

# Non-ingredient Detection in User-generated Recipes using the Sequence Tagging Approach

Yasuhiro Yamaguchi, Shintaro Inuzuka, Makoto Hiramatsu, and Jun Harashima  
Cookpad Inc.

Yebisu Garden Place Tower 12F, 4-20-3 Ebisu, Shibuya-ku, Tokyo, 150-6012, Japan  
{yasuhiro-yamaguchi, shintaro-inuzuka, makoto-hiramatsu, jun-harashima}@cookpad.com

## Abstract

Recently, the number of user-generated recipes on the Internet has increased. In such recipes, users are generally supposed to write a title, an ingredient list, and steps to create a dish. However, some items in an ingredient list in a user-generated recipe are not actually edible ingredients. For example, headings, comments, and kitchenware sometimes appear in an ingredient list because users can freely write the list in their recipes. Such noise makes it difficult for computers to use recipes for a variety of tasks, such as calorie estimation. To address this issue, we propose a non-ingredient detection method inspired by a neural sequence tagging model. In our experiment, we annotated 6,675 ingredients in 600 user-generated recipes and showed that our proposed method achieved a 93.3 F1 score.

## 1 Introduction

At the present time, many people upload their recipes to the Internet. For example, over 6.7 million recipes have been uploaded to Cookpad,<sup>1</sup> one of the largest recipe sharing services in the world. Most of the recipes on the service are posted by ordinary users.

Figure 1 shows an example of a recipe. Note that we use Japanese examples in this study because approximately half of the recipes on Cookpad are written in Japanese. As seen in the figure, a recipe generally consists of a title, ingredient list, and steps. An ingredient list is a set of items that have an ingredient name and quantity.

However, some items in an ingredient list in a user-generated recipe are not actually edible ingredients in a user-generated recipe. For example, the third item 調味料 (seasoning) in Figure 1 is

<sup>1</sup><https://cookpad.com>

| Title   |                |
|---|----------------|
| ナスとピーマンの味噌炒め<br>(Eggplant and Green Pepper Miso Stir-fry) |                |
| Ingredient list   |                |
| ナス (eggplant)   | 5 個 (5 pieces) |
| ピーマン (green pepper)                                       | 5 個 (5 pieces) |
| 調味料 (seasoning)   | N/A            |
| 味噌 (miso)   | 大さじ 3 (3 tbs)  |
| 砂糖 (sugar)  | 大さじ 2 (2 tbs)  |
| 酒 (sake)  | 大さじ 2 (2 tbs)  |
| Steps   |                |
| 1. ナスを輪切りにする<br>(cut eggplants into round slices)         |                |
| 2. ...  |                |

Figure 1: Example of a recipe. The N/A means that the user (i.e., recipe author) has not written the information.

not an ingredient but the heading for the following ingredients. In a user-generated recipe, people freely use the ingredients field to describe ingredients. This noise makes it difficult for computers to use recipes for a variety of tasks, such as calorie estimation.

In this paper, we propose a method to detect non-ingredient items from an ingredient list in a user-generated recipe. Inspired by a sequence tagging approach, our method solves the problem by predicting a label (ingredient or non-ingredient) for each item in an ingredient list sequentially. In our experiment, we annotated 6,675 ingredients in 600 recipes from Cookpad and investigated the performance of our method using the recipes.

## 2 Related Work

The increase in the number of recipes on the Internet has led to an increase in studies on these data, such as recipe analysis (Sasada et al., 2015; Hiramatsu et al., 2019), recipe organization (Kiddon et al., 2015; Jermurawong and Habash, 2015), and recipe

generation (Salvador et al., 2019; Kiddon et al., 2016). Additionally, many recipe-related corpora and datasets have been published recently to promote studies on recipes (Mori et al., 2014; Harashima et al., 2016; Salvador et al., 2017; Yagcioglu et al., 2018).

Among such recipe-related studies, the following two previous works focused on informal text in user-generated recipes, like our study. Harashima and Yamada (2018) converted ingredients written in an user-generated recipe into their canonical forms in an ingredient dictionary. However, in that study, there was no assumption that non-ingredients appear in a recipe, unlike our study.

By contrast, Inuzuka et al. (2018) distinguished non-steps written in a user-generated recipe from actual steps, such as an advertisement for the author’s recipe books, which are not related to cooking. Our study focuses on ingredients in a recipe, unlike their work; that is, we distinguish non-ingredients written in a user-generated recipe from actual ingredients.

Our study is the first to pay attention to non-ingredients in a user-generated recipe. This contributes to a variety of recipe-related studies, particularly based on ingredients in a recipe, such as calorie estimation (Harashima et al., 2020), recipe clustering (Nadamoto et al., 2016), and recipe-related term detection (Chung, 2012).

### 3 Task Definition

The primary task in this study is to classify an item in an ingredient list as an ingredient or non-ingredient. In this work, we define non-ingredient items based on edibility. Figure 2 shows examples of ingredient lists in user-generated recipes. “ホワイトソース” (white sauce) in Figure 2(a) is not an ingredient but a heading. The items below it are ingredients for white sauce. An item without a quantity is likely to be a non-ingredient. By contrast, “お好みの野菜” (favorite vegetables) in Figure 2(b) is an actual ingredient. As shown by this example, an item without a quantity is not always a non-ingredient. “(↑バターでもいいです)” ((↑ you can use butter)) in Figure 2(c) is used as a comment, which mentions the previous ingredient “マーガリン” (margarine). “竹串” (bamboo skewers) in Figure 2(d) is a non-ingredient because it is not edible. In some recipes, kitchenware appears on the ingredient list, like this example. The goal

| Ingredient list       |                |     |
|-----------------------|----------------|-----|
| ...                   |                | ... |
| ローリエ (bay leaf)       | 2 枚 (2 pieces) |     |
| ホワイトソース (white sauce) |                | N/A |
| バター (butter)          | 30g (30 grams) |     |
| ...                   |                | ... |

(a) Example of a cream stew recipe.

| Ingredient list              |                  |     |
|------------------------------|------------------|-----|
| ...                          |                  | ... |
| にんじん (carrot)                | 1 本 (1 piece)    |     |
| お好みの野菜 (favorite vegetables) |                  | N/A |
| 水 (water)                    | 1000cc (1000 cc) |     |
| ...                          |                  | ... |

(b) Example of a hot pot recipe.

| Ingredient list                       |                |     |
|---------------------------------------|----------------|-----|
| ...                                   |                | ... |
| マーガリン (margarine)                     | 60g (60 grams) |     |
| (↑バターでもいいです) ((↑ you can use butter)) |                | N/A |
| 砂糖 (sugar)                            | 40g (40 grams) |     |
| ...                                   |                | ... |

(c) Example of a cookie recipe.

| Ingredient list      |                     |     |
|----------------------|---------------------|-----|
| ...                  |                     | ... |
| サラダ油 (vegetable oil) | 小さじ 1 (1 tsp)       |     |
| 竹串 (bamboo skewers)  | 3 本 (3 pieces)      |     |
| 岩塩 (rock salt)       | 適量 (desired amount) |     |
| ...                  |                     | ... |

(d) Example of a meat roll recipe.

Figure 2: Examples of ingredient lists.

of this study is to detect these inedible items as non-ingredients.

## 4 Proposed Method

In this study, we detect non-ingredient items in an ingredient list using a neural sequence tagging model, shown in Figure 3.

### 4.1 Ingredient Representations

First, we convert each item in the ingredient list into its ingredient representation, which consists of an ingredient name representation and additional features. The former is obtained as follows:

**TF-IDF:** We compute TF-IDF vectors for each item in the ingredient list. The term frequency and inverse document frequency are given as

$$\text{tf}(i, j) = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (1)$$

$$\text{idf}(j, D) = \log \frac{|D|}{|\{d \in D : t_i \in d\}|}, \quad (2)$$

where  $n_{i,j}$  is the number of words  $t_i$  in the  $j$  th ingredient name,  $d$  is the set of tokenized words in the ingredient name, and  $D$  is the set of all ingredient names in the recipe dataset. We tokenize each

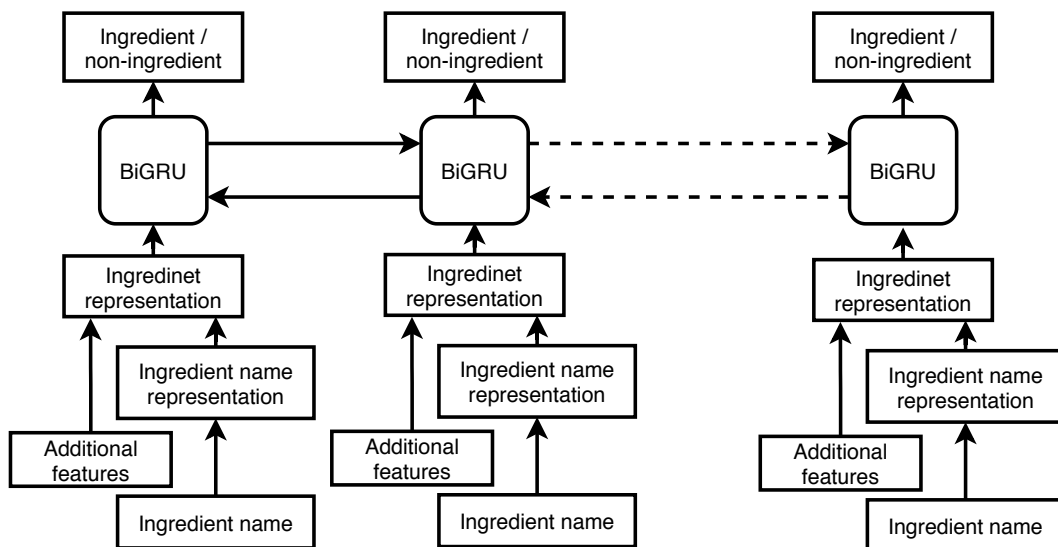


Figure 3: Overview of our method.

ingredient name and compute the TF-IDF values for each token.

**char-CNN:** Instead of TF-IDF, we can also use a CNN-based sequence encoder (Zhang and Wallace, 2017) to obtain the character-level features of ingredient names. We compute the features using different kernel sizes (2, 3, 4, 5) and concatenate them.

Note that we do not use pre-trained embeddings such as GloVe or fastText for the ingredient name representation because these embeddings did not show a good performance in our preliminary experiments.

In addition to the ingredient name representation, we use additional features of the ingredient name and quantity: character count and ingredient name frequency. The ingredient name frequency is computed from the recipe dataset, which is also used for the TF-IDF calculations. We count the ingredients with the same name and use the log-scaled value as the name frequency.

Finally, we concatenate the ingredient name representation and additional features into one vector to create the ingredient representation.

## 4.2 Model

We use the sequence tagging model shown in Figure 3. Each time step corresponds to each item in the ingredient list. The model takes all items in the ingredient list as its inputs in the order that they appear in the recipe. Whereas the inputs of our model are ingredient representations described in the previous section, the outputs of our model are

|                      |       |
|----------------------|-------|
| # of recipes         | 600   |
| # of ingredients     | 5,829 |
| # of non-ingredients | 846   |

Table 1: Statistics of our dataset.

binary predictions, each of which represents an ingredient or non-ingredient.

By performing a non-ingredient detection task as a sequence tagging problem, the model can make predictions by taking items before and after the target item into account. In many recipes, related ingredients are usually written close to each other. As shown in Figure 2(a), if an item in an ingredient list is used as a heading, related ingredients are listed below it.

## 5 Experiment

### 5.1 Dataset

In our experiment, we chose 600 recipes from Cookpad. More precisely, we collected recipes whose ingredient lists contained items without quantity information because such items tended to be non-ingredients in our preliminary investigation. Each ingredient in the recipes was labeled as an ingredient or non-ingredient by three domain-expert annotators. The gold labels were decided by majority vote. Table 1 shows the statistics of our dataset.

### 5.2 Methods

In our experiment, we compared the performance of the following methods using 10-fold cross vali-

| Model              | F1                | Precision         | Recall            |
|--------------------|-------------------|-------------------|-------------------|
| Random Forest      | 87.2 ± 5.1        | 82.8 ± 8.3        | 92.8 ± 4.5        |
| + ingredient freq. | 88.8 ± 4.1        | 86.8 ± 6.6        | 91.4 ± 4.8        |
| BiGRU + TF-IDF     | 90.8 ± 3.1        | 90.1 ± 3.7        | 91.2 ± 3.6        |
| + ingredient freq. | 91.6 ± 3.4        | 89.4 ± 4.8        | 94.1 ± 3.6        |
| BiGRU + char-CNN   | 91.2 ± 2.7        | 91.3 ± 4.8        | 91.4 ± 3.5        |
| + ingredient freq. | <b>93.3 ± 2.3</b> | <b>93.2 ± 3.7</b> | <b>94.1 ± 3.1</b> |

Table 2: Experimental results.

| Name           | Frequency |
|----------------|-----------|
| 砂糖 (sugar)     | 524,647   |
| 塩 (salt)       | 507,766   |
| 水 (water)      | 450,370   |
| 卵 (egg)        | 400,572   |
| 醤油 (soy sauce) | 320,834   |

Table 3: Top 5 ingredient names in our dataset.

dation:

**Random forest** (baseline model): We used *RandomForestClassifier* included in scikit-learn as a baseline model. The input of the random forest model was the ingredient representation described in the previous section. This model predicted a label for each item in an ingredient list independently.

**BiGRU model** (our model): We used two-layer bidirectional GRU (BiGRU) (Cho et al., 2014). The dimension of the BiGRU hidden layer was 128. We trained the BiGRU model for 50 epochs using the Adam optimizer.

### 5.3 Results and Discussion

Table 2 shows the experimental results. We evaluated the methods using F1, precision, and recall. As shown in Table 2, the BiGRU + char-CNN model with the ingredient frequency achieved the highest F1 score of 93.3. The BiGRU-based model was better than the random forest, so this result suggests that a sequence labeling approach is effective for the non-ingredient detection task.

The ingredient frequency improved the F1 scores for both the random forest and BiGRU models. Table 3 shows the most frequent ingredient names, which were calculated from approximately 3 million recipes from Cookpad. As shown in Table 3, many ingredient names that occurred frequently in recipes were actual ingredients, so ingredient name frequency is important for ingredient detection. The ingredient frequency can be an alternative feature of an ingredient dictionary which is usually rarely available.

In most cases, an item without a quantity in the

ingredient list was used as a heading or comment, as described in Section 3. When we predicted an item without a quantity as a non-ingredient, the F1 score was 85.3 in our further investigation. However, some items such as vegetables or fruits, had no quantities although they were actually ingredients. Using the ingredient name frequency, it became possible to predict such items as ingredients properly because names of vegetables or fruits frequently occurred in our recipe dataset.

## 6 Conclusion

In this paper, we introduced a non-ingredient detection task for user-generated recipes and proposed a neural model based on the sequence tagging approach. We used a BiGRU-based model to predict a label for each ingredient over an ingredient sequence. To evaluate our method, we constructed a dataset that contained 6,675 ingredients of 600 recipes from Cookpad. Our experimental results showed that the proposed method achieved a 93.3 F1 score in the task. In future work, we plan to verify the effectiveness of our method for downstream tasks, such as calorie estimation.

## References

- Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)*, pages 1724–1734.
- Young-joo Chung. 2012. Finding Food Entity Relationships using User-generated Data in Recipe Service. In *Proceedings of the 21th ACM International on Conference on Information and Knowledge Management (CIKM 2012)*, pages 2611–2614.
- Jun Harashima, Michiaki Ariga, Kenta Murata, and Masayuki Ioki. 2016. A Large-scale Recipe and Meal Data Collection as Infrastructure for Food Research. In *Proceedings of the 10th International*

- Conference on Language Resources and Evaluation (LREC 2016)*, pages 2455–2459.
- Jun Harashima, Makoto Hiramatsu, and Satoshi Snajo. 2020. Calorie Estimation in a Real-World Recipe Service. In *Proceedings of the 32th Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-20)*.
- Jun Harashima and Yoshiaki Yamada. 2018. Two-Step Validation in Character-based Ingredient Normalization. In *Proceedings of the 10th Workshop on Multimedia for Cooking and Eating Activities (CEA 2018)*, pages 29–32.
- Makoto Hiramatsu, Kei Wakabayashi, and Jun Harashima. 2019. Named Entity Recognition by Character-based Word Classification using a Domain Specific Dictionary. In *Proceedings of the 20th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing 2019)*.
- Shintaro Inuzuka, Takahiko Ito, and Jun Harashima. 2018. Step or Not: Discriminator for The Real Instructions in User-generated Recipes. In *Proceedings of the 4th Workshop on Noisy User-generated Text (W-NUT 2018)*, page 214.
- Jermisak Jermisurawong and Nizar Habash. 2015. Predicting the Structure of Cooking Recipes. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*, pages 781–786.
- Chloé Kiddon, Ganesa Thandavam Ponnuraj, Luke Zettlemoyer, and Yejin Choi. 2015. Mise en Place: Unsupervised Interpretation of Instructional Recipes. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*, pages 982–992.
- Chloé Kiddon, Luke Zettlemoyer, and Yejin Choi. 2016. Globally Coherent Text Generation with Neural Checklist Models. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP 2016)*, pages 329–339.
- Shinsuke Mori, Hirokuni Maeta, Yoko Yamakata, and Tetsuro Sasada. 2014. Flow Graph Corpus from Recipe Texts. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*, pages 2370–2377.
- Akiyo Nadamoto, Shunsuke Hanai, and Hidetsugu Nanba. 2016. Clustering for Similar Recipes in User-generated Recipe Sites based on Main Ingredients and Main Seasoning. In *Proceedings of the 19th International Conference on Network-Based Information Systems (NBIS-2016)*, pages 336–341.
- Amaia Salvador, Michal Drozdal, Xavier Giro i Nieto, and Adriana Romero. 2019. Inverse Cooking: Recipe Generation from Food Images. In *Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2019)*.
- Amaia Salvador, Nicholas Hynes, Yusuf Aytar, Javier Marin, Ferda Ofli, Ingmar Weber, and Antonio Torralba. 2017. Learning Cross-modal Embeddings for Cooking Recipes and Food Images. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*, pages 3020–3028.
- Tetsuro Sasada, Shinsuke Mori, Tatsuya Kawahara, and Yoko Yamakata. 2015. Named Entity Recognizer Trainable from Partially Annotated Data. In *Proceedings of the 14th International Conference of the Pacific Association for Computational Linguistics (PACLING 2015)*, pages 148–160.
- Semih Yagcioglu, Aykut Erdem, Erkut Erdem, and Nazli Ikizler-Cinbis. 2018. RecipeQA: A Challenge Dataset for Multimodal Comprehension of Cooking Recipes. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP 2018)*, pages 1358–1368.
- Ye Zhang and Byron C. Wallace. 2017. A Sensitivity Analysis of (and Practitioners’ Guide to) Convolutional Neural Networks for Sentence Classification. In *Proceedings of the 8th International Joint Conference on Natural Language Processing (IJCNLP 2017)*, pages 253–263.