# Practical Work
# Recommendation System - Food Recommendation

Intelligent Data Analysis Applications in Business

Universitat Politècnica de Catalunya
Faculty of Informatics

26/01/2017

Erol Kazancli

# INTRODUCTION

Food is one of the best pleasures of life. We, in general, don't buy clothes everyday, at least unless we are super rich or addicted to shopping. However, we normally eat three times a day. This is not only to survive but also to experience pleasure. Many of us like cooking, trying new recipes, going to a new restaurant, trying a new product on the supermarket. I, personally, enjoy cooking programs, seeing a new recipe on my facebook page, discovering good and cheap restaurants and I believe there are many people who have the same affection for food.

We know of many websites or mobile applications to recommend a restaurant, or a recipe, or a product, but I believe food is food, whether we eat it in a restaurant, we make it ourselves at home or we buy it from a supermarket. Therefore, I propose a recommender system which will bring all "food-experience" activities under one roof with cleverer approach to adapting to a specific user.

It is possible to make people's lives better by making the right recommendations to the right person. Many of us, who enjoy the experience of discovering new tastes, need recommendations for authentic food. Many of us, who enjoy cooking at home, need new and preferably practical recipes, to break the monotony of cooking a handful of same recipes every week. Some of us, who do not cook, would be glad to do it, if they knew how to. And many of us, who likes saving, needs easy and practical delicious food, would like to know offers in restaurants and supermarkets. Each one of us fits one of these definitions or a mixture of them.

Such a system would would bring dynamism to the food sector by matching the demand and supply. How many times did you change your mind about cooking and settled for a sandwich, because you didn't come up with a simple and delicious recipe? I bet there were times you were unsure whether to cook at home or eat in a restaurant because you were trying to make ends meet. I am pretty sure there are very good restaurants with affordable prices that have difficulty in attracting the right kind of people. So it is possible to make people happier and also make the food sector better and more competitive.

# DESIGN

This recommender system is designed for the consumers and producers alike. The function is to find the correct matching between a consumer and a product, whether it be a menu(or a dish) in a restaurant, a food item in a supermarket or a recipe from a cooking web-site. As long as they are related to food, these options can be further extended. The aim is to bring dynamism to the food sector, both in the benefit of consumers and producers.

The system can be developed as a web application, a mobile application or a mixture of both. Mobile version would be handy for the consumers especially, since they can check anytime what recommendation is available, even when they are on the move. The web version will be an alternative for those preferring the comfort of PCs.

There should be two registration types: Registration for producers and registration for consumers. Producers might be required to pay a certain fee, since they would be benefiting from using it in terms of increases in the sales of their products. Other option would be letting them use the system for free, using only advertisements for the maintenance of the system. However, for the convenience of consumers, it

would be a better option to keep the system advertisement free. As for consumers, they should not be expected to pay for the usage of the system. After all, it is a better strategy to attract as many consumers as possible for the producers to be able to make their products known. We need to keep in mind that there needs to be sufficient incentive on both sides to join the system and it is logical to expect the financial burden should be carried by the producers, since they will be more than recovering it with increased sales.

Once the producer is registered, he can enter the information about the product he/she is offering. In case of a restaurant, they can enter a dish, a drink, a dessert, a special menu, with a normal price or a special offer price. Once they enter an item, they will be required to answer several questions regarding the product. The questions could be the ingredients of the food, the price, if it is a special offer or not, the price before the offer, etc. These properties of the food later will be used to recommend this product to specific user. As for a product entry from a supermarket, the price, the ingredients, the brand, the offer and the previous price should be entered. Likewise for a recipe from a website, the ingredients, the general preparation time, a category(dessert, savory, main dish, etc.), origin (Chinese, mexican, etc.) will be asked from the owner.

It is important to make the entry of the information for the producers as simple as possible, to encourage the utilization of the system. There should be possible ingredient sets already listed as check-boxes to ensure easy entry and standard entry - to prevent misspellings. If a specific ingredient is not present in the list, this ingredient might be indicated in the “other” section.

The entry of false information should be discouraged. There should be a mechanism ensuring the producers to enter information in an honest way and not lie about their products. After all they have interest in entering correct information since if they lie, this will be reflected in the ratings from the consumers and therefore in the recommendations. If a producer is detected as giving false information intentionally, he might also be punished by being banned from using the system. This is important to give the correct recommendation as well as the accuracy of the information. It is very critical to ensure that a consumer trusts the system in terms of the relevance of the recommendation and the accuracy of the information, otherwise it will not be possible to attract the consumers to the system, since they will not believe in the recommendation and the content given.

As for consumers, they should be allowed to enter their preferences to the system. The preferences can be specific ingredients, interest in offers, recipe preparation times, specific categories, preferred brands, price ranges, etc. They should be free not to enter any information but they would be willing to enter their preferences if they want to fully benefit from the system. The user will also have an option to state the properties he does not like. This way, recommendations having these not-preferred properties might be ruled out. Because the system gets preferences beforehand, it will not be a cold-start recommendation system, i.e. it will have certain information to make the first recommendations. In all these respects, it will have a personalized aspect.

We cannot expect a user to remember or know everything he/she likes and enter it. He might discover his likes as he tries new products (novelty). There is also the factor of serendipity, discovering new unexpected things. Therefore, the system needs to learn user preferences not stated as the he/she uses the system. The user should be given the option to like a certain product, menu or recipe just by the description of it. Or he/she might like it, dislike it or find it misleading because he tried it. The degree of like/dislike will be indicated by the person with a rating mechanism. The user might also implicitly show

interest just by clicking on the item to read the description. This will also be taken as a clue in the recommendations. In these respects, the system will be a content-based recommendation system.

Each product, recipe or meal will be represented by a feature vector. The feature vector may contain a category type(menu/dish, product or recipe), a feature for each ingredient (0 or 1 according to the existence of the ingredient), price for products and menus, expected price, simplicity level, preparation time for recipes, general rating, brand name for a product, a type (savoury, sweet, sour), etc.

As can be seen the feature vector changes across categories, but this will not pose a problem since a product will only compared to a product, a menu to a menu, and a recipe to a recipe. There is no need to compare a recipe to a menu because recommendations will be made separately from each category in varying amounts. This amount may be deduced from the interest shown by the user for a specific category. For instance, if the user takes a lot of time reading recipes but shows no interest for a menu in a restaurant then the number of recommendations for recipes will be higher. This preference can also be supplied explicitly by the user.

Although the cross-category comparison will not be made, clues might be obtained from one category to make recommendations from another category. For example, if a person does like a recipe, a menu or a dish from a restaurant containing similar ingredients might be recommended to the user, taking into account also the other preferences.

The recommendations will be made in three ways or a combination of these three. The first, which is also the first to be used as a user registers, will be from the preferences stated. The product, recipes or dishes will be searched based on these preferences. Later as the system warms up with usage by a user, it will have a sufficient database for likes and dislikes. Thus, the second mechanism will come into play, which is comparing other items to the items liked or disliked (item-based), using the feature vectors. Here, a certain level of learning is taking place. The third mechanism (user-based) will be by comparing the user to other users. This comparison will be made from the ratings, likes, or specified preferences. Thus, similarities between users will be found and new recommendations will be made based on others' likes. This comparison might be rather complex and suffers from sparsity, and can be used only if the other mechanisms are not sufficient. In this respect, the system will have collaborative aspect. For item-based similarities, euclidean distance can be used with possible weights associated to features. For user-based similarities, pearson correlation will be used since it will be based on mainly rankings. Normalization of the ratings to the user domain will also be performed.

The recommendations will not always be done based on preferences or similarity. There will also be other partly-random recommendations, comparably less in number, to increase variety and also to allow for nice surprises for a user, which is also called serendipity. This way the user might learn something new about himself. These recommendations can be chosen from popular ones but also not popular products using the general ratings, since popular ones will be already enjoying increased interest, and not popular ones needs some publicity. However, since explainability is important, there should be a reasoning behind these random recommendations.

Whichever of the mechanisms explained above is used, there will be common rules, that will not be violated in any circumstances. First of all, the non-preferred ingredients, should never be in a recommendation. Secondly, there might be cultural, religious or personal sensitivities, such as not eating pork, being vegetarian, etc. These should be taken into account and such products should never be recommended.

From a transparency perspective, the recommendations should be explainable, even the random ones. The user should be convinced that the recommendations are good and relevant, and are not done in anyone's financial interest. This is important for the loyalty of the user and, therefore, the viability of the system.

As for the reliability and robustness of the system, the users have to be genuine users. There should not be fake attempts on the part of producers, to register as consumers and puff the ratings of their items up. There should be mechanisms to deter such action, or detect it with abnormal behaviours as clues. Another method could be disregarding a certain amount of high and low ratings, which could be dubitative. Although ratings are not the main mechanism in this system, high ratings are used when making user-based recommendations. Also, the users can see the general ratings of the recommended products and may base their decision on whether to try the product on these ratings. Therefore, it is important to make these ratings accurate.

The evaluation will be based on the click counts to top-N recommended items or the ratings of these top-N items. If these items receive sufficient clicks or enjoy high-ratings by the recommended user we can conclude the recommending mechanism is functioning at least in finding relevant items (accuracy). The diversity can be evaluated by the increase in the interest in non-popular items, since they are included in recommendations as well. The trust of consumers can be evaluated by simply looking at consumer loyalty, the leaving and the entering ones. Robustness can be kept in check by monitoring abnormal behaviours in the system. The serendipity and the novelty factor can be evaluated by feedback mechanisms that take opinions on customer satisfaction. If any new feature will be added to the system, A/B testing can be used to see the effectiveness of the new feature. Two small representative and similar portions can be selected from the consumer base, and they can be compared using the version with/without the new feature.

## CONCLUSION

The system recommended will be reliable, transparent, creative and intelligent enough to be a good recommender system. It will both benefit the consumers and the producers, allowing consumers to access what they are interested in and allowing producers to make their offers available to interested consumers. There will be enough incentives on both sides to join the system. Evaluations of accuracy, coverage, trust, novelty, serendipity, diversity will all be carried out to ensure viability. The system will be reliable because it will be based on similarities and rankings, not on specific financial interests. It will not allow fake ratings due its nature, because both parties will lose from false declarations. It will be transparent to be able to explain the reason behind recommendations. It will be creative to diversify the recommendations, recommend partly-random items and, therefore, allow for serendipities. It will be sensitive to cultural, religious or personal sensibilities. The system will be intelligent enough, to learn the preferences of consumers to make item-based recommendations. It will also benefit from similarity of users to make user-based recommendations. For the reasons stated it will be both personalized, content-based and collaborative(a hybrid-system). Such a system may benefit from the full power of different types of recommendation systems.

Such a system would find offers for offer seekers, new dishes for experience lovers, simple recipes for novices and creative recipes for cooking experts. It will make the food sector more dynamic, making the producers and consumers meet with mutual gains. It will make the sector also more competitive, since attracting consumers with creative products and nice offers will be the rule.