必要があり、労力がかかるほか、適切な組み合わせが得られるとは限らない。また献立単位で紹介されているサイトであっても、ユーザの嗜好にあった献立を見つけることができるとは限らず、献立に好みではない料理が含まれている場合にも、その料理の入れ替えなどの融通は利かない場合がほとんどである。

そこで本研究では作りたい料理や献立の雰囲気は決まっているが、献立全体を決められないユーザの献立作成を支援することを目的とし、献立全体の雰囲気や作りたい料理を起点とした料理推薦システムを開発している。

献立を扱った先行研究と我々の研究の違いについて触れる。献立を扱った先行研究としてユーザ情報をもとに献立の栄養バランスやアレルギーを考慮して一週間分の献立を推薦する西川らの研究[西川 2013]やレシピサイトから取得したレシピデータとユーザが登録した献立データをもとにユーザが入力したレシピに合った献立を推薦する井出らの研究[井出 2012]がある。

西川らの研究[西川 2013]は栄養面やアレルギーの有無などを主体として献立の推薦を行っている。そのため、レシピ間の相性や献立の雰囲気を主体に献立作成支援を目指している我々の研究とは異なる。また井出らの研究[井出 2012]は我々の研究同様に料理同士の相性やユーザの嗜好をもとに推薦を行っているが、献立の構成情報をユーザが登録するという点で、献立の構成情報にウェブ上のデータを利用して献立の推薦を目指している我々の研究とは異なる。

本稿では献立作成支援システムの実現に向けて、どのような特徴を持つ料理同士が一つの献立を形成するのか、分析を行う。レシピ情報をもとに料理のクラスタリングを行い、献立内に出現する料理レシピ共起頻度とクラスタリング結果を用いて料理ク

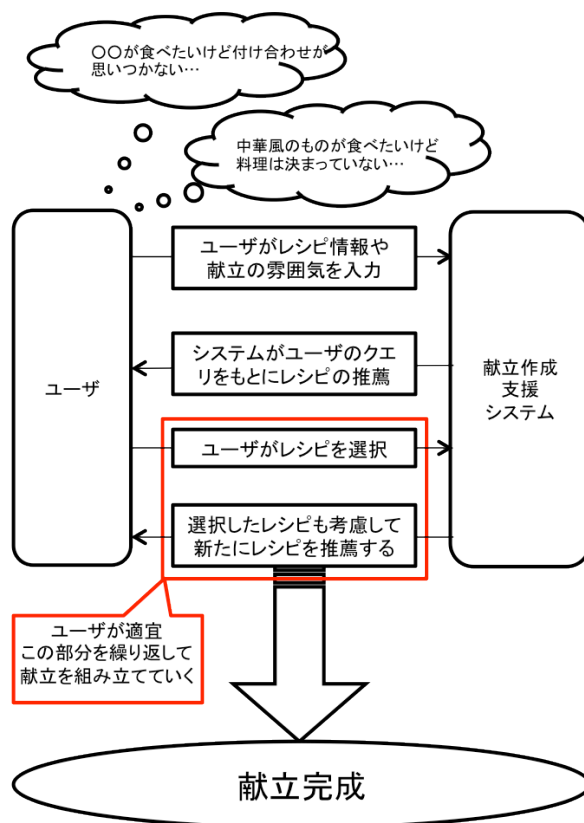


図 1：献立作成支援システムイメージ

ラスト間の相性と雰囲気を分析した結果を報告する。

2. 献立作成支援システム

本研究の目指しているシステムの概要説明とシステム実現のために本稿で報告する分析の全体像を記す。

献立作成支援システムのイメージを図 1 に示す。ユーザが作りたい料理のレシピ情報や雰囲気をクエリとしてシステムに入力すると相性の良い料理レシピを推薦する。ユーザが推薦されたレシピ群からレシピを選択すると、そのレシピも含めたクエリで次の料理を推薦する。それを繰り返して献立作成を支援する。本

システムを使えば、既存の献立内の一部のレシピを入れ替えたいときにも、入れ替えたいレシピ以外のレシピをクエリとしてその献立に合った他のレシピの推薦を行うことができる。

次に本研究における雰囲気と相性の定義について説明する。雰囲気とは「洋風」や「暑い日に」などの言葉にできるものから、言葉にはできないが献立制作者が暗示的に設定したテーマを指す。また相性とは「同じ献立内に汁物が二つ存在することはあまりない」、「焼き魚とみそ汁は同じ献立に登場しやすい」などの料理が持つ特徴同士の関係性のことを指す。

本システムの実現には献立内のある料理がどういった特徴から献立内の他の料理と組み合わせる一つの献立を形成しているのかを知る必要がある。そのため、料理レシピをその特徴に基づいてクラスタリングする。その後、献立の構成情報とクラスタリング結果を用いてクラスタ同士の共起関係を見る。それにより、ある特徴をもつレシピクラスタがどのような特徴を持つ他のレシピクラスタと共に献立に登場しているのか、またレシピクラスタ同士の繋がりから雰囲気を認めることができるかを検討し、その結果を本稿で報告する。

3. 料理レシピのクラスタリング

レシピのクラスタリングを行うために使用したデータと特徴の抽出、クラスタリング手法について説明し、クラスタリング結果について述べる。

3.1 献立とレシピ情報

データは節約レシピ～一週間の献立～[節約レシピ 2015]内で紹介されている 2001 年から 2009 年までの献立データを使用する。節約レシピ～一週間の献立～ は主菜と複数の副菜を組み合わせた献立を一週間ごとに 6 つ紹介しているサイトである。サイト内の献立情報には表 1 のように献立のテーマとその献立を構成しているレシピが記されており、レシピにはその料理を作るためのレシピ情報が記載されている。献立情報には一週間ごとの献立のまとめや献立ごとにテーマの記述があり、献立に関する情報が多いためこのデータを利用した。クラスタリングに使用した献立とレシピ情報の総数を表 2 に示す。

次にレシピのクラスタリングに使用する特徴について説明する。レシピの特徴ベクトルの要素として、レシピ名に現れる単語とレシピに使用されている食材名を用いた。レシピ名に現れる単語の語彙数と、レシピに使用されている食材のユニーク数を表 3 に示す。これらの合計である 6,059 要素が特徴ベクトルの各次元となる。

3.2 レシピ情報ごとの前処理

本節ではクラスタリングに使用するレシピ名と食材名に行った前処理を説明する。

(1) レシピ名に現れる単語

レシピ名は単語に分割し、単語を一つの特徴語とする。形態素解析器 MeCab[MeCab 2015]を用いて文章を形態素に分割し、基本形に戻した。助詞、助動詞、接続詞や「○○入り」の「入り」などの重要ではない単語は除外し、「○○風」の「風」などの単体では重要ではない単語は前の単語と連結し一つの単語とした。

(2) 食材名

食材は単語に分割せず、食材一つを一つの特徴語とする。MeCab を用いて読みのカタカナに変換し、「しお」「塩」などの揺らぎを無くした。

表 1: 献立情報の例

一週間毎のテーマ	梅を使ったさっぱりレシピ
献立テーマ	暑い夏にぴったりのさわやかメニュー
主菜	ジャージャーうどん
副菜 1	厚揚げと搾菜の炒め物
副菜 2	冷やし茶碗蒸し

表 2: レシピのクラスタリングに使用した献立とレシピ情報の総数

献立数	2561
レシピ数	7702

表 3: レシピ名に現れる単語の語彙数と食材のユニーク数

レシピ名に登場する単語の語彙数	2402
食材のユニーク数	3657

3.3 レシピ情報の重み付け

本節では特徴語ごとの重み付け方法について説明する。

(1) レシピ名に現れる単語

レシピ名には出現回数の多い単語でも「サラダ」や「スープ」などレシピを分類するにあたって重要な単語が存在するため、レシピ名に現れる単語同士の重さに差を付けずにレシピに使用された食材が取り得る重さの最大値をレシピ名单語の重みとした。レシピ名に登場した単語 R の重さ W_r は(1)のように表す。

$$W_r = \ln(\text{レシピ数}) \cdots (1)$$

(2) レシピに使用された食材名

食材は調味料などの出現回数が多いので、重み付けを行う。予備実験によりレシピに使用された食材の重みをレシピ名に登場した単語よりも軽くした方が料理レシピの特徴ごとにクラスタリングされたため、レシピ名に登場する単語よりも重みが軽くなるように設定した。食材名を M としてそのレシピに登場した M の重さ W_m は(2)のように表す。

$$W_m = \frac{1}{\text{レシピに使用された食材数}} \times \ln\left(\frac{\text{レシピ数}}{M\text{のレシピ出現数}}\right) \cdots (2)$$

3.4 K-means によるクラスタリング

クラスタリングは統計解析ソフト R (<http://www.r-project.org>) の関数 `kmeans()` を用いてクラスタリングを行った。クラスタ数はクラスタ内のレシピ数が少なすぎると共起関係を求めるための情報量が不足してしまうため、クラスタ内のレシピ数が 5 以下の小さなクラスタができないこと目安に 200 に設定した。クラスタリングにより得られたクラスタの一部を表 4 に示す。表 4 はクラスタ番号とそのクラスタ内に含まれているレシピの数、そのクラスタ内に存在するレシピ名の一部抜粋、クラスタの中心点を参考に影響の強い特徴語を強さの降順に並べたものがカラムとなっている。

「煮る」や「みそ汁」「シチュー」といった料理のカテゴリが最も影響力が強い特徴語として集まったクラスタが多く見られた。それに加えて「クリーム」や「もやし」などのカテゴリを細分化する単語がそれ以降の特徴語として出現した。

一方で表 5 のクラスタのように一つの食材が大きく影響してしまい複数の料理カテゴリが出現したクラスタも存在した。表 5 の

表 4: クラスタリングにより得られたクラスタの例

クラスタ番号	所属レシピ数	クラスタ内のレシピ一部抜粋	影響力が強い特徴語
11	44	シメジとインゲンのご飯, たくあんご飯, じゃこ入り豆ご飯, えびサラダご飯, カリカリベーコン&じゃこご飯, 大根葉ご飯, 小豆ご飯, あおさご飯, サンマご飯, たけのご飯	ご飯, 豆, ひじき, じゃこ, 栗, 雑穀
24	24	キャベ玉みそ汁, なめこもやしのみそ汁, あさりのみそ汁, 豆腐ともやしのみそ汁, 結びかんぴょうのみそ汁, 白菜と大根のゴマ風味みそ汁, 根菜たっぷりおみそ汁, 大根と二色のねぎのみそ汁, 鮭のみそ汁, ごぼうと菊菜のみそ汁	みそ汁, もやし, たつぷり, 大根, 卵
162	20	コーンとかきのクリーム煮, ミートボールのクリーム煮, ほうれん草のクリーム煮, 鱈と白菜のクリーム煮, 豆腐とひき肉のコーンクリーム煮, しめじと鮭のクリーム煮, 青梗菜のクリーム煮, ポテトとホタテのクリーム煮, 豆腐と白菜のクリーム煮, 鶏むね肉のクリーム煮, ハムときのこのチーズクリーム煮	煮る, クリーム, 鶏, 豆腐, コーン
177	29	豆乳シチュー, シチューDEドリア, 迎春シチュー, さつまいも入りクリームシチュー, シチューのパイナップル, 白菜の豆乳シチュー, 牡蠣のクリームシチュー, ぶり大根シチュー, タラとじゃがいものミルクシチュー, タラのトマトシチュー	シチュー, クリーム, 白菜, 豆乳, カボチャ

表 5: 食材名に大きく影響されて料理の特徴がうまく出なかったクラスタ例

クラスタ番号	所属レシピ数	クラスタ内のレシピ一部抜粋	影響力が強い特徴語
110	34	チンゲン菜と厚揚げののんにく炒め, チンゲン菜とほたてのシチュー, ちくわとチンゲン菜のかか和え, 野菜天とチンゲン菜のめんつゆ煮, えのきだけとチンゲン菜の和え物, チンゲン菜のザーサイ和え, チンゲン菜の黒ごま和え	チン, ゲン, 菜, 炒める, 煮る,

表 6: 共起関係の分析には除外したクラスタ

クラスタ番号	所属レシピ数	クラスタ内のレシピ一部抜粋	影響力が強い特徴語
80	921	ポパイピザ, 生麩のペペロンチーノ風, りんごとレーズンの甘煮, ベトナム風はるまき, かりんとう風レバー, 貝割れ菜となめたけのスパゲッティ, ゴボウの唐揚げ, あんず白玉, カレージャーマンピザ, いんげんのマヨ味噌和え, シンプル焼きそば, 苺のムース	鍋, 天ぷら, 餃子, 中華風, カルパッチョ, 水菜, ニラ

ようなクラスタは出現頻度の少ない食材がレシピ名に出現している場合がほとんどであった。また、一つのクラスタに 900 以上のレシピが属する特徴の希薄なクラスタ(表 6)が生成された。このクラスタに所属していたレシピはレシピ名に未知語を含んでいたために、単語分割がうまくいかずにレシピ名が一単語として認識されてしまい情報量が不足したため、大きな一つのクラスタに分類されてしまったと考えられる。

4. クラスタ同士の共起関係

2. で作成したクラスタを献立の構成情報をもとに共起関係を分析し、雰囲気や相性を見ることができると検討する。

4.1 献立構成情報を用いたベクトル空間の作成

クラスタ内の料理レシピが献立に登場するか否かを特徴として 2. で作成したクラスタをベクトル空間で表現する。表 6 のクラスタはクラスタの特徴が希薄で数が多いため、この分析では除外した。

献立構成情報を用いて作成したベクトル空間をもとにクラスタ同士の距離を求める。距離の尺度にはコサイン類似度を利用した。クラスタ A とクラスタ B のベクトルをそれぞれ a, b としたとき $\cos \theta$ を(3)に表す。 $\cos \theta$ を用いてクラスタ A, B 間の距離 $D(A, B)$ は(4)のように表す。

$$\cos \theta = \frac{a \cdot b}{|a| |b|} \cdots (3)$$

$$D(A, B) = 1 - \cos \theta \cdots (4)$$

このベクトル空間でクラスタ同士の距離が近いことはクラスタに属するレシピ同士が同じ献立に出現している頻度が高いことを示している。

4.2 クラスタ間の相性と雰囲気の発見

表 4 で示したクラスタを対象に、各クラスタからの距離が近いクラスタを順に 3 つ並べたものを表 7 に示し、そのクラスタ間の距離を表 8 に示す。クラスタ番号 11 は「ご飯」を特徴に持っており、そのおかげになりそうな「かき揚げ」や「炒め物」を特徴にもつクラスタが共起頻度の高いものとして出現している。また「ご飯」と「かき揚げ」の組み合わせには、日本の家庭料理と表現されるような雰囲気が感じられる。

クラスタ番号 177 は「シチュー」を特徴に持っており、それに合いそうな洋風で汁物ではない「ソテー」や「パイ」、「オムレツ」などの料理を特徴に持つクラスタが共起頻度の高いものとして出現している。またクラスタ間で洋風と表現されるような雰囲気が見て取れる。

表 7: 対象クラスタと対象クラスタに近いクラスタ上位 3 つが持つ特徴語

対象のクラスタ		対象に最も近いクラスタ		対象に2番目に近いクラスタ		対象に3番目に近いクラスタ	
クラスタ番号	影響が強い特徴語	クラスタ番号	影響が強い特徴語	クラスタ番号	影響が強い特徴語	クラスタ番号	影響が強い特徴語
11	ご飯, 豆, ひじき, じゃこ, 栗, 雑穀	18	かき揚げ, えび, 大豆, 三つ葉	115	炒める, 味噌, なす, ピーマン	101	ピリ, 炒める, 辛, 豚肉
24	みそ汁, もやし, たっぷり, 大根, 卵	70	煮物, 大根, ひじき, 大豆	16	キャベツ, ツナ, サラダ, 炒める	15	お好み焼き, たっぷり, チーズ, もやし
162	煮る, クリーム, 鶏, 豆腐, コーン	72	ベビー, リーフ, サラダ, シンプル	91	タマネギ, サラダ, みそ汁, プチトマト	54	ソテー, キノコ, ベーコン, トマト
177	シチュー, クリーム, 白菜, 豆乳	112	パイ, サーモン, 鮭, チーズ	141	マスタード, 粒, ソテー, サーモン	190	チーズ, オムレツ, ミート, アボカド

表 8: 対象クラスタと対象クラスタに近いクラスタ上位 3 つとの距離

対象のクラスタ		対象に最も近いクラスタ		対象に2番目に近いクラスタ		対象に3番目に近いクラスタ	
クラスタ番号	クラスタ番号	クラスタ間の距離	クラスタ番号	クラスタ間の距離	クラスタ番号	クラスタ間の距離	
11	18	0.90569	115	0.91742	101	0.92463	
24	70	0.93066	16	0.93814	15	0.94107	
162	72	0.90000	91	0.91548	54	0.91614	
177	112	0.87620	141	0.89278	190	0.92255	

5. 考察

4. の結果をもとにこれからの課題点について考察する。表 5 のクラスタのように料理カテゴリ以外のものが最も強い影響力を持った場合、複数の料理カテゴリが同じクラスタに分類されてしまう問題が発生した。これは料理レシピを一つのクラスタに分類するハードクラスタリングを用いて分析を行ったためであると考えられる。そのため料理が複数の特徴に分類されるようにソフトクラスタリングでの料理レシピの分類を検討したい。

表 6 の様な特徴が希薄で巨大なクラスタができあがる問題は、単語を正確に分割できないことが主な原因であると考えられ、Mecab の辞書に料理に関連する語を追加することで対応したい。また本稿ではレシピ名に登場する単語とレシピに使用された食材から特徴を抽出し、一対一で料理の相性の分析を行ったが、目的は献立作成支援であるため、レシピが三つ四つと増えたときにも相性の良い組み合わせを求められるかどうか検討したい。

6. おわりに

本稿では献立推薦支援システムの実現に向けて、料理レシピの特徴によりクラスタリングした料理レシピをもとに、献立の構成情報に基づく料理レシピの共起関係から、料理レシピ同士の相性や、料理レシピの組み合わせが持つ雰囲気について検討し、その結果を報告した。「みそ汁」や「サラダ」などの料理カテゴリに関する単語が最も強い影響力を持ったクラスタ同士ではクラスタ間に相性や雰囲気にあたるものを見ることができた。

参考文献

- [cookpad 2015] cookpad, <http://cookpad.com>, 2015.
- [西川 2013] 西川智佳, 伊藤孝行, 永井明彦, 丸山智美: 献立表自動生成におけるユーザのフィードバックに基づく献立再調整アルゴリズム, 第 27 回人工知能学会 全国大会論文集, 人工知能学会, 2013.
- [井出 2012] 井出歩, 大友崇弘, 藤澤伸, 服部隆志: 集合知を用いた献立推薦システム, 情報処理学会第 74 回全国大会, 情報処理学会, 2012.
- [節約レシピ 2015] 節約レシピ, <http://www.s-recipe.com>, 2015.
- [MeCab 2015] MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://mecab.sourceforge.net/>, 2015.