

Proposal & Statement of Work

Capstone Term II AIDI-2005-02

Marcos Bittencourt

January 23rd 2019

Todd McCullough 100593884

Savya Sharma 100770026

Marko Topitsch 100635033

Rahul

Table of Contents

Executive Summary	3
Rationale	4
Problem Statement	5
Ideal	5
Reality	5
Limitations	6
Proposal	6
Data	7
Datasets	7
Data Lists	8
Model Architecture	9
Project Plan	10
Interface	11
References	12

Executive Summary

We have endeavoured to solve what is undoubtedly an age old question; albeit one that has changed somewhat over time due to the rapid improvements in convenience.

What should I eat? This is an all too common problem facing many Canadians daily, morning, noon and night. Even if Canadians prefer to eat home cooked, healthy meals; many feel they have no option but to turn to restaurants or fast food options due to a lack of ideas or time .

As per Statistics Canada, around 54% of Canadians eat out once per week. That's a fairly staggering statistic. And from those Canadians eating out, 40% choose to do so because they don't know how to cook from the available ingredients they have, lack the time or motivation, or simply don't know how to cook. (Statscan Infographic(2019, January 10))

This has an adverse impact on their health and finances. We want to provide people with a way to make food at home easily and quickly by recommending recipes depending on their preferences and what ingredients they have available on hand. Eating out is costly both in terms of time and money. We aim to reduce those costs. Our goal is to inspire people to begin cooking for themselves and their families.

Rationale

There has been a long time running grassroots movement to get Canadians cooking again in their kitchens. It is big business. Anywhere you look, you are bound to find cook books on store shelves, cooking shows on TV or are likely inundated by cooking content on social media.

Yet for all this buzz, Canadians appear to be spending less time in their own kitchens cooking and more time out at restaurants spending money that might be best kept in a savings account.

We want to provide people with a way to figure out what they can eat based on the amount of time they have to cook, the ingredients they have on hand, the type of cuisine they desire, and the nutritional value they are seeking. People all over Canada could be saving money while simultaneously honing a skill everyone should have. That being the ability to cook food for one's self.

As well, they will make an informed choice for eating healthy food. Anyone interested in saving money on eating out while simultaneously becoming more independent will benefit from this.

Problem Statement

Canadians today are strapped for time and money, but they want to eat healthy and have a variety of choices. With the seemingly unlimited options available out there, we propose making an assistant that can help them make the best choice possible when it comes to choosing a meal.

Ideal

The user tells the recommender program what ingredients they have on hand, how much time they have, the style of cuisine they desire, and the nutritional value they want out of the meal. The program then uses the input given to recommend the best meal recipes that best suit the user's choices.

Reality

People do not know what they can cook with the time they have available, and generally need new ideas for what they can make and eat. Currently, Canadians aren't planning meals in advance and when they arrive at the point that they need to eat, they are frustrated by the decisions that need to be made in order to make a meal. They settle on the path of least resistance, choosing often to forego making meals at home, and instead choosing to eat out.

In general, Canadians face these general meal preparation problems;

- Do not know what they can cook with the time they have available - [metric - time]
- Do not always consider the ingredients they have in their home - [metric - ingredients]
- Do not know the nutritional value of what they are eating [metric - nutritional value]

- Are unsure of what regional dishes they may be able to make [metric - cuisine style]

Limitations

Allrecipes.com does not in fact have all the recipes in the world. Some recipes people desire may not be in the database. The algorithm does not consider individual cooking skill level. Perhaps the user doesn't have the best ingredients for the cuisine that they have chosen. Could they be instead recommended a choice with similar or interchangeable ingredients?

Proposal

If someone in Canada wants to eat a meal but isn't sure what they can make with the ingredients on hand, the user can input this information into the application, and will get an appropriate recommendation. Canadians today are strapped for time and money, but they want to eat healthy and want to have a variety of choices. With the seemingly unlimited options available out there, we propose making an assistant that can help them make the best choice possible.

Data

In order to solve this problem, we can gather data from user input and compare it to recipes listed on [allrecipes.com](https://www.allrecipes.com). From this we can create an algorithm to provide meal recipe recommendations to the user.

The parameters we will be working with are;

- the amount of time the recipe takes to prepare
- the nutritional value desired by the user
- the ingredients available on hand
- The style of cuisine they desire

With this data we can build a recommender system in which the user inputs the amount of time they have available, the nutritional value they may be seeking, the type of cuisine they desire, and the ingredients they have available to them. Based on these inputs we will generate a list of recipes that fit their preferences.

Datasets

The data set we retrieved from [allrecipes.com](https://www.allrecipes.com) is the primary and only source of data at this period in time.

Allrecipes.com, Inc. is a food focused social network, which generates revenue from advertising. Users of the site post recipes that are tagged based on many different categories.

Data Lists

1. Requirements for data eligibility for our purposes must include a list of ingredients, the nutritional value of the recipe, cuisine type, and a rating or user reviews.
2. Note: Some of your data requirements can include engineered features. I.e., “dataset must include the number of days alive on planet”. This data point is almost never included in a dataset, but can be easily created if you know the person’s date of birth.
3. Our Main dataset will be from allrecipes.com. It is a very large dataset and should suffice for our purpose. If we could figure out a way to incorporate data from other recipe websites that would definitely be beneficial for the efficacy and strength of the recommender system.
4. We’re assuming that the 60,000+ recipes from allrecipes.com will be sufficient to develop a strong recommender system. A limitation of this data set is that it may be missing some recipes people may want.

Model Architecture

We are building a recommender system using collaborative filtering.

We will supplement this with content based recommendation to ensure quality items in the long tail get recommended as well and not lost in the data.

Collaborative Filtering

Collaborative filtering is one of the most, if not the most widely used method of recommending. It does not require any domain knowledge since it learns all it needs solely from the user-item-rating relations. It is serendipitous (it can return items the user likes, but never would have thought of). The downfalls of the algorithm come from its cold start problem (a difficulty recommending for new users or items) and the lack of constraints you can put on the result through the algorithm. If the user wants only one general type of recipe, you will need to programmatically remove all the unwanted types before the recommendations make it to the user.

Content Based Recommendation

Content Based Filtering has some benefits over other algorithms. While it does require data from the user seeking the recommendation, it does not require data about the other users. Scalability is a key factor, it is simple to scale the process to search through larger amounts of users. However, it does often require data engineering. It is also not a serendipitous recommender; it will not return items the user would like, but never think of.

Project Plan

PHASE	DETAILS												
		JAN		FEB				MAR					
		27	3	10	17	24		2	9	16	23	30	
1	Asses Project Needs		4										END
	Evaluate Existing Model		2										
2	Identify Website Needs		4										
	Organize & Design Database		4										
	Train model/architecture			8									
	Test model/architecture				8								
	Refine Python Code				8								
	Refine Model/architecture					8 hrs							
3	Design Database					8 hrs							
	Design Website					12 hrs		12 hrs					
	Evaluate model/architecture						4						
	Refine model/architecture						4	8 hrs					
4	Develop Final Report							10 hrs					
	Final Report								6				

Interface

Users will access the recommender system through an online or offline web-portal. In order to make recommendations, the user will need to be logged in, and have previously rated recipes in the database.

There however may be the possibility to make a guest account that can be used as a default, that takes from the highest rated recipes. This concept will be explored further to access if this is viable.

The interface itself, while yet to be designed, will match the esthetics of all the major cooking/recipe influence sites on the internet.

The user will have the ability to narrow the search results down with filters in the following categories. The number of these could potentially increase;

Cuisine	Meal type	Ingredient	Diet Restrictions
Canadian	Appetizers	Milk	Vegetarian
Chinese	Breakfast	Fish	Vegan
Italian	Dinner	Cheese	Paleo
Indian	Dessert	Pasta	Low-Carb
French	Drink	Chicken	Food Allergies
Mexican	Salad	Broccoli	Atkins
etc	etc	etc	etc

At this point in time, the recommender chooses all recipes from users that share similar rating characteristics. The filters then take that data and filter it down further to get the results that the user is looking for.

References

Statscan Infographic (2019, January 10). Retrieved from

<https://www150.statcan.gc.ca/n1/pub/11-627-m/11-627-m2019003-eng.htm>