

## **RECIPE4U: An Android Application Using Web Scraping**

Ahmad Farid Najmuddin, Siti Nur Maisarah Bahaman, Ireen Munira Ibrahim, Siti Salihah Shaffie, Siti Rozanae Ismail and Anisah Abdul Rahman

Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA, Perak Branch, Tapah Campus, Tapah Road, 35400 Perak, Malaysia

**To Link this Article:** <http://dx.doi.org/10.6007/IJARBS/v10-i9/7988> DOI:10.6007/IJARBS/v10-i9/7988

**Published Date:** 29 September 2020

### **Abstract**

Cooking is a daily activity that some people do. To do so, some people might need to refer to recipes. However, it is sometimes difficult and time consuming to find the right recipe, especially local Malaysian dishes. Although there are solutions such as apps to quickly search for recipes were developed, existing apps and solutions seem to lack the feature to suggest recipes to the user based on their preferred ingredients. This project aims to develop a new mobile recipe application, called Recipe4U using web scraping framework where this application is expected to suggest recipes based on the user's preferred ingredients. The app will include Malaysian dishes recipe with both English and Malay language selection. The project implemented regular expression to search the ingredients based on the keywords entered by the user. This project uses Android Mobile Application with Java language and a web scraping application scripts in R language and SQLite database for recommender function. The user acceptance test results proved that this project is well accepted by the end users. A survey was conducted on app's acceptance test with 83 respondents, to obtain their feedback on the app's functionality and usability. Based on the survey, 89% of the respondents are very satisfied with the overall function of the app. For future work, more sources of the recipes will be added, recipe sharing, and bookmark feature will be implemented to improve the usability of the app.

**Keywords:** Mobile Application, Recipe Suggestion, Android Mobile, Web Scraping, Recommender System.

### **Introduction**

Cooking is a basic thing to do in daily life. Preparing menus may require people to search and think for suitable recipe according to their preferences or maybe they are searching for a new recipe to try. Recipe made cooking easier as it provides all necessary steps and ingredients to cook the foods as intended (Linda, 2019). Traditionally, recipes can be obtained through

printed medium such as cooking magazines or books. As the technologies improved day by day, the technologies such as the Internet and smartphone give people an easy way to browse and access information at their fingertips (Nath & Mukherjee, 2015). This information such as cooking recipes, can be found easily through the internet. Smartphone has become a great medium for finding recipe on the go.

According to (Yanai et al., 2014), cooking recipe websites, blogs and social media such as Facebook and Instagram, have become common nowadays since user can access them using their smartphones, hence allowing them to access information anywhere and anytime. There are many known websites and mobile applications currently available for free to find recipe such as BBC Good Food, Tasty, Big Oven, Cookpad and more, which provide recipe suggestion, steps to cook (with media instructions), list of ingredients needed and users are able to share their favorite recipe or their own recipe with others too. Despite of the plenty recipes provided, most of the apps or webs cannot provide or recommend recipes according to the user preferred ingredients. Besides, the apps and web sites as mentioned only focusing on Western and some Asian cuisine recipes. Hence, the focus of the study is to provide users with more familiar local Malaysian dishes and ingredients.

The objective of this study is to develop a new mobile recipe application using web scrapping framework where this application is expected to suggest recipes based on the user's preferred ingredients and focused on Malaysian dishes recipe with English and Malay language selection option. This project will help users to get an ideal recipe based on their ingredient availability and save their time searching for the right recipe.

The remainder of this paper will cover on the review of the previous mobile recipe application and techniques for web scrapping. Section 3 elaborates the development of mobile recipe application. Section 4 discusses the results obtained and the analysis done. Finally, Section 5 concludes the paper with the future research.

## **Literature Review on System Development**

### **Web Scraping**

Web scraping is also known as data mining and web harvesting. It is a method to extract and fetch data from websites. Glez-Pena et al. (2013) also defines web scraping as procedure of extract and combine contents from Web in a systematic way. Web scraping contains of variations of programming technique and technologies such as data analysis and information security (Mitchell, 2015). There are numerous methods and algorithms available for performing web scraping such as manual scrapping, website scraping API and regular expression.

Regular expression has been chosen for this project because it can choose more than one element and it can infer appropriate regular expression to recognize that element in the page while manual method is costly and website scraping API have limited amount of output document formats, and also have deep learning curve. Regular expression can recognize strings or patterns in unstructured text (Ferrara et al., 2014) and writing a regular expression on HTML pages have several criteria: word boundaries, HTML tags, tables structure, etc. According to (Goyvaerts, 2007), regular expression requires an "engine" which is a piece of software that can process the regular expressions syntax. This engine will try to match the regex pattern to the given string and usually the engine is part of a larger application where user do not access the engine directly, but the application will access it when needed and making sure that the right regular expression is used for the right file or data.

### **Recommendation System**

As technology growth over the years, recommendation system has been widely used in businesses and social media such as Facebook. According to (Li et al., 2018), recommendation system is an information filtering technology where the user can search a significant information. It helps the user to search for information accurately based on user preferences. Recommendation systems manage excess numbers of data for user by providing them with personalize, exclusive content and service recommendation (Isinkaye et al., 2015).

There are two recommender algorithms considered in this study; Content Based Filtering Algorithm (CBFA) and Collaborative Filtering Algorithm (CFA), and hybridization of those two algorithms is also taken into consideration. CBFA will recommend items based on the relationship between content of the items and user profile (Karbhari et al., 2017). Meanwhile CFA will recommend items based on what multiple users preferred (Kembellec et al., 2014). While hybrid filtering algorithm is a combination of the advantages of both algorithms. All this algorithm has their own pros and cons, however after reviewing pros and cons among those, CFA is chosen as it is suitable algorithm for this project. It is easy to implement this algorithm and capable to improve the prediction of the performance.

### **Collaborative Filtering Algorithm**

According to (Isinkaye et al., 2015), collaborative filtering is a domain-independent prediction technique for content that cannot easily and adequately be described by metadata such as movies and music. Collaborative filtering technique or sometimes called a social filtering, is a method that suggest information to users based on other user's preferences. Based (Sharma et al., 2016) collaborative technique used a neighborhood concept which the contents are recommended based on other user's preferences that have similar taste with the users. Collaborative technique is classified into two categories: neighborhood based, and model based. Neighborhood based method or sometimes called memory-based algorithm will choose the data that have been choose by the similarity with the user. It is a process to predict ratings by referring to the user's which ratings are similar to the queried user (Mustafa et al., 2017). While model-based technique makes a predictive model based on the other user ratings. Reason for this method is to copy and model the interactions of user item with factors representing the hidden characteristics of the users and the items in the system (Shah et al., 2017).

### **Mobile Application Approach**

Mobile application can be defined as a software application that run on a smartphone's operating system. Currently mobile application is high on demand because of their popularity among end users (Ahmad et al., 2017) There are three common ways to develop the application: web approach, native approach and hybrid approach. Based on the reviewed literature, the approach that was chosen for the project was Native Mobile Application (NMA) since it has a better experience and run faster than Web Mobile Application and Hybrid Mobile Application.

### **Native Mobile Application**

NMA is an application that develop on specific platforms and are written using languages that the platform accepts such as Objective-C for iOS and Java for android. NMA are fast and flexible and has the best performance. There are two main mobile operating system platforms

for native application: iOS and Android. There are several advantage and disadvantage using native mobile application. Table 1 shows the advantages and disadvantages of the NMA. For this project, android platform is chosen because it is preferable compared to iOS for apps development.

Table 1

*Advantages and disadvantages of NMA*

Advantages	Disadvantages
Distributed in app stores and play stores	Costly (not the best option for simple apps)
More interactive and run much smoother	Difficult language to learn
Better user experience	

**Mobile Application Development****Project Architecture**

The system architecture was divided into three tiers: User Application tier, Web Server tier, and Database Server tier as shown in Figure 1. The User Application tier is the Android mobile application, where the users used this application for searching the recipe suggestions. The second tier is the Web Server tier, where the main process was done. The Web Server tier received the searched keywords from the mobile application, and the web server performed web scraping from the internet, in order to collect the data.

The data that have been collected were cleaned and filtered. The scraped data were the list of the recipe that matched the keywords provided by the user in the search. The list will be sent back to the mobile application and displayed to user. The searched ingredients keywords were also sent to the Database tiers where the database stores the searched ingredients, which were used for displaying on the ingredients page.

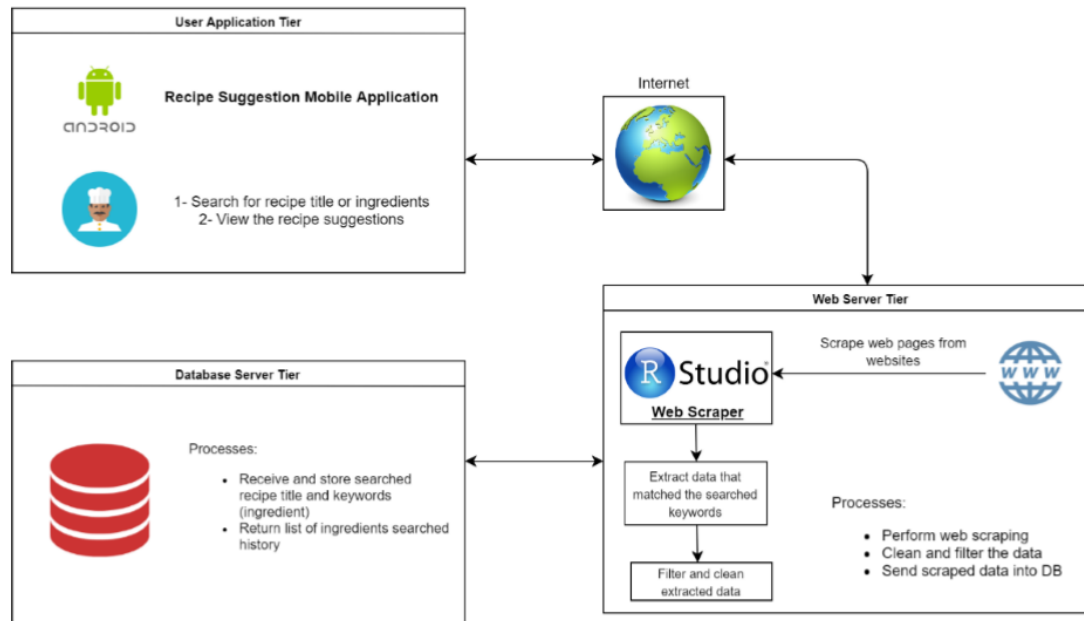


Figure 1. Project Architecture Diagram

### Mobile Application Flow

The first process is getting the input from user as shown in Figure 2(a). User will enter the preferred ingredients or recipe into the text boxes as the keywords. Then the keywords will be sent to the website to search for matching keywords. Then it will copy the whole website search result pages and the source code. Next, the websites HTML codes will be extracted to get only recipe name, image, description and link. Then it will be cleaned to get the raw text. Lastly, it will combine and structure the texts and displays the list of recipes as shown in Figure 2(b). To scrape the recipes from the websites, R language was used and the coding is done with RStudio. Before scraping the recipes, two packages from R were installed which were *stringr* and *rvest*. Before displaying the search result, the program will check the accuracy point of the recipe, if it is more than 0, which means at least 1 keyword matched the title or description, then the recipe will be displayed. On the recipe lists, the missing keywords will also be displayed together.

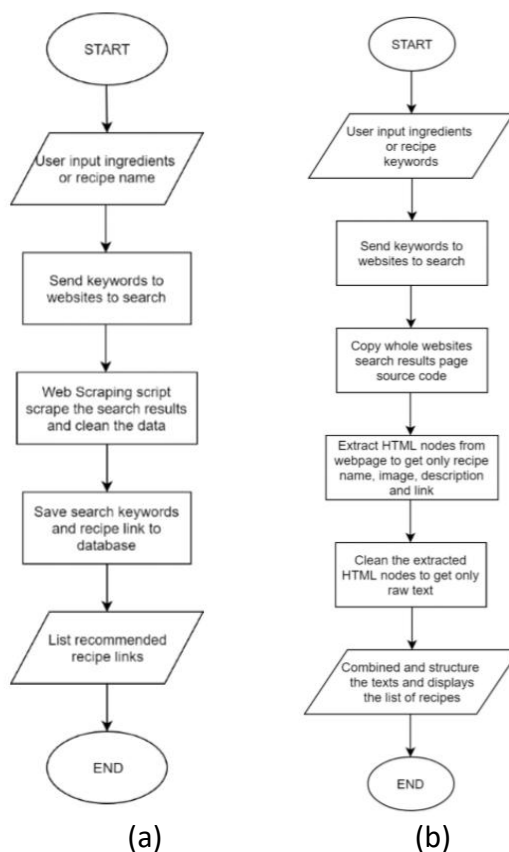


Figure 2. (a) Recipe Searching Flow (b) Web Scraping

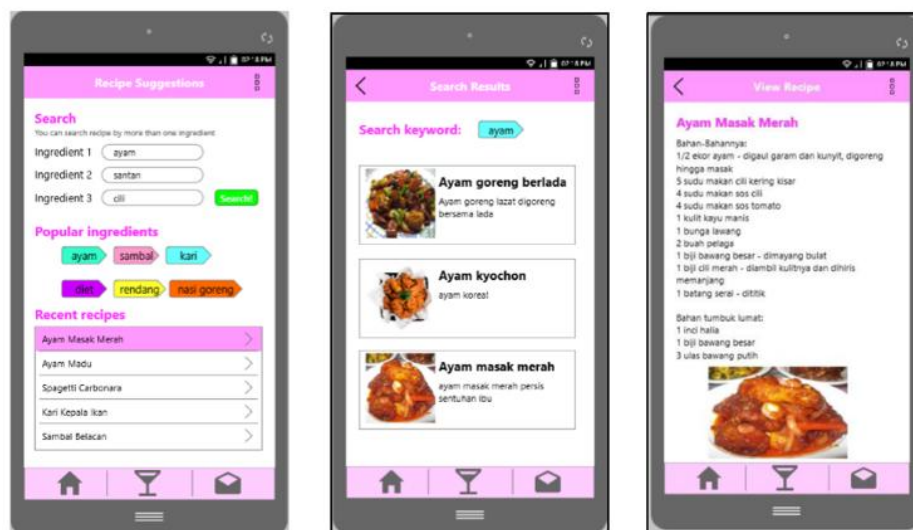


Figure 3. User Interface for Recipe Searching

### Implementation and Testing

There are two testing were conducted during the development of the mobile app: Functionality Testing and Usability Testing. The functionality testing was conducted to confirm that the recipe app functioning well and runs as intended. Based on the result obtained, all functions in the mobile app operated as required after a number of

modifications. Meanwhile, Usability Testing was performed to analyse end users satisfaction towards the app. A survey was conducted with 89 respondents participated in this test. The test consists of four sections. The sections are Section A – App Installation, Section B - User Interface Design, Section C – App Usability and Section D - Overall App Development.

### App Installation

In this section, respondents were asked either the installation of the app is succesful or otherwise. Based on the survey, 93% of the respondents (83 respondents) are able to install and launch the app as shown in Figure 4. Only 7% of them (6 respondents) were unable to install due to incompatibility of their smartphone with the app (they are using Android 4.04 and Android 6.0.1 versions) and some of them are unsure what happened since they have not received any error message.

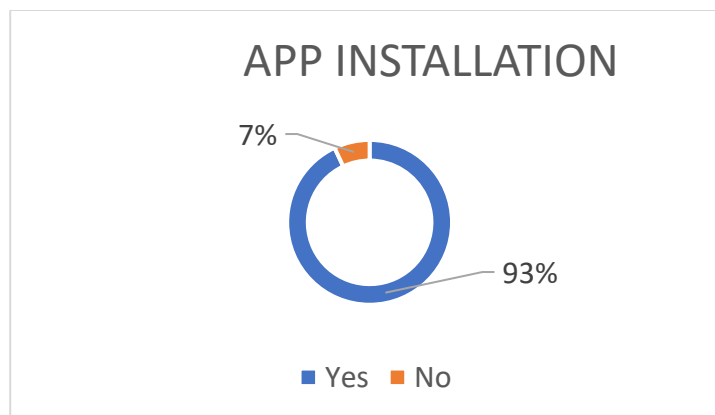


Figure 4. Percentage of respondents for App Installation

### User Interface Design

In this section, the testing was conducted to assess user's satisfaction towards the interface design of the app based on: suitable interface (font and color), user friendly and understandability, and common function buttons. The results were shown in Figure 5. Based on the results, most of the respondents are very satisfied with the criteria evaluated. 71 % of the respondents are very satisfied with the interface, 80% are very satisfied that the app is user friendly and understandable. 80% of them are very satisfied with all functions button in the app. There are no % of responses under disagree and strongly disagree scale.

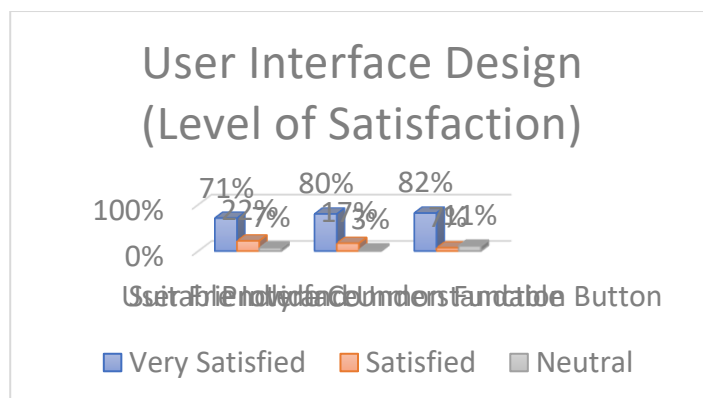


Figure 5. Percentage of user’s level of satisfaction towards User Interface Design

**App Usability**

In this section, the testing covered how well the app met user requirements based on: accessibility (time and place), helps in finding Malaysian’s recipe, searching time, and recommended suitable recipe. Based on the results as shown in Figure 6, 90% of the respondents are very satisfied with the app as it helps them finding Malaysian’s recipes. Meanwhile, 85% of them are very satisfied that this app recommends suitable recipe based on their preferred ingredients. Beside, 83% are ver satisfied that this app reduced their recipe searching time because this app can be accessed anywhere anytime.

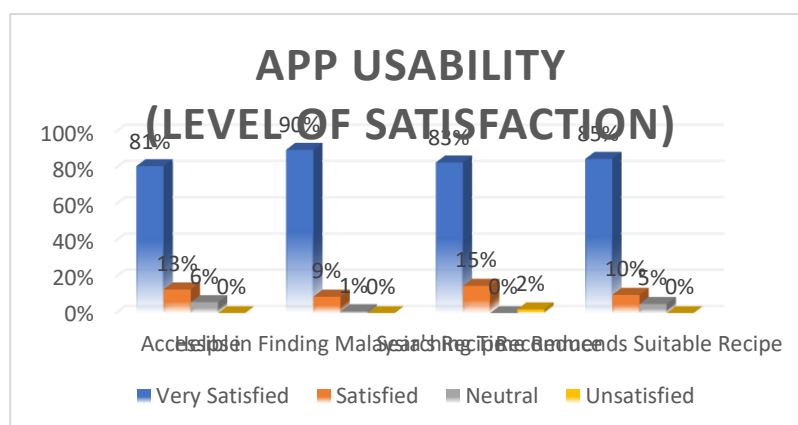


Figure 6. Percentage of user’s level of satisfaction towards App Usability

**Overall App Development**

The last section assess the overall user’s level of satisfaction towards the implementation of the app. As shown in Figure 7, 89% of the respondents very satisfied with the development and implementation of the app in helping them to find the suitable recipe based on their preferences. User are also being asked for their suggestions and opinions towards the app, which hopefully can help giving ideas for future works or improvements.



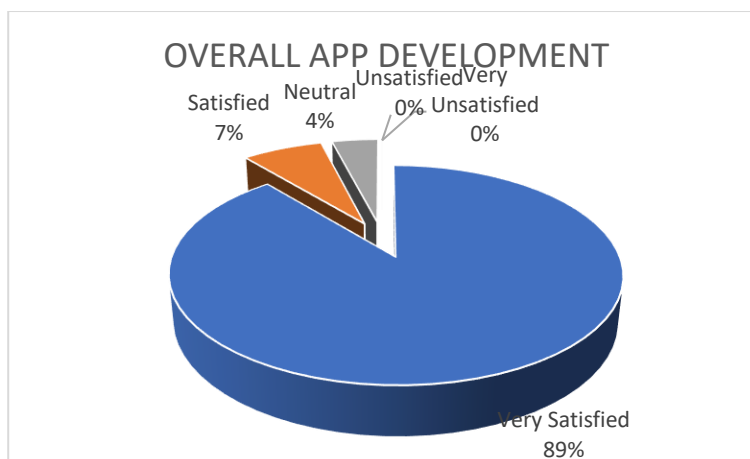


Figure 7. Percentage of user's level of satisfaction towards Overall App Development

### Conclusion

Recipe4U app consists of two components, the Web Scraping component, which do all the processes of the application, and the Android mobile application which serves as the container for running and using the project on Android smartphones. Recipe4U was developed to help people to find Malaysian dishes recipe easily at their fingertips, in one place. It is aimed to solve the problem of hassle to open multiple websites just to find a recipe. This will help user to save their time to search for the right recipe. Furthermore, another key feature of Recipe4U is the ability to recommend popular ingredients and recipes to the user. This helps user to find out what kind of foods other users usually search for or what recipe they need with their available ingredients.

Even though the project has been successfully developed and all the objectives were achieved, there are some limitations for Recipe4U. The first suggestion is the lack of recipe search results list pagination. Currently, user need to scroll down a long list of the recipes list in the app. This can become troublesome if the list is too long, and even can make the app becomes overload. This will affect the user experience while using the app, which also affect the usage of the application. Moreover, the next limitation is the lack of save or bookmark recipe feature. This feature is useful for user to save their search results. As example, if a user search for a recipe that might be decided for later use, the user can save or bookmark the recipe first, so the user does not have to search for the recipe again. Currently, if the user wants to use the recipe again, he or she need to perform the search again. This feature will help to save time and improve satisfaction.

Once the Usability Testing results were evaluated and the limitations of the project has been identified, the next step is to propose some solutions or recommendations for implementation in the next version. The first recommendation is the app will add more website sources for scraping the recipes including social media websites such as Instagram and Facebook. This is because social media nowadays also become one of the popular platforms for people to share their recipes as social media is accessible by anyone. Then, for the second recommendation is to implement the pagination feature. The search result list can be limited to show only a few lists, and then user can expand the list by navigating to the next page, or just keep scrolling and new results will be auto populated. This feature will greatly

improve user experience of the app. Next is to implement the bookmark or save recipe feature. Since the project already have a database, then the database can also be used to store the list of recipes that user decided to save. This feature will also make it easy for the users to look for searched recipes in the history sections or bookmark section.

The final recommendation is to implement the sharing feature. The Recipe4U app can integrate with third-party social media platform for the sharing feature so that user can share their recipes directly to their own social media account. Furthermore, with the implementation of this feature can help to promote the app to other users, which will increase the success rate of the applications as more users are using the app.

### **Acknowledgment**

The authors would like to thank the Universiti Teknologi MARA Cawangan Perak and the Faculty of Computer and Mathematical Sciences for the supports and facilities provided in the completion of this project.

### **Corresponding Author**

Ahmad Farid Bin Najmuddin is a Computer Science lecturer at Universiti Teknologi MARA Perak Branch, Tapah Campus, Tapah Road Perak, Malaysia. Email: ahmad824@uitm.edu.my.

### **References**

- Ahmad, A., Feng, C., Ma, T., Yousif, A., & Shi, G. (2017). Challenges of mobile applications development: Initial results. *IEEE International Conference on Software Engineering and Service Sciences* (pp. 464–469). DOI: 10.1109/ICSESS.2017.8342956.
- Ferrara, E., De Meo, P., Fiumara, G. & Baumgartner, R. (2014). Web data extraction, applications and techniques: A survey. *Knowledge-Based Syst*, vol. 70, 301–323. DOI: 10.1016/j.knosys.2014.07.007.
- Glez-Pena, D., Lourenço, A., Lopez-Fernandez, H., Reboiro-Jato, M., & FdezRiverola, F. (2013). Web scraping technologies in an API world. *Brief in BioInfo*, 15(5), 788-797. DOI: 10.1093/bib/bbt026.
- Goyvaerts, J. (2007). *Regular Expression: The Complete Tutorial*, Lulu Press Inc. <http://www.princeton.edu/~mlovett/reference/Regular-Expressions>.
- Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. A. (2015). Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 16(3), 261-273. DOI: 10.1016/j.eij.2015.06.005.
- Karbhari, N., Deshmukh, A., & Shinde, V. D. (2017). Recommendation System using Content Filtering: A case study for college campus placement. *International Conference on Energy, Communication, Data Analytics and Soft Computing* (pp. 963–965). DOI: 10.1109/ICECDS.2017.8389579.
- Kembellec, G., Chartron, G., & Saleh, I. (2014). *Recommender Systems*, John Wiley & Sons. DOI: 10.1002/9781119054252.
- Li, Z., Hu, J., Shen, J., & Xu, Y. (2016). A Scalable Recipe Recommendation System for Mobile Application. *International Conference on Information Science and Control Engineering' 03* (pp. 91-94). DOI: 10.1109/ICISCE.2016.30.
- Linda, L. (2019). "How to Read a Recipe for Great Results". <https://www.thespruceeats.com/how-to-read-a-recipe-479944>.

- Mitchell, R. (2015). *Web Scraping with Python, Collecting Data from the Modern Web*. Sebastopol CA: O'Reilly Media, Inc.
- Mustafa, N., Ibrahim, A. O., Ahmed, A., & Abdullah, A. (2017). Collaborative filtering: Techniques and applications. *International Conference on Communication, Control, Computing and Electronics Engineering*. DOI: 10.1109/ICCCCEE.2017.7867668.
- Nath, A., & Mukherjee, S. (2015). Impact of Mobile Phone / Smartphone : A pilot study on positive and negative effects. *International Journal of Advance Research in Computer Science and Management Study*, 2(3), 294-302. <http://ijseas.com/volume2/v2i3/ijseas20160353.pdf>.
- Shah, K., Salunke, A., Dongare, S., & Antala, K. (2017). Recommender systems: An overview of different approaches to recommendations. *International Conference on Innovations in Information, Embedded and Communication Systems* (pp. 1–4). DOI: 10.1109/ICIIECS.2017.8276172.
- Sharma, S., Sharma, A. T., Sharma, Y., & Bhatia, M. (2016). Recommender system using hybrid approach. *IEEE International Conference on Computing, Communication and Automation* (pp. 219–223). DOI: 10.1109/CCAA.2016.7813722.
- Yanai, K., Maruyama, T., & Kawano, Y. (2014). A Cooking Recipe Recommendation System With Visual Recognition Of Food Ingredients. *International Journal of Interactive Mobile Technologies*, 8(2), 28-34. DOI: 10.3991/ijim.v8i2.3623.