Column Editor’s note: Most people these days routinely use the internet to look up recipes. Data scientists Shuyang Li and Julian McAuley offer food for thought in this issue’s Recreation in Randomness that the availability of recipes online is small potatoes compared to recent developments in searching, reconstructing and personalizing recipes to improve the experience of cooking in one's home.

Keywords: cooking recipes, recipe writing, food recommendation, home cooking, assistive cooking

Over 1.8 million years of human history, cooking has played an integral role in the development of society and civilization (Wrangham, 2009). During months of quarantine, millions found themselves returning to their home kitchens—often for the first time in years—to provide their daily meals. An April survey indicated that nearly half of American adults have started cooking at home more frequently (Oaklander, 2020). As a result, online recipe websites such as Allrecipes and Tasty have seen a surge of traffic (Lundstrom, 2020). In light of global food/ingredient shortages and price fluctuations (Dahir, 2020), it has become all the more important to help those at home cook healthy dishes given the ingredients at hand. For some, cooking is a welcome distraction from the outside world; for others, it is a necessity of daily life.

As much as traditional print media (e.g. Bon Appetit, Cook’s Illustrated) have extolled the virtues of culinary training and expertise in the home, the majority of home cooking revolves around simple, accessible recipes that can be found on online recipe aggregators (Hune-Brown, 2016). As simple as these recipes may be, they still pose significant challenges for new home cooks. These cooks often have little to no experience in the kitchen (Krishna, 2020), lacking the intuition born from expertise. They are also frequently limited by their pantries, ingredient availability, and lack of kitchen appliances.

Despite the scale of recipe websites, they are not exhaustive, and it takes human effort to catalog and upload recipes resulting in many different varieties of the same dish involving different ingredients. Early efforts to programmatically assist cooks in recipe discovery, preparation, and meal planning date back to the early rise of computational efficiency and accessibility (Hammond 1986). While these assistive technologies remain relevant today, they have yet to achieve broad applicability and adoption by the public. Recent research and techniques in data science and natural language processing (NLP) have shown promise in tackling this myriad of challenges, from ways to suggest and retrieve healthy recipes to generating brand new recipes from scratch given the ingredients in one’s pantry and catering to one’s taste.

Searching for Healthy and Nutritious Recipes

How can we help home cooks find healthy dishes without needing to search through and scrutinize millions of recipes online?
Even when presented with a list of ingredients, it can be difficult for humans to identify the fat, sugar, and/or caloric content of a recipe without memorizing a table of nutritional facts (Elsweiler et al., 2017). Automated tools, however, can easily access nutritional facts and health assessments for recipes at a moment's notice. Ueta et al. (2011) built a system to help users search for recipes that target specific health conditions. The authors measured the co-occurrence between nutrients and medical keywords (e.g. ‘bone,’ ‘acne’), which were then paired with a table of nutritional values for cooking ingredients. When a user searches for a broad class of recipes using colloquial language (e.g. ‘I want to recover from fatigue’), the system ranks recipes based on their nutritional values given the keywords found in the query and returns the top results.

While certain nutrients may be associated with specific health conditions, simply recommending recipes that contain large numbers of specialized ‘healthy’ ingredients may not be suitable for most home cooks. Inagawa et al. (2013) proposed an alternative, assistive system for customized healthy recipe recommendation. Rather than relying on a general user query, their model used nutritional constraints. These rules could be dietitian-specified or provided by the user via common diet and nutrition applications (e.g. MyFitnessPal). Their system recommends a small set of dishes that satisfy those constraints, from which the cook can efficiently pick their favorite. Yang et al. (2017) created a recipe recommender system with a similar mode of input: they sought to infer personal taste and nutritional preferences from a short survey presented to the user. The system learned a policy derived from images presented to a user to quickly narrow down their preferences.

Ahn et al. (2011) approached the problem of food pairing from a chemical angle, creating a flavor network consisting of ingredients and the chemical flavor compounds they contain. The authors then used this flavor network to analyze ingredient substitutability and complementarity in a set of 50,000 recipes from online aggregators. They observed a systematic difference between ingredient combinations used in different cuisines across the globe. For example, primary ingredients in East Asian cuisine—garlic, scallions, soy sauce—contain disjoint sets of flavor compounds; meanwhile, in North American and West European cuisine, primary ingredients including milk, eggs, and butter all tend to share flavor compounds. This study also highlights the limits of existing recipe aggregators and datasets, especially with regards to cuisine and geographic diversity.

Indeed, regional preferences dictate not only the types of dishes found on recipe websites but also the types of food images uploaded to these sites. In a study of recipe aggregators from Germany, China, and the United States, Zhang et al. (2019) found that food photographs preferred by users of each platform are correlated with different sets of visual features (e.g., brightness, color palette, and framing). While regional and cultural cuisines use ingredients and flavors in different ways, recipe aggregators tend to be dominated by dishes whose look and taste are dictated by their primary demographic.
Finding Recipes in the Wild

Can machine learning help us reconstruct the recipes behind our favorite food photos?

The techniques mentioned above can help users find healthier recipes while searching recipe databases. While traditional search focuses on text, people increasingly interact with food media via images and photographs. Alongside the proliferation of ‘foodies culture’ and the casual gourmand, people are posting large volumes of food photos on social media. This includes a wealth of inspiration for home cooks—as of June 2020, over 13 million photos on Instagram have been uploaded under the tag #homecooking. Recent work on cross-modal understanding of images and text has inspired work toward image-based search systems for recipe retrieval (Marin et al., 2019).

This is a natural direction toward more powerful and expressive ways for home cooks to find recipes. Elsweiler et al. (2017) studied 218 adults and found that their recipe preferences were informed by and correlated well with visual properties of food photos, including brightness and color palettes. This suggests that cooks could better express their preferences and the food they are trying to make via images. A recent line of research in image-to-recipe retrieval (Salvador et al., 2017) has shown the promise of automated systems for accurately identifying recipes from pictures of finished dishes. The authors showed that one such machine learning system learned to represent images and recipe texts in a semantically aligned way. Essentially, this means that subtracting the representation for the word ‘cake’ from a picture of a chocolate cake and then adding the representation for ‘cupcake’ results in a representation close to that of a picture of ‘chocolate cupcake.’

Salvador et al. (2017) also showed that their model could identify recipes from images more accurately than crowdsourced human workers from the Amazon Mechanical Turk platform, particularly for dishes with easily-recognizable component ingredients. We can see that by introducing an automated system to identify recipes from images, we may help cooks search for recipes much more efficiently, in a more comfortable modality. Myers et al. (2015) showed that computer vision models can also infer the caloric content of photographed dishes, which means that our hypothetical image-based recipe search engine could also allow us to find lower-calorie and healthier recipes without needing additional inputs.

These methods can link a picture to a likely recipe from a database, but they cannot reconstruct an arbitrary dish from a picture. Such a system would allow home cooks to make a much wider variety of foods and not be limited to already-uploaded recipes—one could recreate one’s favorite restaurant or fast-food dish without complicated experimenting and reverse-engineering. Two recent papers suggest that this might not be very far off. Chen et al. (2017) pioneered a machine learning system that learns to scrutinize small sections of a food photo to identify which ingredients and cooking techniques have likely been used. The authors demonstrated that their system could distinguish between sets of
recipes that use the same ingredients but very different cooking and preparation methods (e.g. shredding vs. chopping, grilling vs. boiling). Such a model can also explain which parts of an image lead it to predict certain ingredients or cooking techniques, much as a human would.

We can take this idea and go one step further: writing the complete recipe with ingredients, techniques, and ordered steps. Salvador et al. (2019) explored a system for generating these full recipes from images by first predicting the constituent ingredients and using those to infer cooking instructions. The authors showed that given the same image, human raters consistently preferred recipes returned by their model compared to similar recipes retrieved from existing recipe databases. Key to this system is treating the recipe inference task as language modeling, or predicting the next word of a sentence (or recipe step). This allows a model to write a recipe that more accurately represents the dish being photographed than any existing recipe on recipe websites.

Since we can parse food photographs, a natural extension would be building a similar type of understanding around videos. Cooking videos are increasingly popular among young adults (Delgado et al., 2014), but home cooks can be put off by perceived complexity and the audio-visual format. Oftentimes these videos are not accompanied by written recipes, which can make it hard for home cooks to reproduce these dishes. Fujii et al. (2020) explored the idea of reconstructing written recipes via captioning videos with the cooking instructions being performed at any given point. While such systems are still not ready for public consumption, this research represents an encouraging step in the direction of democratizing cooking for the average consumer.

**Personalized Recipes**

Can new recipes be created in an automated manner to satisfy individual palates?

So far, we have focused on retrieving appropriate recipes accurately described via a search query or image. But cooks also decide what to cook based on their individual preferences; this concept helps explain why so many different forms of barbeque developed around the world—each with varying levels of sweetness, spice, and acid from very different ingredient combinations. In addition to regional differences, recipe popularity, appeal, and staying power depend on a host of other human factors such as gender and sensitivity to seasonal trends (Kusmierczyk et al. 2015, Rokicki et al. 2016).

For a system to help home cooks consistently, it must provide personalized recipe suggestions tailored to the user’s palate and preferences.

This type of personalized recipe suggestion falls squarely into the realm of Recommender Systems which learn to predict user preferences and rank recipes accordingly. Harvey et al. (2013) used user ratings collected from recipe websites to predict how a user would rate a recipe they have never seen before. The system could then compare the predicted ratings for a set of search results to re-rank them in order of enjoyment for a specific user. The authors also explored adding nutrition information
to the recommender system, allowing it to learn to what extent a given user prioritizes health and nutritional value when choosing dishes to cook. Ueda et al. (2014) presented a similar system for extracting previously consumed recipes from users’ browsing and cooking histories to learn which ingredient combinations a user prefers.

These recommender systems can learn specific user preferences, but they restrict user agency by presenting a set of suggestions based solely on a cook’s historical activity. We thus seek a system where the user can specify some constraints—e.g. ingredients they have at hand—and receive an appropriate and personalized suggestion. Research into such systems falls in the field of recipe generation: creating a brand-new set of ingredients and cooking instructions. In discussing the work of Salvador et al. (2019) to reconstruct recipes from images, we have already considered the benefits of such an abstractive system for accurately representing recipes that may not exist verbatim in existing websites.

Writing a recipe comes with many challenges—recipes consist of real-world ingredients that must be combined in a specific, ordered manner to create the final dish. Most cutting-edge text generation models tend to be opaque, giving users little control over the text being produced. However, a good recipe generation model must produce coherent recipes where the instructional steps can be followed by a home cook. Several recent papers point to methods for improving the coherence of generated recipes.

These papers assume that a user has provided a rough description of their desired recipe as well as the desired ingredients. Kiddon et al. (2016) proposed a mechanism for keeping track of which of the specified ingredients have been used already and which ones must be used in following instructions. They based their method on the idea of tracking ingredient usage with checklists. The authors provided human evaluators with sample recipes from their model and found that their method created recipes that were more fluent, grammatical, and appropriate given the user query and provided ingredients when compared against extant text generation models.

Bosselut et al. (2018) approached the task from a different angle: since recipes are procedural instruction sets, the authors treated each step as a series of cooking actions acting on ingredients. By learning how each action changes the physical qualities of an ingredient, their model tracked the desired physical state of provided ingredients and returned the corresponding English-language instruction. The authors posited that while existing generative techniques are unable to learn the ‘common sense’ ways in which complex items like ingredients are used and interact with one another in recipe instructions (Levy et al., 2015), their techniques allow the model to explicitly learn these relationships.

More recently, Majumder et al. (2019) sought to combine the above work from recommender systems and text generation. In addition to being provided ingredients, their model also learned user
preferences from sequences of previously consumed dishes (tracked via social media acknowledgments and 'favorites'). The authors showed that their system can leverage these preferences to generate different recipes for different users, even when presented with the same ingredients. They also explored the idea of relaxing user constraints: by allowing a user to only specify some primary ingredients (e.g. chicken or pasta), the model could create a recipe with customized herbs, flavorings, and other accoutrements.

But no matter how good recipe generation algorithms become, they depend heavily on the quality and diversity of the openly accessible recipes uploaded to recipe websites. This problem is accentuated by the fact that large online recipe collections and websites tend to be dominated by recipes from a few major cuisines (Ahn et al., 2011). For example, while Mongolian cuisine contains a rich variety of protein-based and noodle soup dishes, Allrecipes only contains 14 recipes marked as Mongolian—consisting solely of varieties of the American-Chinese dish 'Mongolian Beef,' with no relation to Mongolian food. A user searching for lamb or noodle recipes would never be exposed to a large variety of simple yet delicious recipes from under-represented cuisines. Fortunately, recent work in cross-lingual understanding (Conneau and Lample, 2019) has opened the door for ways to learn recipes from multi-lingual collections. Many cultures have a rich written tradition for passing down recipes in their language, and such models could learn from these collections to drastically increase the diversity of generated or suggested recipes.

Conclusion

We have only begun exploring the myriad of ways that cutting-edge data science and machine learning techniques can help us cook and stay healthy. Existing models can track the nutritional value of recipes to nudge cooks toward healthier recipes. They can help recreate recipes from one’s favorite restaurant, blogger, or food photographer. Such systems can also encourage people to cook more often by suggesting recipes that fit their tastes.

Ultimately, these lines of research converge to make cooking techniques, ingredients, and diverse cuisines more approachable for the home cooks of today. Their methods are used in several consumer-facing applications today, including tools to analyze the structure of recipes and explore recipes in a graphical interface (Chang et al., 2018), as well as online recipe generation sites that allow users to specify ingredients and a recipe name and receive a set of cooking instructions (Lee et al., 2020).

The impact of such research can extend beyond software technologies. Food anthropologists and historians could benefit from computational methods to analyze trends in ingredient usage, cooking techniques, and consumption not only for contemporaneous populations but also historical societies. Research in personalization and preference elicitation can inform public health policy, with more effective dietary guidelines and potentially meal planning to improve health in schools and hospitals.
Commercial applications span personalized ready-to-eat meals or individualized versions of existing meal kits (e.g., HelloFresh, BlueApron).

Some researchers have proposed even more futuristic ways to integrate cooking technology into daily life: Hamada et al. (2005) and Neumann et al. (2017) developed multimedia systems to assist cooks with ingredient measurement and parallel workflows during cooking, and Mizrahi et al. (2016) discussed the application of digital instruments such as laser cutters and 3D fabrication devices to allow home cooks even more fine-grained control over flavors and textures. But even if such futuristic techniques become commonplace, it is unlikely that humanity would abandon the institution of home cooking. Instead, we can anticipate the development of assistive systems to make healthy, delicious home cooking simpler and more accessible for the general public.

Disclosure Statement

The authors have nothing to disclose for this article.

Acknowledgements

This work was inspired in part by research in recipe generation done by the authors at UC San Diego alongside Bodhisattwa Majumder, Jianmo Ni, and Yufei Li.

References


This article is © 2020 by Shuyang Li and Julian McAuley. The article is licensed under a Creative Commons Attribution (CC BY 4.0) International license (https://creativecommons.org/licenses/by/4.0/legalcode), except where otherwise indicated with respect to particular material included in the article. The article should be attributed to the authors identified above.