

CookOverflow



Ahmad Droobi, Ataa Shaqour

Supervisor: Dr.Mona Demaidi
(Phd in AI)

Department of Engineering
An-Najah National University

This dissertation is submitted for the degree of
Computer Engineering

May 2022

I would like to dedicate this thesis to my loving parents ...

Declaration

I hereby declare that except where specific reference is made to the work of others, We would thank our supervisor Dr.Mona for her kind, support and our Computer Engineering Department for all the courses they offered us and not forget to thank our friends.

Ahmad Droobi, Ataa Shaqour

May 2022

Acknowledgements

And I would like to acknowledge future computing technologies lab at Clemson and the companies that offered us training to make this project done.

Abstract

Cook-Overflow is social media platform, specialized in food viewing, recipe sharing, interacting between users and using its data for a recipe recommender system, which we give it the recipe we have, and it gives us what to cook using machine learning approach, more over we built the platform infrastructure from zero that matches our requirements for the data collection and recipe recommendation.

Users can share their recipes, posting them on our platform. In addition, other users can interact with other users posts with using like and comment on the recipe posts, and even following other users.

the platform has friendly, easy to use UI, reliable and scalable infrastructure that make it very useful for lovers of food in general, and students who live away from their families to cook their own meals. And the users (lovers of food) could communicate with each other using simple Mobile application Messenger that makes it easier for them to share experiences and get the help from Professionals (chefs), guiding them to the best and easy ways to make their daily meals or Parties that they could invite their friends and share the best moments with the food they made with different Nationalities all over the world.



Fig. 1 components of Research Process

Table of contents

List of figures	x
1 introduction	1
1.1 Problems	1
1.2 Objectives	1
1.3 Scope	2
1.4 importance	2
2 Constraints, Standards/Codes and Earlier course work	3
2.1 Constrains	3
2.1.1 Dataset	3
2.1.2 Hardware Problems	4
2.1.3 Lack of similar abstractive approach in the architecture	4
2.2 Standards/Codes	4
2.2.1 Design and visual identity guidelines	4
2.2.2 Coding conviction	5
2.3 Earlier course work	5
2.3.1 Database	5
2.3.2 Web Development	5
2.3.3 Software Engineering	5
2.3.4 Artificial Intelligence	6
2.3.5 Advanced Software Engineering	6
2.3.6 Security	6
2.3.7 Mobile	6
2.3.8 Machine Learning	6
3 Literature Review	7
3.1 Data	7

3.2	Algorithm and Approach	8
3.3	Model	10
3.3.1	Knowledge Requirement for Modeling	10
3.4	Modelling the Architecture	11
4	Methodology	14
4.1	Languages and Frameworks	14
4.1.1	Back-End-Framework	14
4.1.2	Back-End-Security	15
4.1.3	Back-End-URL Structure	15
4.1.4	Front-End	16
4.2	UML Diagrams	16
4.2.1	Use-Case Diagram	16
4.2.2	Class Diagram	17
4.2.3	sequence Diagram	18
4.2.4	State Diagram	19
4.2.5	Activity Diagram	20
4.3	Dataset	21
4.4	Text Processing(Spacy) and NLP	21
4.5	Search Engine (Query Algo)	21
4.6	TimeLine	21
4.6.1	Week 0	21
4.6.2	Week 1	22
4.6.3	Week 2	22
4.6.4	Week 3	23
4.6.5	Week 4	24
4.6.6	Week 5	24
4.6.7	Week 6	24
4.6.8	Week 7	25
4.6.9	Week 8	25
4.6.10	Week 9	25
4.6.11	Week 10	25
4.6.12	Week 11	27
4.6.13	Week 12-14	27
4.7	Web Application	27
4.7.1	Login, Registration and Forget Password Page	27
4.7.2	Home page	30

4.7.3	Post Preview page	30
4.7.4	Notification	31
4.7.5	Tags	31
4.7.6	Find Friends	32
4.7.7	Profile	32
4.7.8	Edit Profile	33
4.7.9	Direct Messages	34
4.8	Jupyter Note Book	35
5	Conclusion/ Recommendations and Future Works	38
5.1	Conclusion	38
5.2	Future Works	38
References		40
Appendix A	Attachment A	41

List of figures

1	Component Of Research Process	vi
3.1	JSONFormatte	7
3.2	ResultOfScrapedRecipes	8
3.3	Recipe Data Flow Chart	9
3.4	Data Science Process	10
3.5	Modeling Architecture	11
3.6	Pycharm	12
3.7	Jupyter	12
3.8	GitHub	12
3.9	Trello1	13
3.10	Trello1	13
4.1	Django-Python	14
4.2	Model View Template	15
4.3	Use Case Diagram	16
4.4	ClassDaigram	17
4.5	Sequence diagram	18
4.6	State Chart Diagram	19
4.7	Activity Diagram	20
4.8	Registration	28
4.9	Verification Email	28
4.10	Login	29
4.11	forgetPassword	29
4.12	HomePage	30
4.13	postPreview	30
4.14	Notification	31
4.15	Tags	31
4.16	findFriends	32

4.17 userProfile1	32
4.18 userProfile2	33
4.19 editProfile	33
4.20 editProfile2	34
4.21 Direct Messages	34
4.22 search Results 1	35
4.23 search Results	35
4.24 search Results 2	36
4.25 search Results 3	36
4.26 search Results 4	37
4.27 search Results 5	37

Chapter 1

introduction

This chapter will talk briefly about the problems that faced the development period, objectives, scope and the importance of this platform.

1.1 Problems

The idea comes from that there isn't a specialized Social media platform in food industry in general that they could interact, share experiences and meet Professionals in the industry of food. And, of course, the machine learning approach make it so simple for them to find the meals they would like to cook, with the whole instructions and recipes which make it free and available all the time. More over, Rating of recipes as more people liked them make it easier for them to find the most popular recipe to cook and for students who didn't have any resources to know how to cook could find all Kinds of recipes and the instructions all in one in our platform, and there is no website that can offer a feature that search a recipe by its ingredients as what we offer in our platform.

1.2 Objectives

This research aims to design a platform that Help users to search for recipes by providing a search engin system with the ingredients of the wanted recipe.

This Platform comes with many features for the user, which the users can provide the search engin the ingredient of the wanting recipe. Then, it will return the mostly matched recipes that have the ingredient that the user provided.

Moreover, the users could interact, share, follow and rate the recipes that the other users shares in the platform. Each user has an account that he could brows all the content in

the websites, first he goes to explore page then he could follow the content he likes more, and his account id secured and verified by email, and the users could reset there password in case they forget it, In addition, They could edit there data, or posting the recipes they made, providing as with Data we need in the search engin, and manipulate these data in a recommender system to provide them with the wanting recipes.

1.3 Scope

The product we come up with is Software + Research Recommender system, this platform built mainly in (Django Python in the Back-End, PostgreSQL as Database, MVT (Model ,View ,Template) architecture, and the Front-End built with HTML5, CSS3, JS, Bootstrab, JQury, Vanilla JavaScript and Tailwind CSS. For communication between the MVT architecture uses the built in API's Django REST Framework. Moreover, This platform has mobile application, Which built with Flutter, the porpus of this application is to chatting between users and provide the mainly features of our platform in users mobile.

In the End For communication and make it easy to progress in building this platform. the best choise is using the distributed version control and source code management functionality Git /Github in team work. In addition, For The work process we used Scrumban, which is an Agile development methodology that is a hybrid of Scrum and Kanban and organizing our work and tasks to be done and completed in Trello.

1.4 importance

There is no social media platform that specialize in food industry, recipe sharing and data collecting to benefit other users from the search Engine, so this platform provide an easy and reliable way to let users in general and students in particular to search there recipes using the Search engine(Recommender System) by ingredients, as we mentioned previously this platform aimed to help students who are far from there families, wives ,husbands in saving time by using the search engine of ingredients.And money by providing the search engine with the available ingredients of recipes. And of course have fun and spend some time learning and trying new recipes for beginners, Moreover, Share knowledge and experience of Professional users and amateurs.

Chapter 2

Constraints, Standards/Codes and Earlier course work

2.1 Constraints

Cook-Overflow is not the only social Media Platform that has posts, comments, or interactions. But, the Main Problem is that how to add features to make it different. That was the most difficult point this platform faced.

Adding search Engine infrastructure based in its specialization was the key, and marketing it as the only website that contain Food recipes with interactions and communication with others and were even the beginners could contribute to its contents. Moreover, We customize the infrastructure that been built from zero with the AI search engine that been successfully deployed in the backend to the next Phase of the Research and developments Engineering.

2.1.1 Dataset

Basically we faced a challenge in preparing our Data. The idea is not just got a ready Dataset from Kaggle, because we need it to be harmonic with what we built in the website (Food Social Platform). So, we prepared our data with what's compatible with our vision from the Platform. And we made web Scraping to some recipes from trusted websites. which are:

1. Foodnetwork.com
2. Epicurious.com
3. Allrecipes.com

And we Scraped more than 72,000 recipes using the libraries of python 3 setuptools==28.6.1, beautifulsoup4==4.5.3 and there was three JSON files from three different websites divided to four Columns: Title, Picture URL, Instructions and Ingredients. Then, After we Processed the data with topic modeling and Text Rank and analyses the ingredients with instructions we added the tag Column.

2.1.2 Hardware Problems

When we tried to use our machines (laptop's) to Pooling and TF-IDF vectorized by recipe it took very long time and didn't completed, so we used Google Colab to upload all our JSON Files, Then. we converted them to CSV File. Not to mention that we used Clemson University Supercomputer (Palmetto) as we had access to there Supercomputer to train our model, Moreover, In Web scraping phase,internet connection was not enough. so we go to NNU university to make that happen, because we scraped more than 1 Tera then after Cleaning it reached the half.

2.1.3 Lack of similar abstractive approach in the architecture

To use and prepare dataset from your own built website and then Deploy it again and again so we come up with micro services approach that our model talk with website database Using API ,which make the architecture design more complex so we could use DOS approach which our Model in different server, and the Platform in other Server so that these servers could communicate with each others.

2.2 Standards/Codes

Most companies and institutions have charisma and standards that they need to follow to Add any Feature in the Product. And, In CookOverflow, we keep adding any possible improvement in the process of development itself, from Code Quality, Clean Code, Design Patterns and Time Organization.

2.2.1 Design and visual identity guidelines

First of all we collect the requirements, designed the UML at Draw.io ,the state diagram, the sequence diagram and the User interface (Front end). Then, We assigned tasks for each one in our team and we write complete description for each task at Trello, After that, We Chose a framework matches with our requirements and system infrastructure, Then, We used Version

control (Git/Github) to make our work as professional as possible, In addition, We wrote clean code, reviewed each dependencies. And, created a python environment to download the needed libraries for our platform, All in all, We lived the software life cycle from collecting requirements to designing to implementing and testing and so on.

2.2.2 Coding conviction

When it comes to code there are many recommended standards and best practices and sometime these practices might vary depending on your language of choice. For example, In CookOverflow we used Django (Python) as our backed framework and with these came a set of guidelines that we recommended to adhere to you can see the full set of guidelines and best practices for Django in this reference Django [1]

2.3 Earlier course work

We tried to use our cumulative course work and extra curriculum and training. We learnt during four years at the college at Computer Engineering Department the following:

2.3.1 Database

One of the most important courses in the university that we benefit from it to design our 27 Tables, draw the UML Diagrams, and SQL to the Back-End of our manipulation of the data we want, we used PostgreSQL for the Database.

2.3.2 Web Development

Web development course was one of the most helpful courses for this project since we are building a web platform for the end-users. along with a set of extra resources especially the largest we learned Django to use it for our Back-End Platform, And learned how to make verification to email and make our website as secure as possible.

2.3.3 Software Engineering

All the development operations (DevOps) that were implemented during the project was out of the things we learned during the Software engineering course, the code standards and app workflow design.

2.3.4 Artificial Intelligence

the Artificial Intelligence course was the core for all progress we made in our research since without it we won't be able to understand any concept while taking further advanced courses to achieve our goal of building an AI search Engine for recipe ingredients.

2.3.5 Advanced Software Engineering

Digging deeper in the development and architecture of the website and the pipelines building, the signal API architecture for the notifications in the Platform and adding likes and so on.

2.3.6 Security

Even though we didn't take the security course but we took the basics during the Network courses and we extended our knowledge through online tutorials to implement Django authentication.

2.3.7 Mobile

We built simple flutter app for chatting between users of the website.

2.3.8 Machine Learning

We took course with Clemson University helped us to implement the unsupervised learning and clustering technique and built the recipe Search Engine AI.

Chapter 3

Literature Review

In this project we tried to add a special feature that all website aiming to have from its users which is collecting data from website and processing them in a Machine learning model. Then deliver them to users as a recommendations. So in this Chapter we will discuss the AI/ML Life cycle from data collection ,cleaning ,parsing ,modeling and make the end to end connecting points with the Back-End. Then the deployment of the model making every step corresponding with the requirements of the product.

3.1 Data

We scraped our data from three websites, 72 000 recipes as JSON then during the preprocessing and cleaning, we convert them to CSV file.

Scraping process Done by these Two liberties: `setuptools==28.6.1` and `beautifulsoup4==4.5.3`. In this format as Shown in the Code :

```
return {  
    'title': title,  
    'ingredients': ingredients,  
    'instructions': instructions,  
    'picture_link': picture_link,  
}  
  
## get_all_recipes_fn(page_str, page_num):  
    base_url = 'http://www.foodnetwork.com'  
    search_url_str = 'recipes/a-z'  
    url = '{}//{}//p/{}'.format(base_url, search_url_str, page_str, page_num)  
  
    try:  
        soup = BeautifulSoup(request.urlopen(  
            request.Request(url, headers=HEADERS)).read(), "html.parser")  
        recipe_link_items = soup.select('div.o-Capsule_m-Body ul.m-PromoList li a')  
        recipe_links = [r.attrs['href'] for r in recipe_link_items]  
        print('Read {} recipe links from {}'.format(len(recipe_links), url))  
    return recipe_links
```

Fig. 3.1 This is just a long figure caption for the minion in Despicable Me from Pixar

Here are the Result of The Scraped Recipes:

```

1  [
2   "05zEpbSqcse0rcnCJWyZ90gdH0MLby": {
3     "ingredients": [
4       "12 egg whites",
5       "12 egg yolks",
6       "1 1/2 cups sugar",
7       "3/4 cup rye whiskey",
8       "12 egg whites",
9       "3/4 cup brandy",
10      "1/2 cup rum",
11      "1 to 2 cups heavy cream, lightly whipped",
12      "Garnish: ground nutmeg"
13    ],
14    "picture_link": null,
15    "instructions": "Beat the egg whites until stiff, gradually adding in 3/4 cup sugar. Set aside. Beat",
16    "title": "Christmas Eggnog"
17  },
18  "mF5Szmoqxf4wtI1hLRvzuKk.z6s7P2S": {
19    "ingredients": [
20      "18 fresh chestnuts",
21      "2 1/2 pounds veal stew meat, cut into 2x1-inch pieces",
22      "4 tablespoons olive oil",
23      "1 1/2 cups chopped onion",
24      "1 1/2 tablespoons chopped garlic",
25      "1 bay leaf",
26      "2 1/2 cups canned low-salt chicken broth",
27      "3/4 cup dry white wine",
28      "6 medium carrots, peeled, cut into 1-inch pieces",
29      "3 tablespoons chopped fresh sage"
30    ],
31    "picture_link": null,

```

recipes_raw_nosource_epi.json: tokenization, wrapping and file handling have been turned off for this large file in order to reduce memory usage and avoid freezing or crashing.

[Don't Show Again](#) [Forcefully Enable](#)

Fig. 3.2 This is just a long figure caption for the minion in Despicable Me from Pixar

3.2 Algorithm and Approach

First of all, We scraped our dataset from the following websites:

1. All Recipes.
2. Epicurious.
3. Food Network.
4. Our Own website.

Secondly, We processed the dataset as the following:

1. Filtering out incomplete recipes and converting the dictionary to string.
2. Add stop words after recipe title, ingredients and instructions.
3. Fixing the length of recipes by removing too large recipes and padding the smaller ones with a special character. Finally left with 72000 recipes.

4. NLP : We have used character level RNN's where we have first vectorized the input dataset into numbers and then converted them into a Tensorflow dataset for final training.
5. Then we split up the dataset into batches of 64 and final started the training.[7]
6. We have applied different Deep learning:
 - (a) Optimizers: Adam, RMSProp & SGD(Stochastic Gradient Descent).
 - (b) Learning rate: varying from 0.005 to 0.001.
 - (c) Loss Function: sparse_categorical_crossentropy & categorical_crossentropy We were able to successfully achieve a best accuracy of 93.20% using RMSProp Optimizer, sparse_categorical_crossentropy as the loss function and 0.0005 as the Learning Rate. (But this Was under Research, we didn't Apply it to the Final Version of the Product).

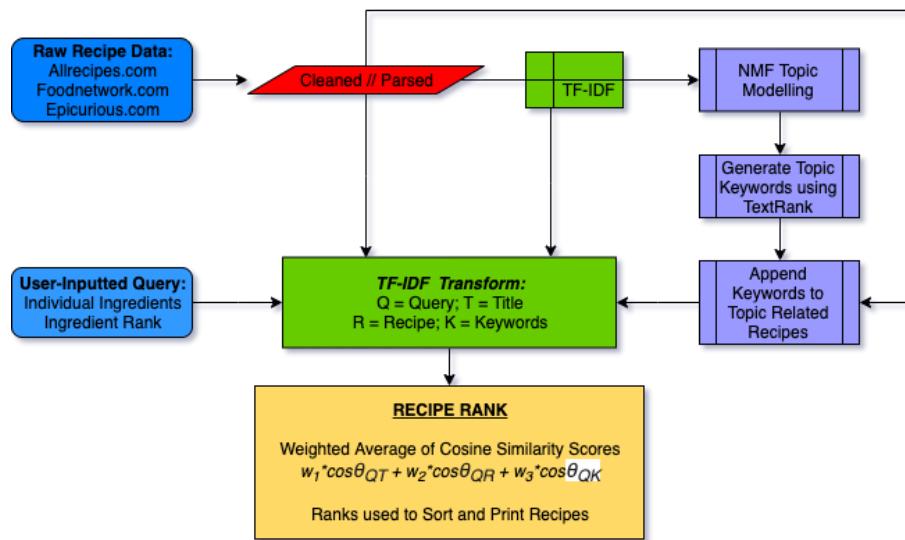


Fig. 3.3 CookOverflow Data Flow Chart

3.3 Model

The Data Science process

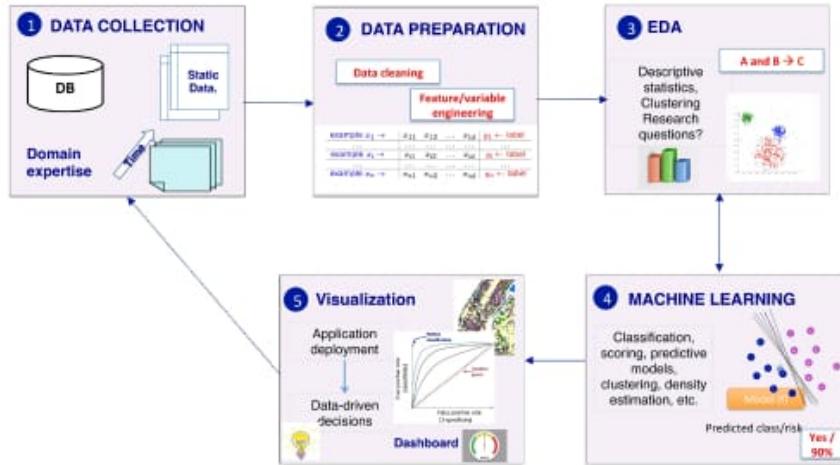


Fig. 3.4 Data Science Process and Life Cycle

The system we have built will be used for Recommendation of recipes based on ingredients from the given set of input and thus also requires proper designing. The Design requirements for the project are:

3.3.1 Knowledge Requirement for Modeling

There are many knowledge requirements for modeling our Platform:

1. Knowledge of Web Development:

To develop the registration process and the recipe recommendation and Search AI algorithm process, the knowledge of Web Development was required. The team has knowledge of Web Development, and has used in CookOverflow for the front-End: JavaScript, CSS, AJAX and HTML5. And Django a Python framework for the server-side scripting needs. Finally PosgresSQL for the database.[4]

2. Knowledge of Machine Learning Especially Unsupervised Learning:

Since the application required designing a model which could help clustering the recipe based on NLP tags , the knowledge of Unsupervised Learning was required.

Our team had knowledge of Neural Networks, and our application specifically used RNN's(Recurrent Neural Network).

3. Knowledge of UML Diagrams:

Knowledge of UML diagrams was essential to create the different types of models which is an essential component of Software Engineering and Design Architecture.

3.4 Modelling the Architecture

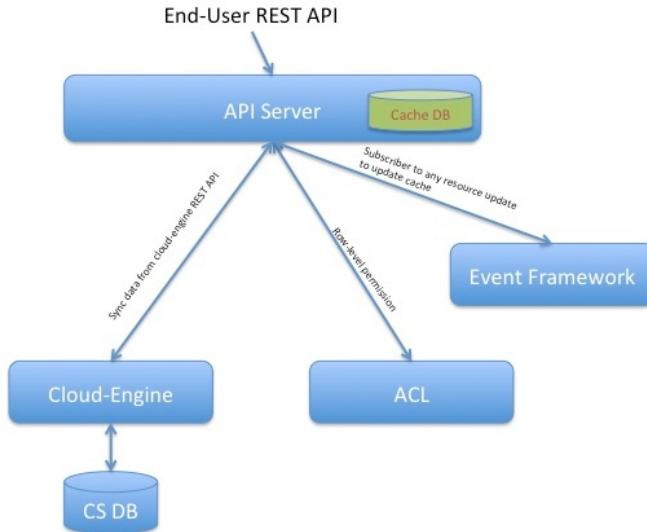


Fig. 3.5 API Server Architecture

1. Server:

The system requires servers to process applications and to host databases, to and from which the querying is done. For development purposes the team has used their local machines and For Microservices Architecture the ML Model Was in Supercomputer.

2. Web Hosting:

The application and databases are required to be hosted on the web, to provide anytime access to them. For development, the team has used the localhost to host applications.

3. IDEs:

Since various technologies are being used, specialized IDEs and tools are used for them. The team has used:

- (a) Pycharm and VSC and Web Development.



Fig. 3.6 Pycharm IDE

- (b) Jupyter Notebook for Development of Python Script.



Fig. 3.7 Jupyter IDE

4. tools :

- (a) Git/GitHub Version Control (DesvOPs).

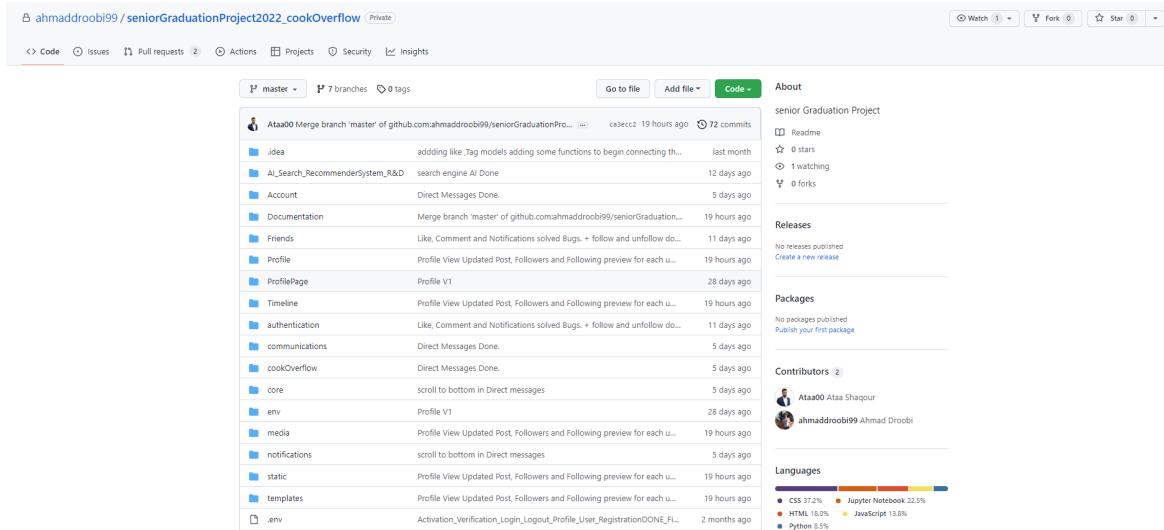


Fig. 3.8 Our GitHub

- (b) Google Colab, Palmetto (Super Computer of Clemson).
- (c) Draw.io For UML Drawing.
- (d) Overleaf (Design The Paper on Cambridge Template), LaTeX.
- (e) PowerPoint (Presentation).
- (f) Trello (Agile Development Process) For Team work, Assignment of Tasks.

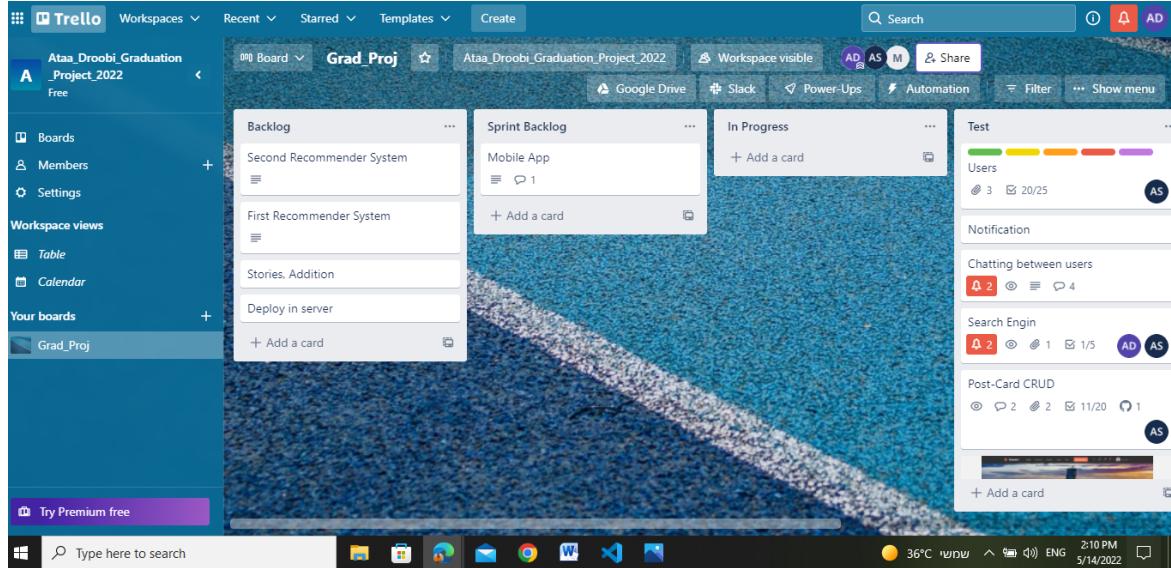


Fig. 3.9 This is just a long figure caption for the minion in Despicable Me from Pixar

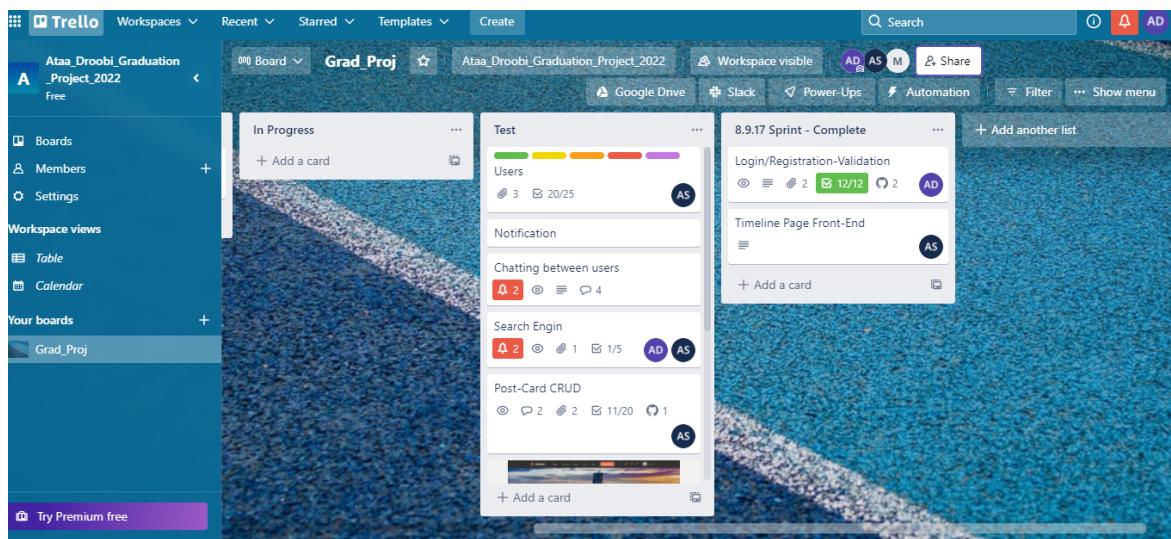


Fig. 3.10 This is just a long figure caption for the minion in Despicable Me from Pixar

Chapter 4

Methodology

This chapter will talk about the process of development that this platform goes on. By mentioning the languages and frameworks that we use and the process of creating the search engine model.

4.1 Languages and Frameworks

This section will have the main languages and frameworks that we used in our development cycle.

4.1.1 Back-End-Framework



Fig. 4.1 Django-Python

We choose Django/Python this platform is a big project, which Django is Good choice for that. And to make it easy to deal with the search engine which written in Python language which is the same language that used in Django. Moreover, Its scalable and its excellent in CDN (A content delivery network (CDN) refers to a geographically distributed group of servers which work together to provide fast delivery of Internet content.) connectivity and content management.[2]

4.1.2 Back-End-Security

One of Django power is security. It contains clickjacking protection in the form of the X-Frame-Options middleware which in a supporting browser can prevent a site from being rendered inside a frame. In addition. In this platform we used CSRF protection which is middleware and template tag provides easy-to-use protection against Cross Site Request Forgeries. For the login registration, we used the built in Django authentication , which provides both authentication and authorization together.

4.1.3 Back-End-URL Structure

Django REST framework is a powerful and flexible toolkit for building Web APIs. Some reasons you might want to use REST framework: The Web browsable API is a huge usability win for your developers. Authentication policies including packages for OAuth1a and OAuth2.

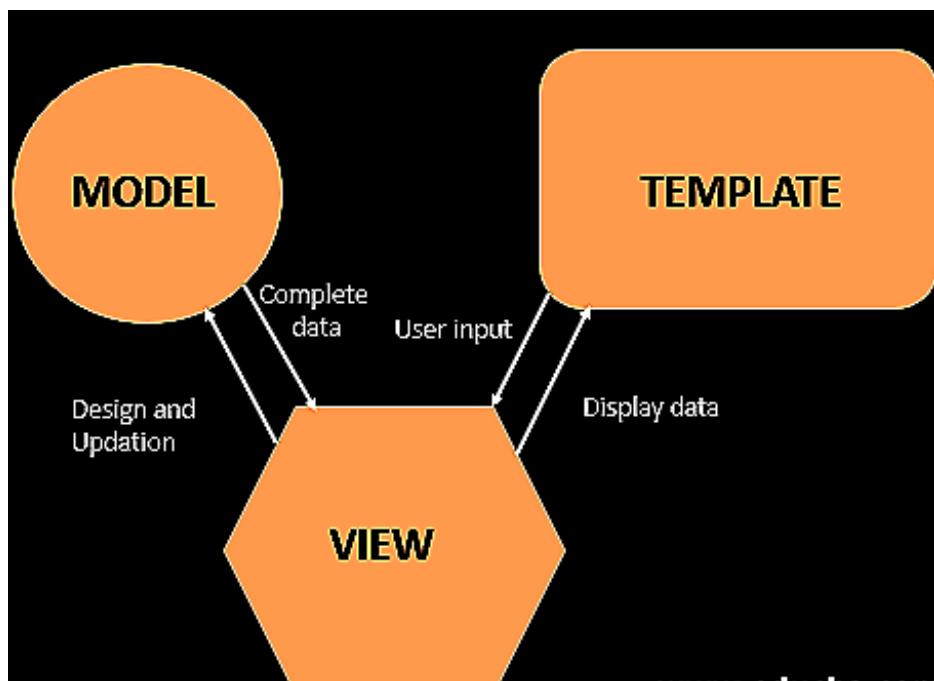


Fig. 4.2 Data Science Process and Life Cycle

4.1.4 Front-End

For the Front-End we used HTML5, CSS which are (Bluma), Vanila JS. And for sure, Django came up with one or several template engines.[8] Django ships built-in backends for its own template system, creatively called the Django template language (DTL), and for the popular alternative Jinja2. Backends for other template languages may be available from third-parties. It makes easy to write your own custom backend.[3]

4.2 UML Diagrams

This section will talk about the UML diagrams of the search engine. Which are Use-Case, Class, sequence, state and Activity diagrams.

4.2.1 Use-Case Diagram

The use case diagram starts with the logging in and then entering entering the symptoms values.

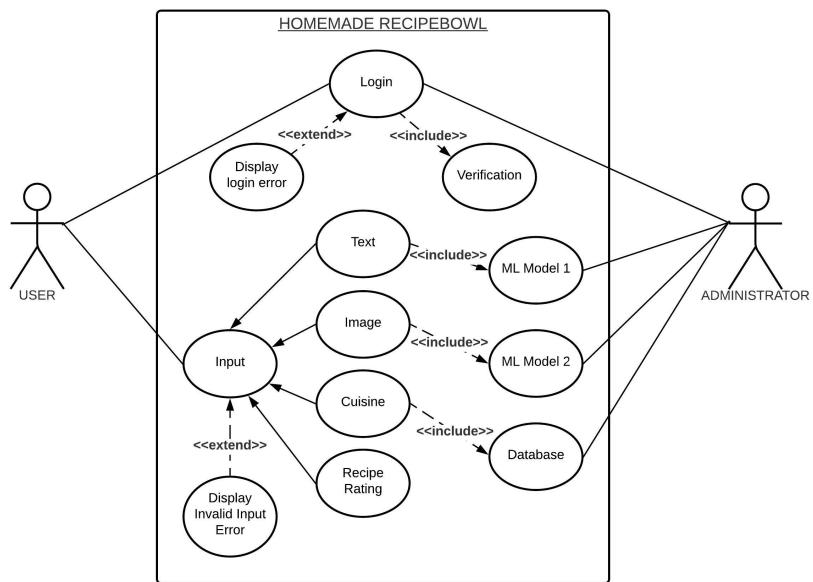


Fig. 4.3 This is just a long figure caption for the minion in Despicable Me from Pixar

4.2.2 Class Diagram

Shows the various components and interactions of various classes and their associated cardinality.

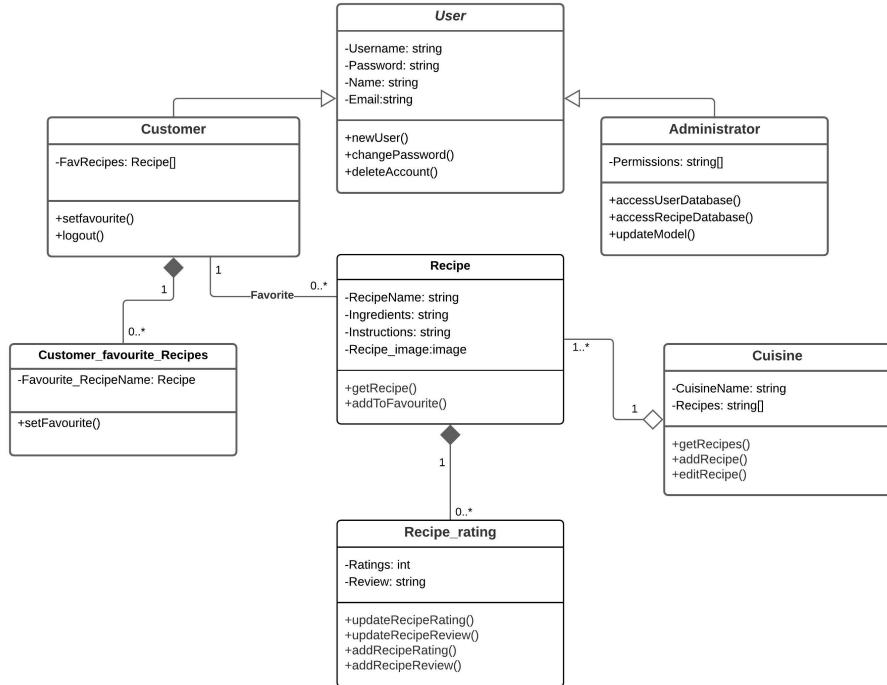


Fig. 4.4 This is just a long figure caption for the minion in Despicable Me from Pixar

4.2.3 sequence Diagram

A sequence diagram is a type of interaction diagram because it describes how—and in what order—a group of objects works together.

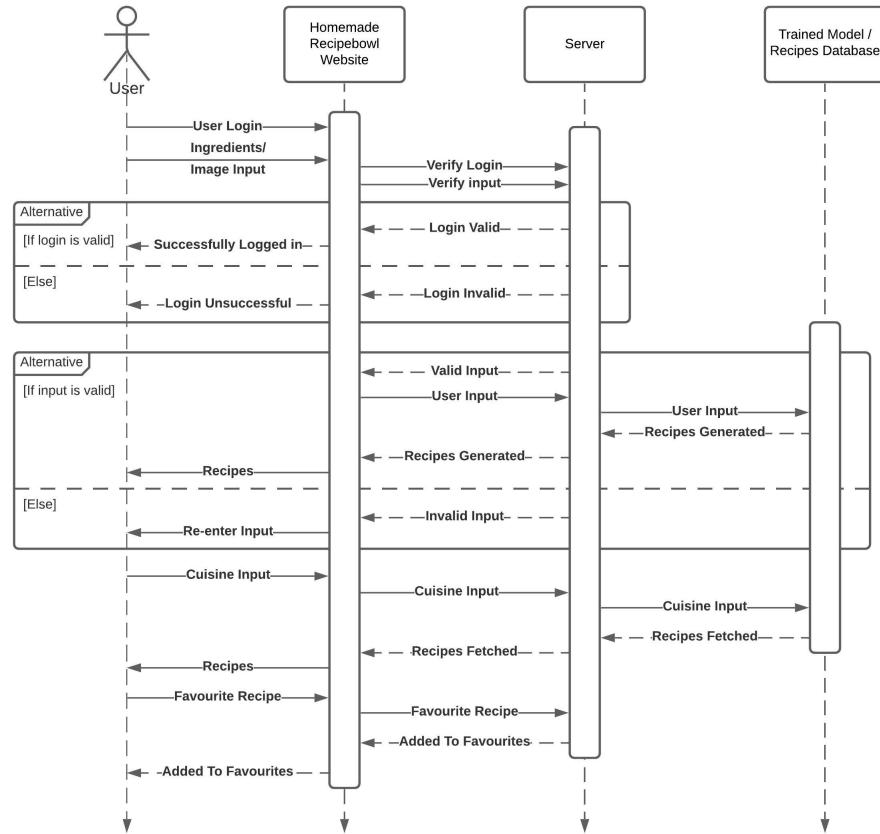


Fig. 4.5 This is just a long figure caption for the minion in Despicable Me from Pixar

4.2.4 State Diagram

They are behavioral diagrams used to represent the conditions of the system or part of the system at finite instances of time.

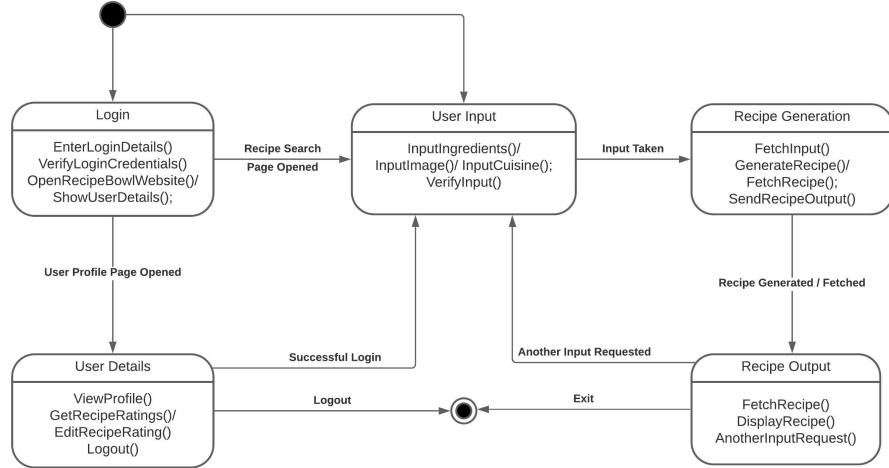


Fig. 4.6 This is just a long figure caption for the minion in Despicable Me from Pixar

4.2.5 Activity Diagram

An activity diagram is a behavioral diagram i.e. it depicts the behavior of a system.

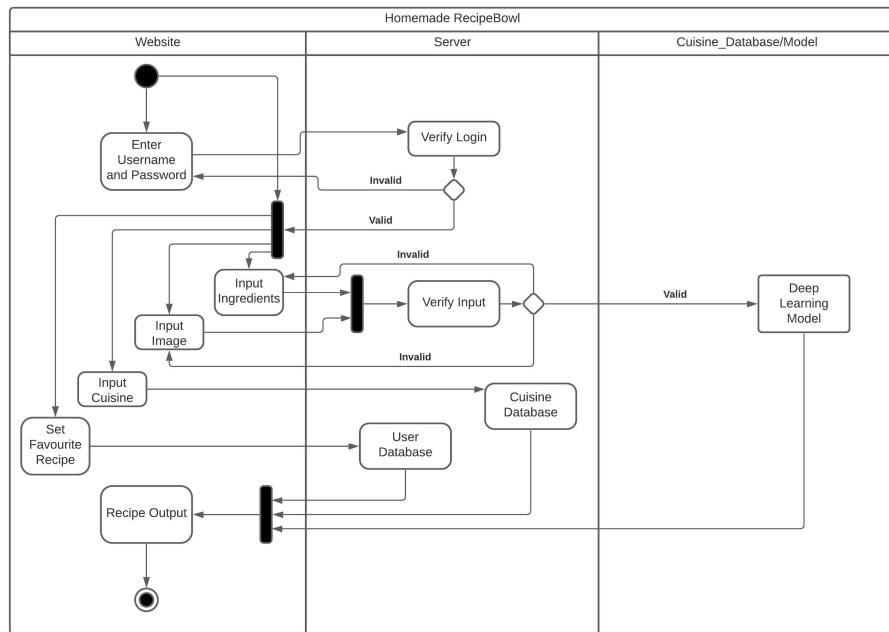


Fig. 4.7 This is just a long figure caption for the minion in Despicable Me from Pixar

4.3 Dataset

we Web scraped Three Json File from Three websites we already mention them above ,and we convert the Post Table in the Database to Json and append it to the ML Model to Process the new Data every Week for example in different server ,by this approach we Trained our model in Real data , and from factual Data from Our own , built from zero website

4.4 Text Processing(Spacy) and NLP

Tokenizing Using Spacy: For this tokenization, we will lemmatize the words. This is will help create a denser word embedding. However, no POS tagging, know entities, or noun_phrases will be parsed and added. and then we Used TF-IDF Modeling as the words just Represent themselves and there's no correlations between the ingredients with each others .[Khan]

4.5 Search Engine (Query Algo)

The final product presented is a search algorithm that takes in a list of ingredients or categories, and uses the query to return relevant recipes that utilize those ingredients or are similarly related to other ingredients and those recipes.[JAGITHYALA]

4.6 TimeLine

This section will talk about more details of our process in this platform.

4.6.1 Week 0

Team formation and Mentor Selection:

1. Team Member Formation.
2. Choosing the field/technology of interest after a series of meetings with the team members.
3. Choosing the field/technology of interest after a series of meetings with the team members.
4. Selecting the appropriate mentor best suited for the technology to be used.

4.6.2 Week 1

Project Idea Discussion and Synopsis Drafting, We came up with three project ideas:

1. A gaming project which uses physical movement detection for moves in the game.
This would collaborate exercise with gaming.
2. Personal voice assistant software and hardware for blinds to detect the objects around.
3. Platform for AI Search Recommendation recipes based on given set of ingredients.
This was influenced by the cooking trend during lockdown situation.

Finally, we decided to take up the third idea, as we found it to be a more practical and useful idea. And also it was feasible according to our knowledge and skill set.

4.6.3 Week 2

Discovering Project Requirements:

1. Looking onto the existing work in this related field, if done.
2. Deciding the further contributions we can make.
3. Looking for availability of datasets for training purposes of our models to be able to predict the required output.
4. Looking for platform to perform high end computations like web scraping and machine learning model training.
5. Developing a user interface through which the user can access all the functionality easily and innovatively.

Software Requirements Specification (SRS):

1. Purpose: to develop a system which primarily is capable of generating extensive recipes out of given set of ingredients.
2. Intended Audience:
 - (a) Any person of any age group, any region and any profession who loves exploring food.
 - (b) A very large size audience!!!, students, Professionals (Cheifs), our Mothers ...ects.

3. Scope:

- (a) Fetch/Scrap relevant dataset.
- (b) Decide upon at deep learning technique for text generation and Modeling And Searching.
- (c) Train the ML Model on variations of hyper parameters.
- (d) Develop high tech and easy-to-use user interface system.

4. Product Features:

- (a) Ingredients-to-recipe feature.
- (b) Full Socail Media Platfrom (Notifications, Posts ,Comments,Rating ,Tags...).
- (c) Image-to-recipe feature(under Rsearch).
- (d) Nutritional analysis for every recipe(under Rsearch).

5. User Interface Requirements:

Easy to use dynamic website which runs on exhaustive set of devices and which is capable of handling heavy user load.

6. Hardware Requirements:

High end computational machine required for Web Scraping and ML Model training if required.

7. Software Requirements:

3rd Party dependencies required for Web Scraping like (BeautifulSoup, JSoup), Model Training script like (Tensorflow, Keras, Pytorch) and Web interfaces like (React or Vue for frontend, Flask, Express or Django for backend).

4.6.4 Week 3

Exploring datasets and websites for scrapping (if required):

1. Looked up for the existing datasets on the internet.
2. Found one nearly exhaustive dataset by MIT, but the access was not public, wasn't accessible.
3. Could not find any other dataset related to the Recipe name along with ingredients and instructions.

4. Created the Github Repository for the project -Cookoverflow (Currently the repository is private)

4.6.5 Week 4

Scraping Websites:

1. Shortlisted three websites suitable for ingredient-recipe dataset:
 - (a) Epicurious
 - (b) All Recipes
 - (c) Food Network
2. Developed the code for scraping in python.
3. Scraped the websites using the BeautifulSoup module in python.

4.6.6 Week 5

Designing UML Diagrams:

1. Use Case Diagram
2. Class Diagram
3. Sequence Diagram
4. Activity Diagram
5. State chart Diagram

Deep Learning Course:

1. Took a brief overview of the Machine learning specialization on Coursera by Andrew NG.
2. Also referred to some online articles and tutorials regarding NLP.

4.6.7 Week 6

Writing code for training the model:

Drafted code in python using Google Colab for model training on input ingredients and giving recipes as the output.

4.6.8 Week 7

Training and Testing of the Model:

1. Trained the Machine learning model over 20 epochs on Google Colab.
2. Tested the model and retrieved the output from the model.
3. Requested Prof. Melissa Smith from Clemson access for further training on Palmetto (Clemson super Computer due to limited resources(13 GB RAM)in Google Colab.

4.6.9 Week 8

Trained and tested our model Palmetto GPU using docker and Teamviewer.

4.6.10 Week 9

Started developing the Website:

1. Started developing the website's front end using the Front End Basic Technologies.
2. Website Layout designed before the start of actual implementation by team discussions and brainstorming.

4.6.11 Week 10

Developing the Website Back-End and Integrating with front-End:

1. Developing the website's Back-End using the Django framework of python.
2. Started integrating the Front-End and the Back-End.
3. Continued with the Front-End of the website.
4. Updated the Project SRS according to further scope of the project. And Updated Software Requirements Specification (SRS).
5. Purpose:
to develop a system which primarily is capable of generating extensive recipes out of given set of ingredients.
6. Intended Audience:

- (a) Any person of any age group, any region and any profession who loves exploring food.

- (b) A very large size audience!!!

7. Scope:

- (a) Fetch/Scrap relevant dataset.
- (b) Decide upon apt deep learning technique for text generation.
- (c) Train the ML Model by varying hyper parameters.
- (d) Develop high-tech and easy-to-use user interface system.

8. Product Features:

- (a) Ingredients-to-recipe feature.
- (b) Image-to-recipe feature (under Research).
- (c) Nutritional analysis for every recipe (relevant dataset was not available and web scraping was not feasible).
- (d) Nutritional analysis for every recipe (relevant dataset was not available and web scraping was not feasible).
- (e) Adding few other features to improve user experience.

9. User Interface Requirements:

Easy to use dynamic Social Platform which runs on exhaustive set of devices and which is capable of handling heavy user load.

10. Hardware Requirements:

- (a) High end computational machine required for Web Scraping and ML Model training if required.
- (b) Used Google Colab and Clemson University Supercomputer Palmetto (using docker and Teamviewer for training ML Models).

11. Software Requirements:

3rd Party dependencies required for Web Scraping like (BeautifulSoup, JSoup), Model Training script like (Tensorflow, Keras, Pytorch) and Web interfaces (HTML, CSS, Javascript, Bootstrap, JQuery, frontend, Django, for backend).

4.6.12 Week 11

Completed scrapping and started preparing model training code:

Continued development of website by adding certain features to it(login page).

4.6.13 Week 12-14

Further Developed the Cookoverflwo Website:

1. Developed the website's Front-End.
2. Added additional features/components like login page, user dashboard, favourite recipe feature, about us page, notifications, Profiles, Tags, Follow/Unfllow, Likes, Comments, Update Profile, Post (Image ,Text ,Video ,Tag),Search, Friends View.
3. Integrated front end components with each other.
4. Integrated front end with Back-End ,And sending Info From Our Website to the Model to update its recipe Data Set.

4.7 Web Application

This section will show snippets of our CookOverflow Web-Platform.

4.7.1 Login, Registration and Forget Password Page

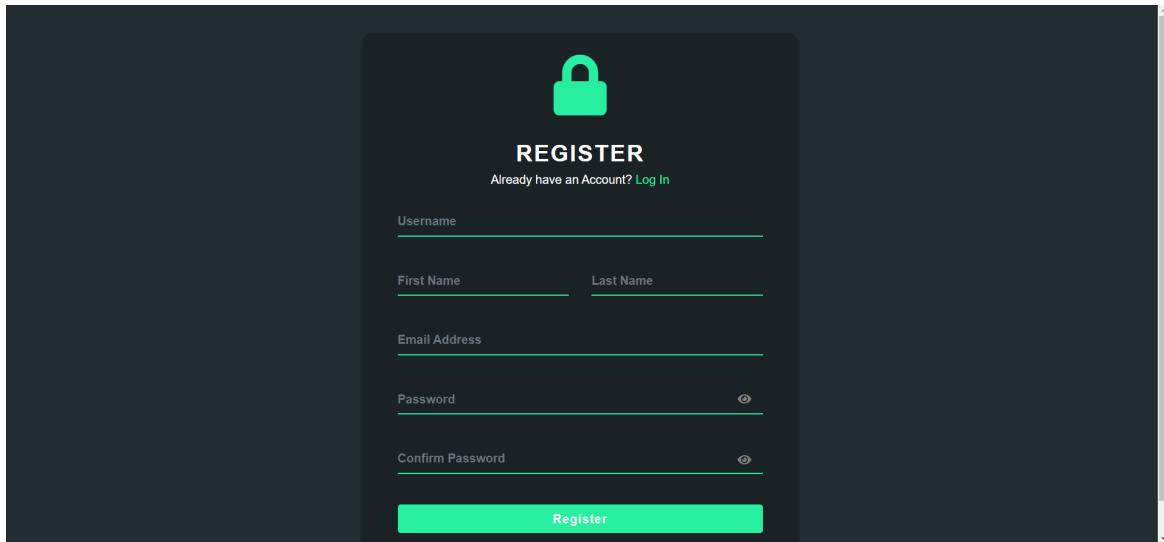


Fig. 4.8 Registration Page

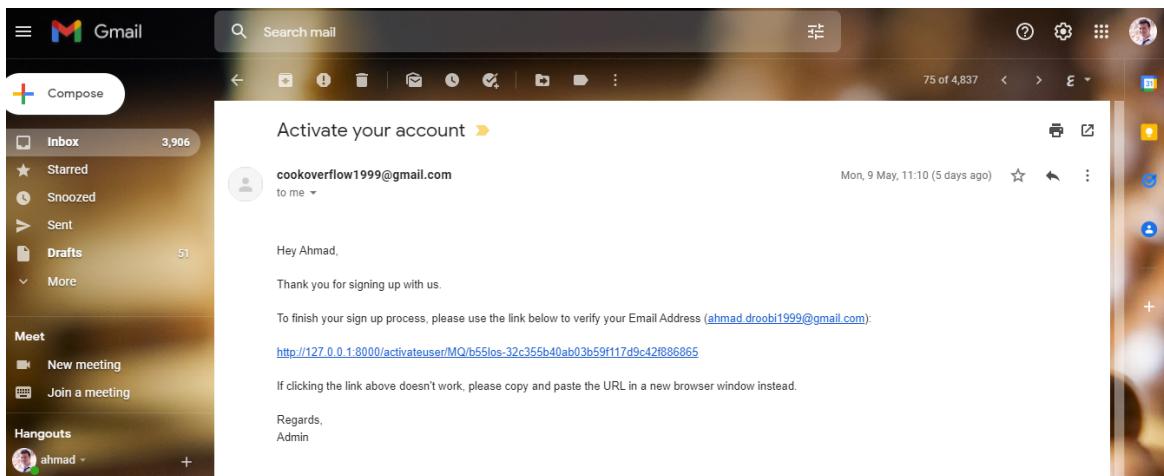


Fig. 4.9 Verification Email

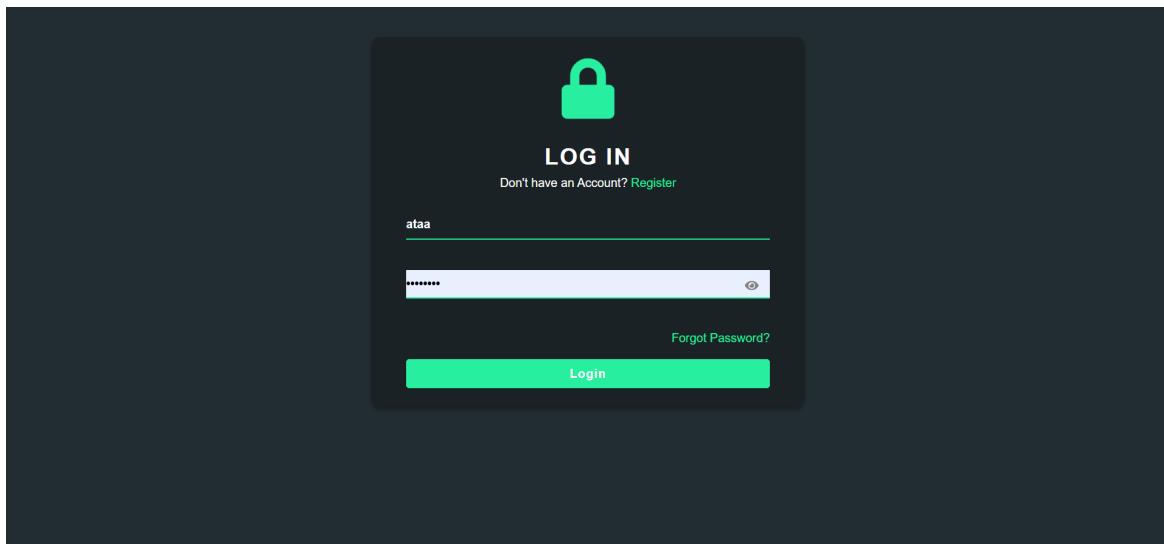


Fig. 4.10 Login Page

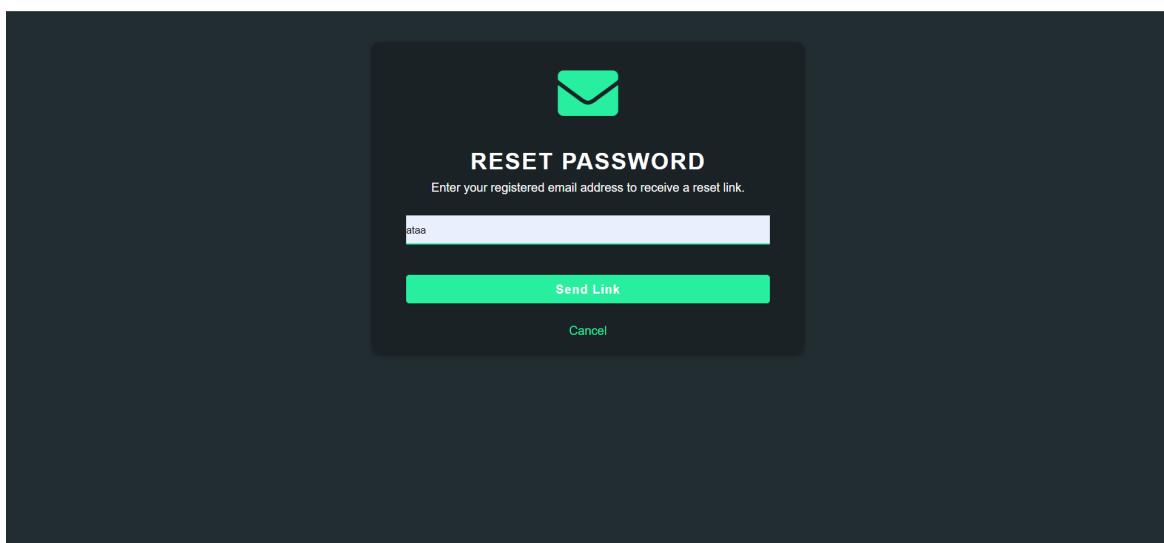


Fig. 4.11 Forget Password Page

4.7.2 Home page

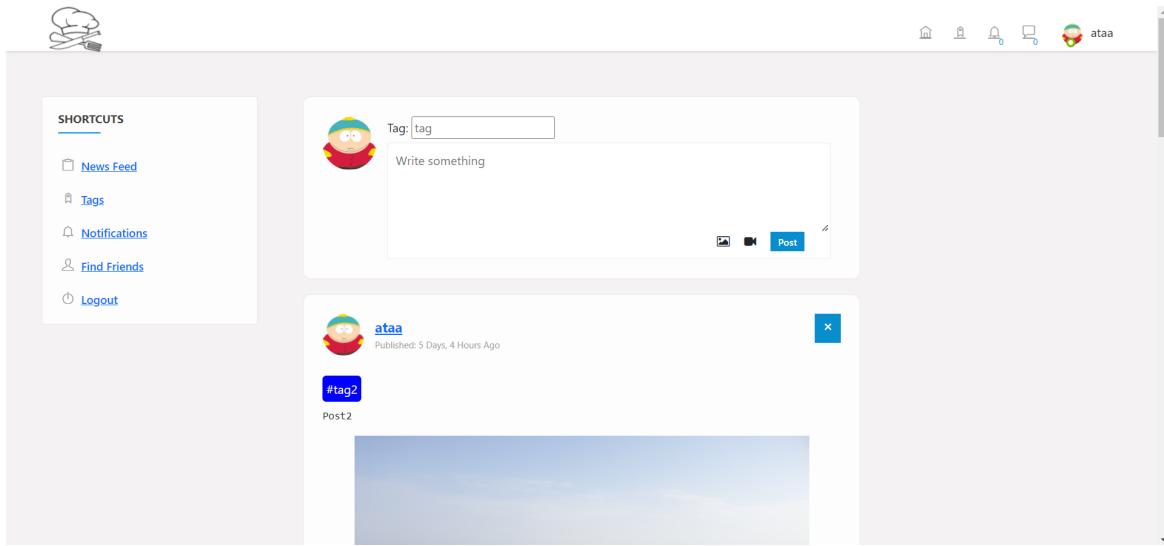


Fig. 4.12 Home Page

4.7.3 Post Preview page

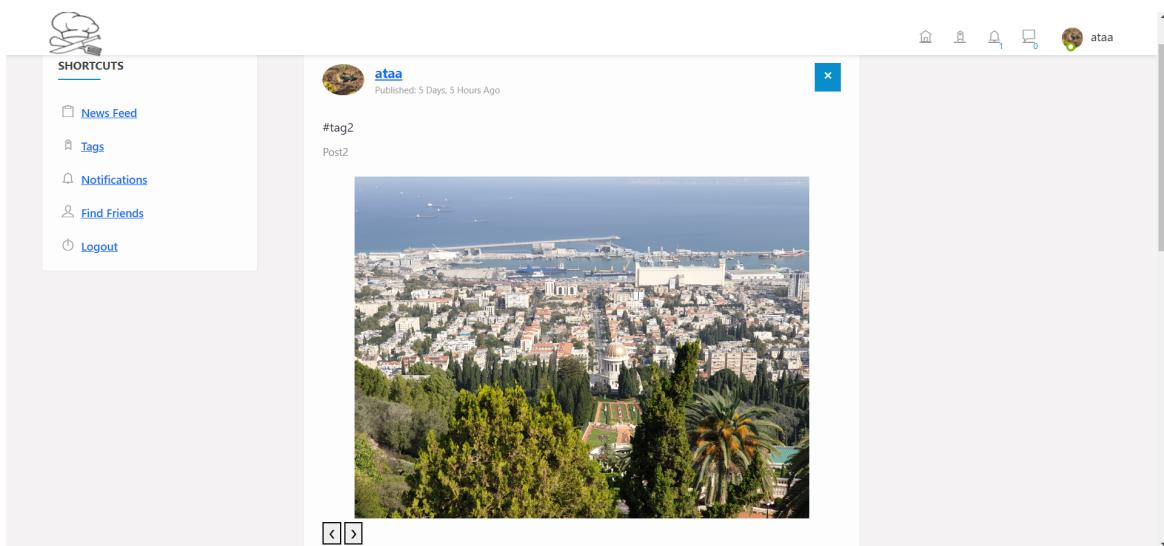


Fig. 4.13 Post Preview Page

4.7.4 Notification

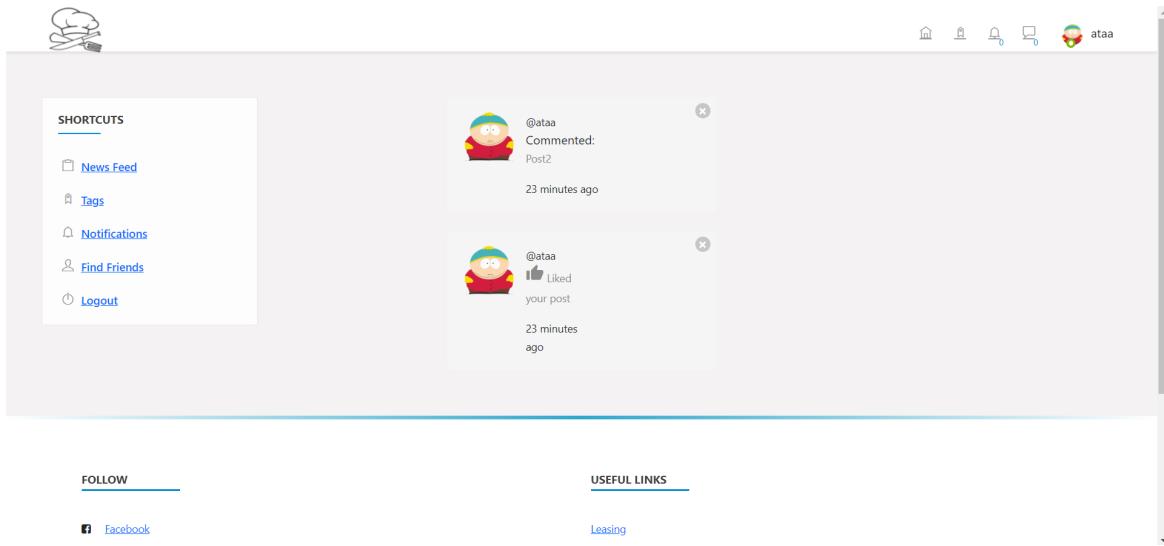


Fig. 4.14 Notification Page

4.7.5 Tags

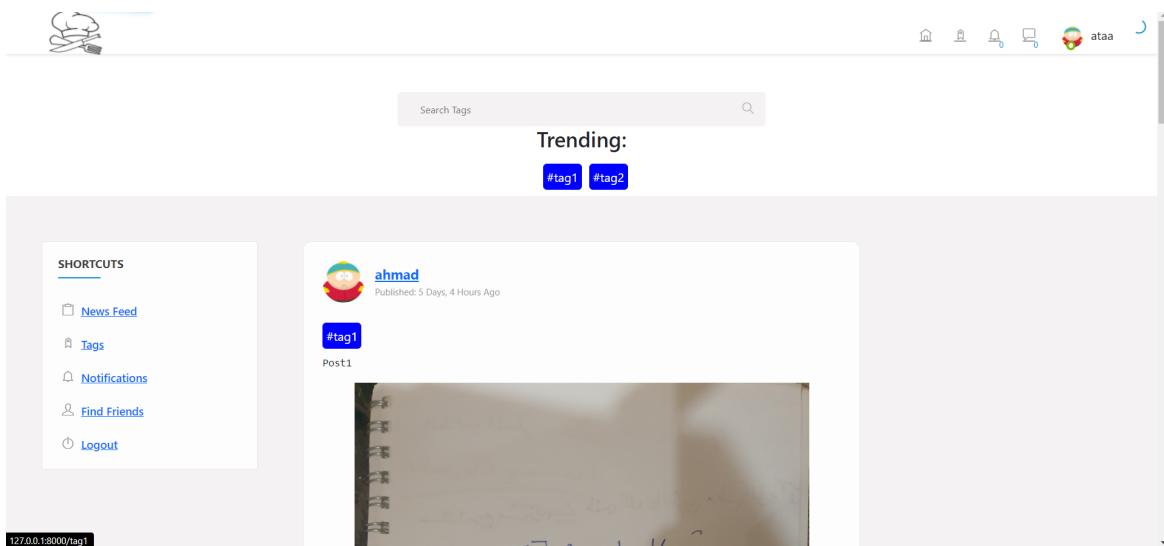


Fig. 4.15 Tags Page

4.7.6 Find Friends

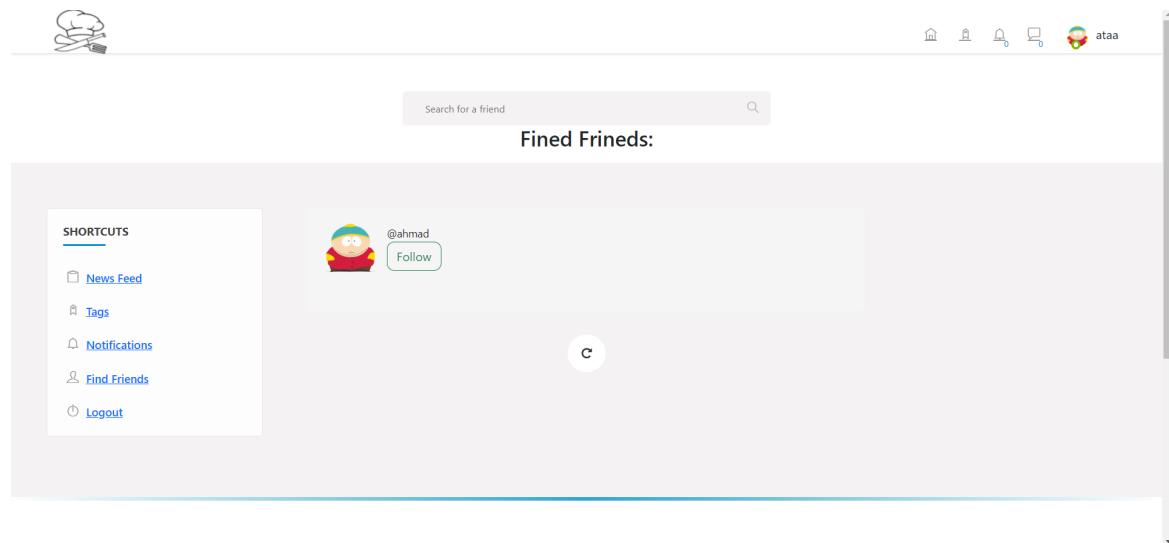


Fig. 4.16 Find Friends Page

4.7.7 Profile

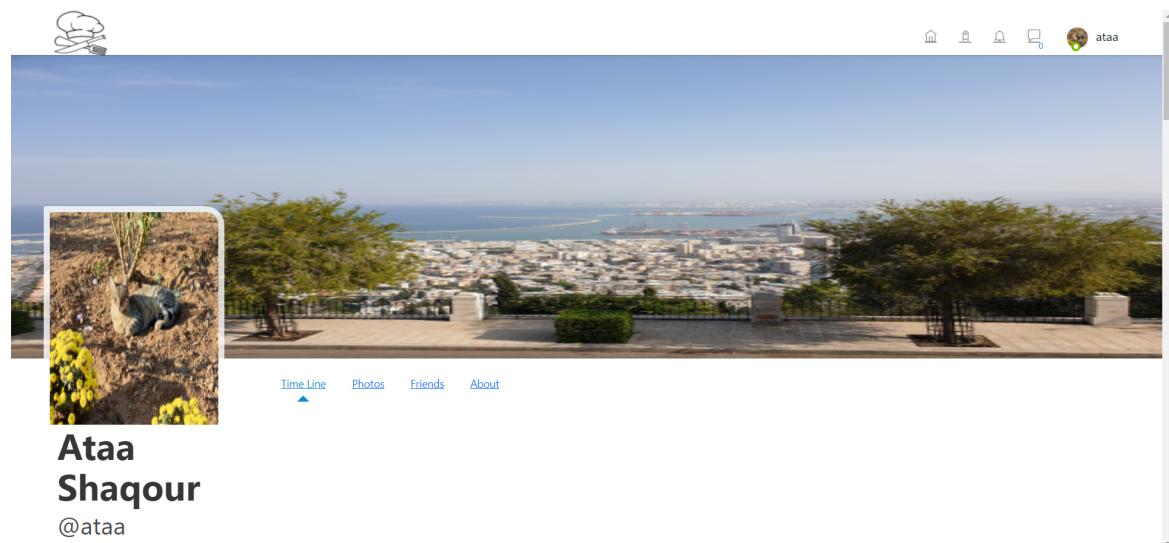


Fig. 4.17 User Profile Page

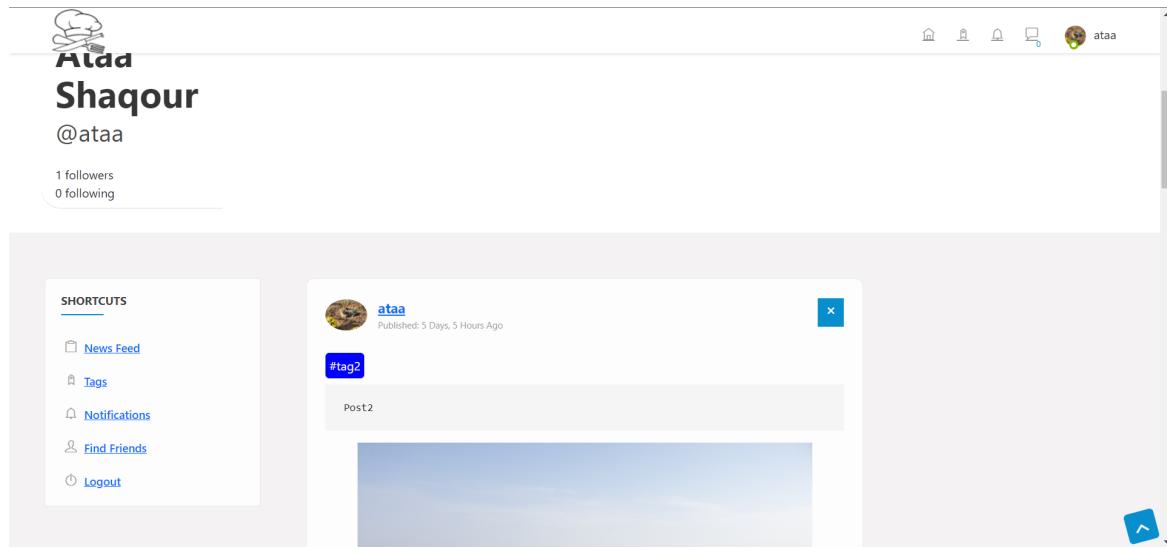


Fig. 4.18 User Profile Page

4.7.8 Edit Profile

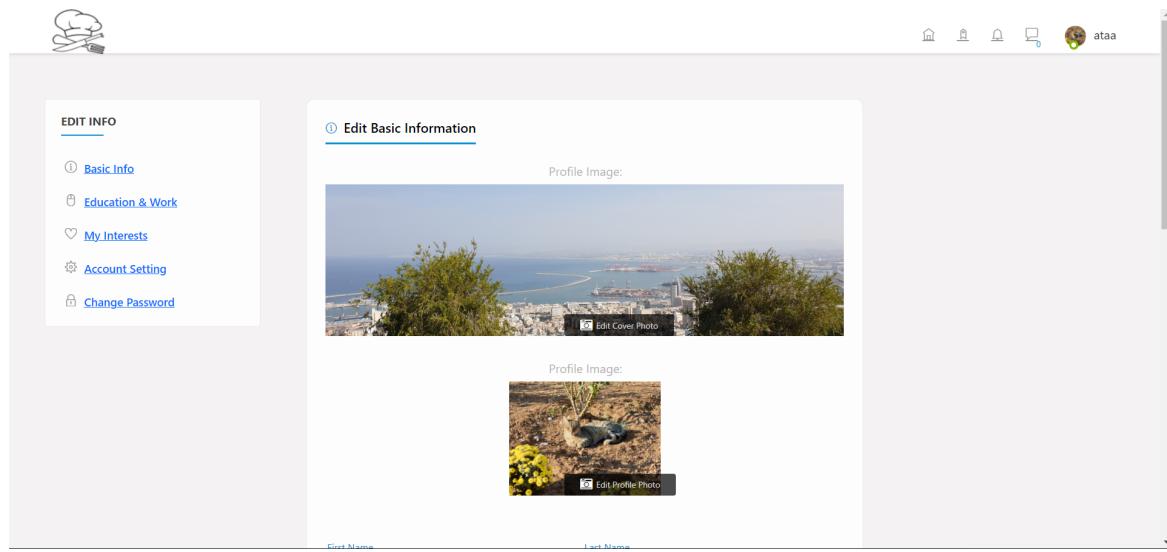


Fig. 4.19 Edit Profile Page

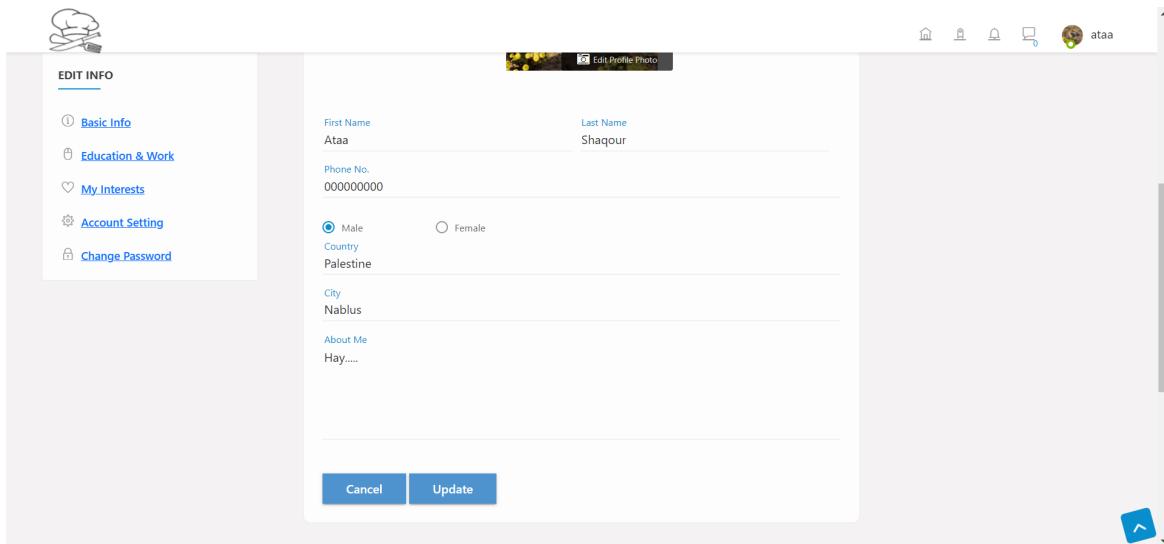


Fig. 4.20 Edit Profile Page

4.7.9 Direct Messages

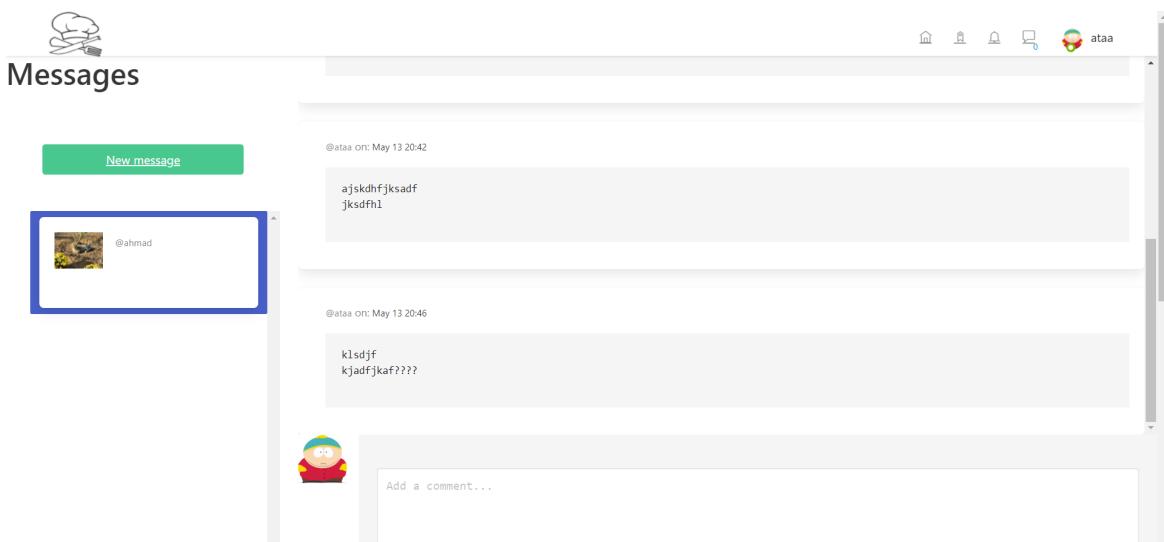


Fig. 4.21 Direct Messages Page

4.8 Jupyter Note Book

This section will show snippets Results of the Search Engine.

```
query = ['pepper','apple','pork']
Search_Recipes(query, query_ranked=False, recipe_range=(0,3))

Search Query: ['pepper', 'apple', 'pork']

Recipe Rank: 0 Apple Butter Pork Loin

Ingredients:
2 (2 pound) boneless pork loin roast; seasoning salt to taste; 2 cups apple juice; 1/2 cup apple butter; 1/4 cup brown sugar; 2 tablespoons water; 1/4 teaspoon ground cinnamon; 1/4 teaspoon ground cloves

Instructions:
Preheat the oven to 350 degrees F (175 degrees C).
Season the pork loins with seasoning salt, and place them in a 9x13 inch baking dish or small roasting pan. Pour apple juice over the pork, and cover the dish with a lid or aluminum foil.
Bake for 1 hour in the preheated oven. While the pork is roasting, mix together the apple butter, brown sugar, water, cinnamon, and cloves. Remove pork roasts from the oven, and spread with apple butter mixture.
Cover, and return to the oven for 2 hours, or until fork-tender.

Recipe Rank: 1 Shredded Apple Pork

Ingredients:
1 cup MUSSELMAN'S® Apple Butter; 1/4 cup MUSSELMAN'S® Apple Cider Vinegar; 1/4 cup brown sugar; 2 tablespoons garlic, minced; 2 tablespoons whole grain Dijon mustard; 1 teaspoon salt; 1 teaspoon pepper; 2 medium apples, cored and chopped; 1 medium onion, chopped; 6 boneless pork chops

Instructions:
In small bowl, whisk together apple butter, vinegar, brown sugar, garlic, Dijon mustard, salt and pepper.
Place chopped apples and onions in slow cooker. Add pork chops and pour apple butter mixture over top.
Cook on low for 5 to 6 hours, covered. Shred pork with two forks in cooker, and let meat marinate on low in sauce for 20 minutes, uncovered. Serve warm over rice or mashed potatoes.
```

Fig. 4.22 search Results 1

```
In [104...    ### No Category Weight
query = ['cream','banana','cinnamon']
Search_Recipes(query, query_ranked=False, recipe_range=(0,3))

Search Query: ['cream', 'banana', 'cinnamon']

Recipe Rank: 0 Banana Cinnamon Roll Casserole

Ingredients:
4 large stale cinnamon rolls; 1 1/2 cups milk; 5 eggs; 2 teaspoons vanilla extract; 2 teaspoons ground cinnamon; 2 bananas, divided; 1/2 cup brown sugar

Instructions:
Slice cinnamon rolls into 1-inch squares. Whisk milk, eggs, vanilla extract, and cinnamon together in a bowl.
Slice 1 banana and layer on bottom of a 9x13-inch baking dish; top with brown sugar. Layer cinnamon roll squares on top of banana. Pour egg mixture over cinnamon rolls; press cinnamon rolls with a spatula to allow egg mixture to fully coat. Cover with aluminum foil and refrigerate until egg mixture is absorbed, 1 hour to overnight.
Preheat oven to 375 degrees F (190 degrees C).
Bake in preheated oven until firm, about 50 minutes. Remove aluminum foil and bake until golden brown, about 10 minutes more. Remove from oven and let stand for 10 minutes. Slice remaining banana and place on top to serve.

Recipe Rank: 1 Cinnamon Banana Cream Cake

Ingredients:
Cake: 1 (18.25 ounce) package vanilla cake mix; 1 cup water; 3 eggs; 1/3 cup oil; 3 very ripe bananas, mashed; 2 teaspoons ground cinnamon; 1 teaspoon baking soda; Frosting: 1 cup milk; 1 (3.5 ounce) package instant banana cream pudding mix; 1 (8 ounce) container low-fat frozen whipped topping (such as Cool Whip® Lite), thawed

Instructions:
Preheat oven to 350 degrees F (175 degrees C). Grease a 13x9-inch baking dish.
Beat cake mix, water, eggs, and oil together in a bowl using a electric mixer until batter is smooth and well mixed, about 2 minutes. Mix bananas, cinnamon, and baking soda together in a separate bowl; stir into batter. Pour batter into the prepared baking dish.
Bake in the preheated oven until a toothpick inserted in the center of the cake comes out clean, 35 to 40 minutes. Cool cake completely.
Beat milk and pudding mix together in a bowl until smooth; fold in whipped topping until frosting is smooth. Spread frosting onto cooled cake.
```

Fig. 4.23 search Results

```
### Only Category Weight
query = ['apple','blueberry']
Search_Recipes(query, query_ranked=False, recipe_range=(0,3))

Search Query: ['apple', 'blueberry']

Recipe Rank: 0 Carrots a la Orange

Ingredients:
1 (16 ounce) package carrots, peeled and sliced; 1 tablespoon white sugar; 1 teaspoon cornstarch; 1/4 teaspoon salt; 1/4 teaspoon ground ginger; 1/4 cup orange juice; 2 tablespoons butter

Instructions:
Bring a pot of lightly salted water to a boil in a saucepan and cook carrots in the boiling water until just tender, about 15 minutes. Drain. Transfer carrots to a serving dish and keep warm.
Whisk sugar, cornstarch, salt, and ginger in a saucepan until thoroughly combined; turn heat to medium and whisk in orange juice. Bring to a simmer, whisking constantly, and cook until sauce is smooth and thickened, about 3 minutes; stir in butter. Pour sauce over carrots, stir to coat, and serve.

Recipe Rank: 1 Cumin and Orange Glazed Carrots

Ingredients:
3 navel oranges; 4 pounds medium carrots (24), peeled; 3 1/2 tablespoons vegetable or olive oil; 1 1/2 teaspoons cumin seeds; 1 3/4 cups water; 2 tablespoons fresh lemon juice; 1 tablespoon sugar; 1 1/4 teaspoons salt; 1/2 teaspoon black pepper; Special equipment: parchment paper or wax paper

Instructions:
Cut out a round from parchment paper to fit just inside a 12-inch heavy skillet (3 inches deep), then set round aside. Cut peel, including all white pith, from oranges with a sharp paring knife. Working over a bowl, cut segments free from membranes, letting segments fall into bowl, then squeeze 1/2 cup juice from membranes into another bowl (discard membranes). Chop orange segments and reserve. Cut a 1-inch-thick diagonal slice from a carrot and set aside, then roll carrot away from you 90 degrees and cut another 1-inch-thick diagonal slice. (Shape will resemble a trapezoidal log. If carrots are very thick, cut slices in half lengthwise as well). Cut up rest of carrot in same manner, then repeat with remaining carrots. Heat oil in skillet over moderate heat until hot but not smoking, then add cumin seeds and cook, stirring, until fragrant, about 1 minute. Add carrots, water, orange juice, sugar, salt, and pepper. Cover carrots directly with parchment round and simmer, stirring occasionally, 30 minutes. Remove parchment.
```

Fig. 4.24 search Results 2

```
### Only Category Weight
query = ['japanese']
Search_Recipes(query, query_ranked=False, recipe_range=(0,3))

Search Query: ['japanese']

Recipe Rank: 0 Crabby Bisque

Ingredients:
3 (10.5-ounce) cans restaurant-style condensed crab bisque (recommended: Bookbinder's); 2 1/4 cups plus 6 tablespoons heavy cream; 3 tablespoons fresh lemon juice; 3 tablespoons chopped fresh parsley leaves; 3 (4.25 ounce) cans crabmeat, picked over; Salt; Cayenne pepper

Instructions:
Watch how to make this recipe.
In a medium saucepan, combine bisque, 2 1/4 cups heavy cream, lemon juice, parsley, and crabmeat with juices and bring to a simmer over medium heat for 5 to 10 minutes. Season soup, to taste, with salt and cayenne pepper. Divide soup among 6 bowls. Add 1 tablespoon of cream to each bowl and swirl with butter knife to create a pretty design. Serve immediately.

Recipe Rank: 1 Chocolate Hazelnut Praline Torte with Frangelico Whipped Cream

Ingredients:
1 recipe Hazelnut Praline; 3 ounces fine-quality bittersweet chocolate (not unsweetened), chopped coarse; 3/4 stick (6 tablespoons) unsalted butter, softened; 1 teaspoon salt; 1/2 cup sugar; 4 large eggs, separated; 1 teaspoon vanilla extract; 1 cup well-chilled heavy cream; 2 tablespoons Frangelico (hazelnut-flavored liqueur)

Instructions:
Preheat oven to 350°F. Butter an 8 1/2-inch springform pan (2 inches deep) or other cake pan with same dimensions and line bottom with a round of wax paper. Butter paper and dust pan with flour, knocking out excess flour. In a food processor pulse praline until ground fine. Reserve 1/4 cup ground praline and transfer remaining praline to a bowl. In food processor pulse chocolate until finely ground and add to bowl of praline. In a bowl with an electric mixer cream together butter, salt, and 1/4 cup sugar until light and fluffy. Beat in yolks, 1 at a time, beating well after each addition, and beat in vanilla. Stir in praline chocolate mixture (mixture will be very thick). In another bowl with cleaned beaters beat whites until foamy and add a pinch salt and remaining 1/4 cup sugar in a steam, beating until meringue just holds stiff peaks. Fold one third meringue into yolk mixture to lighten and fold in remaining meringue gently but thoroughly. Pour batter into pan and smooth top. Bake torte in middle of oven 45 to 55 minutes, or until it begins to pull away from side of pan. Cool torte in pan on a rack (torte will fall slightly and continue to set as it cooks) and remove from pan. To serve may be made 1 day ahead and chilled, covered. Bring torte to room temperature before serving. In a bowl with an electric mixer beat cream until it just holds soft peaks. Stir in Frangelico and half of reserved praline.
Preheat oven to 350°F. Butter an 8 1/2-inch springform pan (2 inches deep) or other cake pan with same dimensions and line bottom with a round of wax paper. Butter paper and dust pan with flour, knocking out excess flour.
```

Fig. 4.25 search Results 3

```

### Only Category Weight
query = ['japanese']
Search_Recipes(query, query_ranked=False, recipe_range=(0,3))

Search Query: ['japanese']

Recipe Rank: 0  Crabby Bisque

Ingredients:
3 (10.5-ounce) cans restaurant-style condensed crab bisque (recommended: Bookbinder's); 2 1/4 cups plus 6 tablespoons heavy cream; 3 tablespoons fresh lemon juice; 3 tablespoons chopped fresh parsley leaves; 3 (4.25 ounce) cans crabmeat, picked over; Salt; Cayenne pepper

Instructions:
Watch how to make this recipe.
In a medium saucepan, combine bisque, 2 1/4 cups heavy cream, lemon juice, parsley, and crabmeat with juices and bring to a simmer over medium heat for 5 to 10 minutes. Season soup, to taste, with salt and cayenne pepper. Divide soup among 6 bowls. Add 1 tablespoon of cream to each bowl and swirl with butter knife to create a pretty design. Serve immediately.

Recipe Rank: 1 Chocolate Hazelnut Praline Torte with Frangelico Whipped Cream

Ingredients:
1 recipeHazelnut Praline; 3 ounces fine-quality bittersweet chocolate (not unsweetened), chopped coarse; 3/4 stick (6 tablespoons) unsalted butter, softened; 1 teaspoon salt; 1/2 cup sugar; 4 large eggs, separated; 1 teaspoon vanilla extract; 1 cup well-chilled heavy cream; 2 tablespoons Frangelico (hazelnut-flavored liqueur)

Instructions:
Preheat oven to 350°F. Butter an 8 1/2-inch springform pan (2 inches deep) or other cake pan with same dimensions and line bottom with a round of wax paper. Butter paper and dust pan with flour, knocking out excess flour. In a food processor pulse praline until ground fine. Reserve 1/4 cup ground praline and transfer remaining praline to a bowl. In food processor pulse chocolate until finely ground and add to bowl of praline. In a bowl with an electric mixer cream together butter, salt, and 1/4 cup sugar until light and fluffy. Beat in yolks, 1 at a time, beating well after each addition, and beat in vanilla. Stir in praline chocolate mixture (mixture will be very thick). In another bowl with cleaned beaters beat whites until foamy and add a pinch salt and remaining 1/4 cup sugar in a stream, beating until meringue just holds stiff peaks. Fold one third meringue into yolk mixture to lighten and fold in remaining meringue gently but thoroughly. Pour batter into pan and smooth top. Bake torte in middle of oven 45 to 55 minutes, or until it begins to pull away from side of pan. Cool torte in pan on a rack (torte will fall slightly and continue to set as it cooks) and remove from pan. Torte may be made 1 day ahead and chilled, covered. Bring torte to room temperature before serving. In a bowl with an electric mixer beat cream until it just holds soft peaks. Stir in Frangelico and half of reserved praline.
Preheat oven to 350°F. Butter an 8 1/2-inch springform pan (2 inches deep) or other cake pan with same dimensions and line bottom with a round of wax paper. Butter paper and dust pan with flour, knocking out excess flour.

```

Fig. 4.26 search Results 4

Testing the Algorithm

```

query = ['cinnamon', 'cream', 'banana']
Search_Recipes(query, query_ranked=True, recipe_range=(0,3))

Search Query: ['cinnamon', 'cream', 'banana']

Recipe Rank: 0  Banana Cinnamon Roll Casserole

Ingredients:
4 large stale cinnamon rolls; 1 1/2 cups milk; 5 eggs; 2 teaspoons vanilla extract; 2 teaspoons ground cinnamon; 2 bananas, divided; 1/2 cup brown sugar

Instructions:
Slice cinnamon rolls into 1-inch squares. Whisk milk, eggs, vanilla extract, and cinnamon together in a bowl. Slice 1 banana and layer on bottom of a 9x13-inch baking dish; top with brown sugar. Layer cinnamon roll squares on top of banana. Pour egg mixture over cinnamon rolls; press cinnamon rolls with a spatula to allow egg mixture to fully coat. Cover with aluminum foil and refrigerate until egg mixture is absorbed, 1 hour to overnight.
Preheat oven to 375 degrees F (190 degrees C).
Bake in preheated oven until firm, about 50 minutes. Remove aluminum foil and bake until golden brown, about 10 minutes more. Remove from oven and let stand for 10 minutes. Slice remaining banana and place on top to serve.

Recipe Rank: 1 Cinnamon Bread I

Ingredients:
2 cups all-purpose flour; 1 cup white sugar; 2 teaspoons baking powder; 1/2 teaspoon baking soda; 1 1/2 teaspoons ground cinnamon; 1 teaspoon salt; 1 cup buttermilk; 1/4 cup vegetable oil; 2 eggs; 2 teaspoons vanilla extract; 2 tablespoons white sugar; 1 teaspoon ground cinnamon; 2 teaspoons margarine

Instructions:
Preheat oven to 350 degrees F (175 degrees C). Grease one 9x5 inch loaf pan.

```

Fig. 4.27 search Results 5

Chapter 5

Conclusion/ Recommendations and Future Works

5.1 Conclusion

In this research we built a food social website completely from zero to match our requirements and customization to the product we tried to Market, More over, We live two Life cycles Software Life Cycle from collecting requirements, Designing, Coding and Testing then deploy.

And the second is Data Science Life cycle from Collecting data, to cleaning, parsing, modeling then choosing the best Algorithm for training. And After that test the model in our real data and building an algorithm for search which been captured from text ranking and query algorithm, we hope to lunch the website to get feedback for end users.

5.2 Future Works

Much can be done as future work regarding our our approach in Search engine for recipes and its use in a Social Media Platform. With the concept of Recipe matching being a fundamental component in daily cooking routine, we aim to develop a more accurate and human-similar recipe detection model. Furthermore, regarding the stages Recipe Search engine, future work includes a more thorough investigation into visual features to increase the system's performance and computation time using DOS and microservices architecture. More importantly, a contribution can be done by collecting a new and higher quality Recipe dataset, and providing its corresponding user-based and golden-standard annotations and of course adding it more accurately from our Platform . Finally, we plan on enhancing the

application by enlarging and improving its Recipe storage repository, and improving Ai search and making it a more user-customized environment. And Adding Arabic Languages to the recipes after collecting and creating Data from the Platform it self after its release

References

- [1] Django (2008). Django best practices. [online] <https://django-best-practices.readthedocs.io/en/latest/>.
- [2] Django documentation (2008). Django documentation. [online] <https://docs.djangoproject.com/en/4.0/>.
- [3] Elzero Web School (2008). Elzero web school channel. [online] <https://www.youtube.com/channel/UCSNkfKI4cU-55Nm-ovsvOHQ>.
- [4] HackerRank (2008). Hacker rank website. [online] <https://www.hackerrank.com/>.
- [JAGITHYALA] JAGITHYALA, A. Recommending recipes based on ingredients and user reviews.
- [Khan] Khan, M. S. J. R. M. K. M. I. B. H. A. M. S. R. H. Food recipe recommendation based on ingredients detection using deep learning.
- [7] Resnick and Varian (1997). recommender systems. 40.
- [8] W3school (2008). W3school website. [online] <https://www.w3schools.com/python/>.

Appendix A

Attachment A

Proposed Disclaimer Statement Format

The report is a document written by the student(s) and should reflect expertise in different aspects of research methodology and technical writing skills. The supervisor's job is to guide the student so that she/he can achieve the objectives in an efficient way while gaining the skills sought. While maintaining credit the disclaimer statement is simply a statement protecting the Department and the University from any legal liability claims associated with the use of the results and the methods presented. Its format is as follows:

DISCLAIMER

This report was written by Ahmad Droobi and Ataa Shaqour at the Computer Engineering Department, Faculty of Engineering, An-Najah National University. It has not been altered or corrected, other than editorial corrections, as a result of assessment and it may contain language as well as content errors. The views expressed in it together with any outcomes and recommendations are solely those of the student(s). An-Najah National University accepts no responsibility or liability for the consequences of this report being used for a purpose other than the purpose for which it was commissioned.