I Came, I Cooked, I Conquered:

Recommending personalized recipes to minimize food waste

Adriel Saporta asaporta@stanford.edu

Claire Pajot cpajot@stanford.edu

Jenia Borisenko evgeniya@stanford.edu

Jiang Peicun 11jiang@stanford.edu

Stanford University

Introduction

Background

We are a group of Stanford students passionate about food.

Our project is a small but meaningful step towards fighting the global food crisis by helping consumers better manage perishable items in their kitchens.

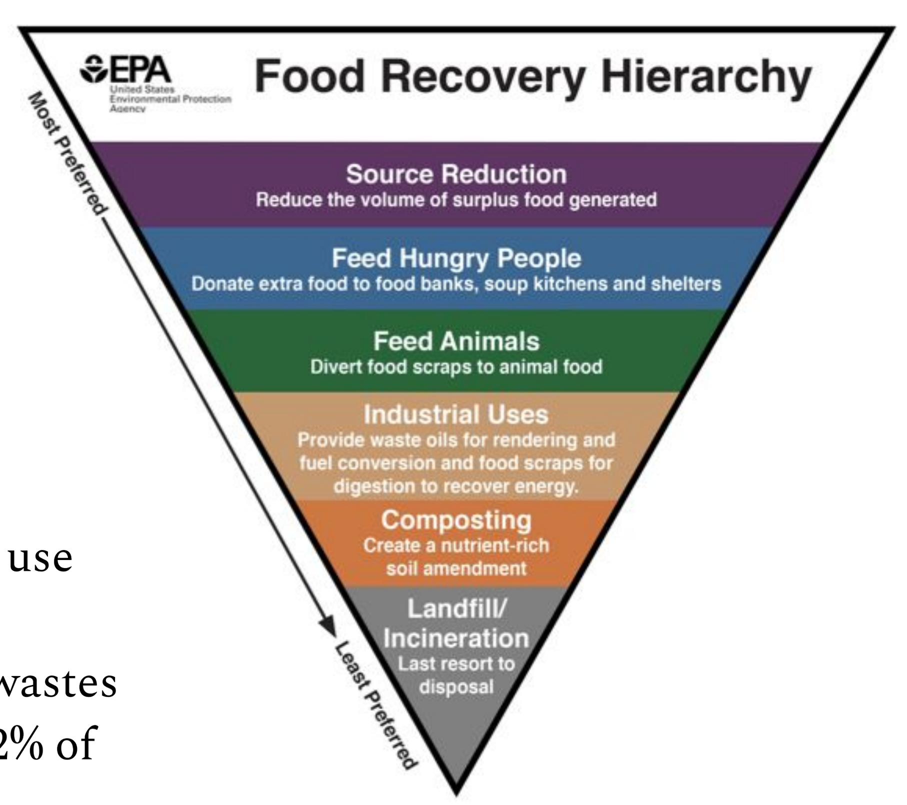
Social Impact

- Our project recommends recipes tailored to users' tastes and preferences while making use of ingredients that they already have.

"Be mindful of old ingredients and leftovers you need to use up. You'll waste less and may even find a new favorite dish. Shop in your refrigerator first! Cook or eat what you already have at home before buying more."

-Environmental Protection Agency

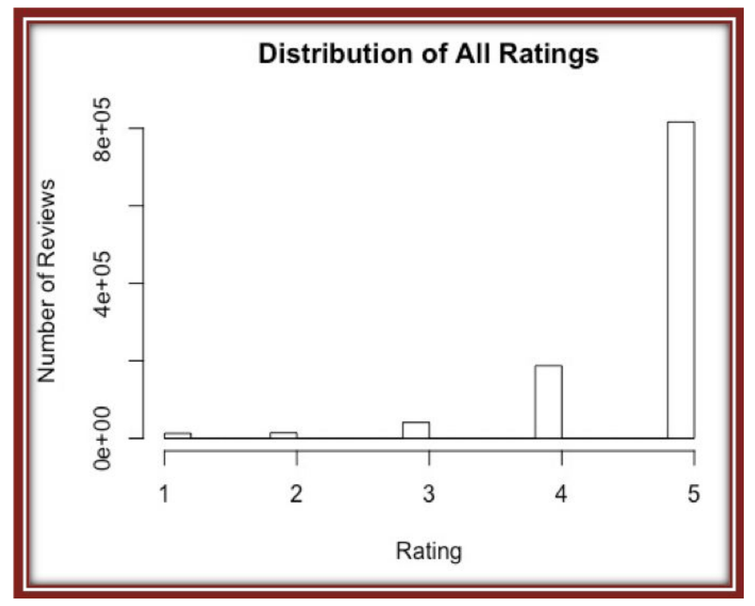
Prevention (reducing amount of uneaten food) *VS* ***Disposal*** (discarding uneaten food generated)

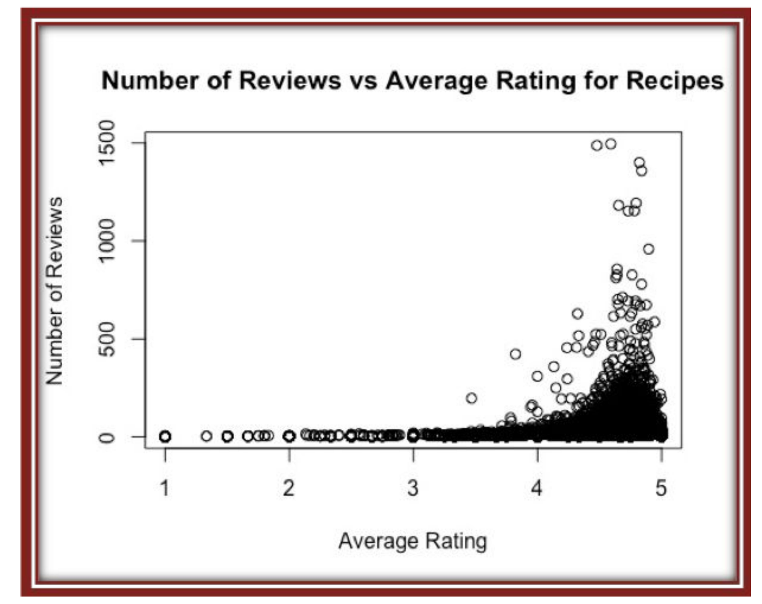


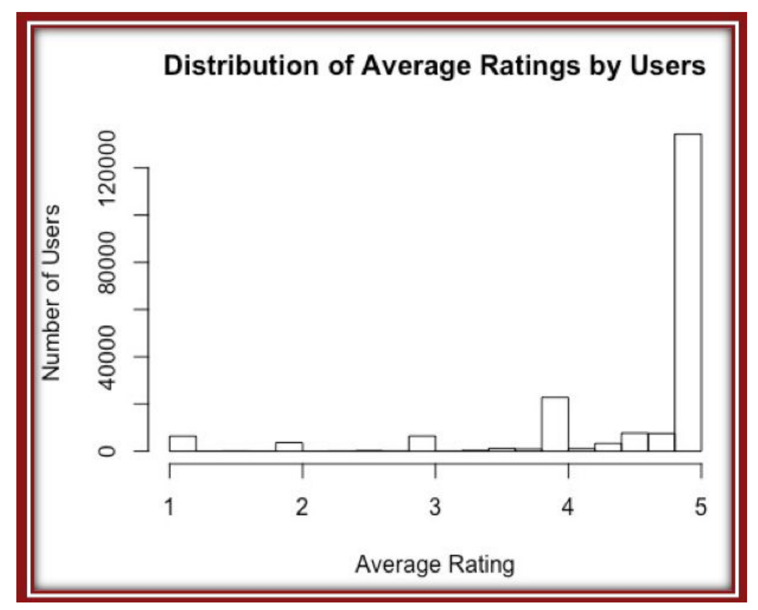
- By focusing on prevention, our project empowers users to adopt food management practices that collectively leads to the greatest possible reduction of food waste.
- Consumers are also incentivized to use our product for economic reasons.
  - USDA: the average American wastes $371 worth of food per year, 9.2% of their food spending
- The EPA highlights the importance of consumer education; our project indirectly teaches users how to use food items they may otherwise discard.

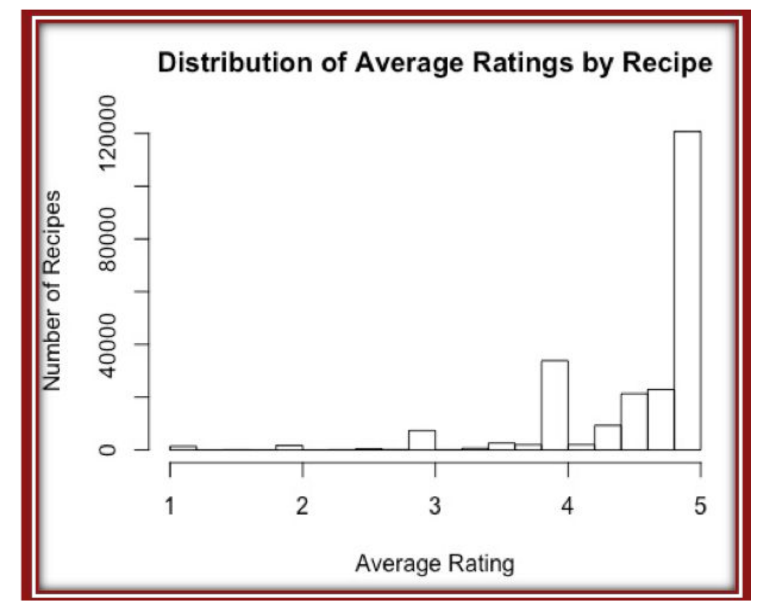
Data

- Source: database from Food.com's online recipe aggregator
- Two datasets: *recipes* (ingredients and tags like "health") and *interactions* (users' recipe ratings on 0-5 scale and comments).
- Removed all reviews with rating of 0 to reduce noise in our data since users often rate 0 when not providing a meaningful rating.



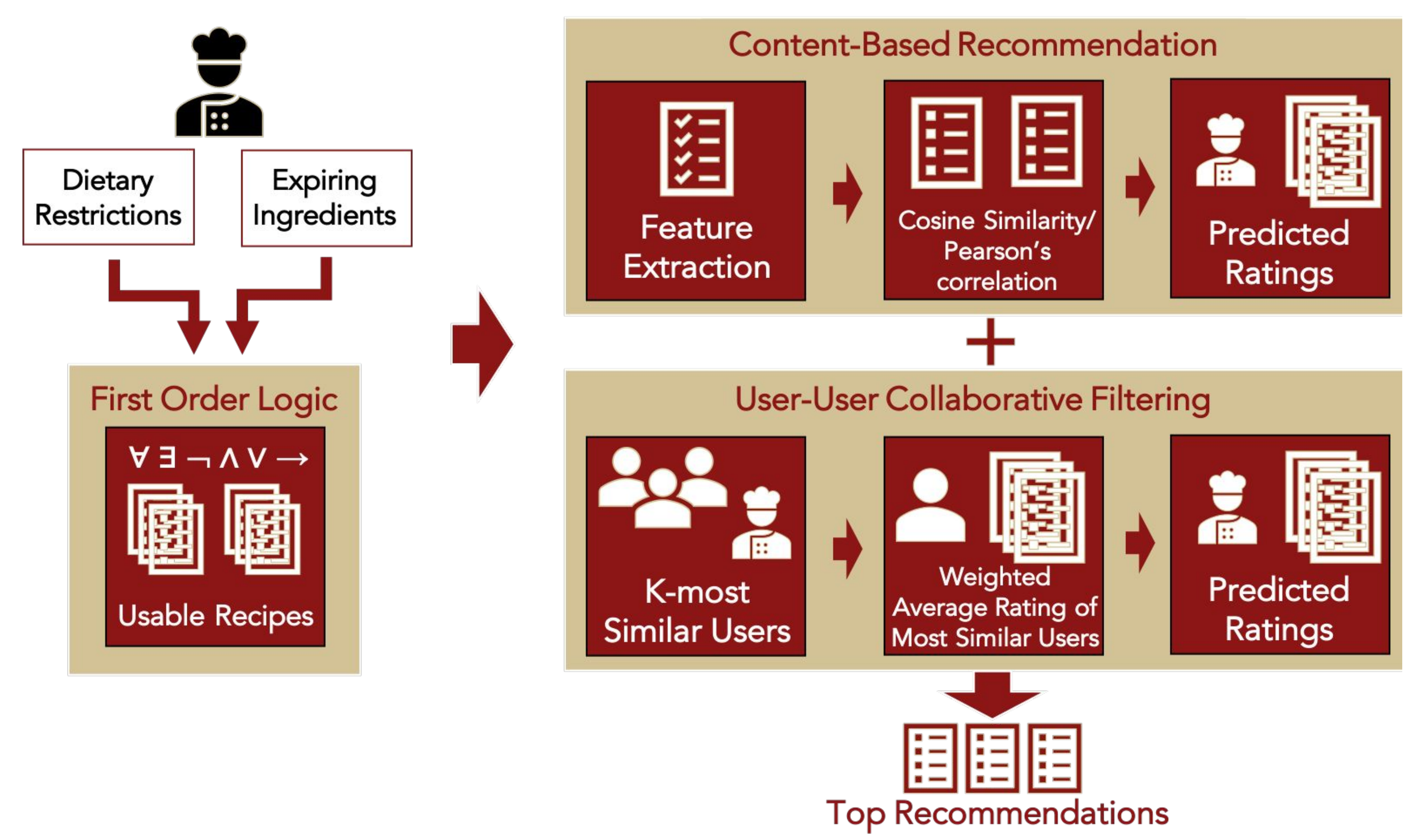






Algorithms and Models

Final Architecture



Model Explanation

- **Overview:** We consider recipes that use as many of the user's ingredients as possible and satisfy any dietary restrictions. We then predict the rating for each candidate recipe, and recommend the 10 recipes with the highest ratings.
- **Rules-based/First Order Logic:** We first determine the set of recipes that could be used. This is done with simple first-order logic to derive the set of recipes that meet dietary restrictions and use the most expiring ingredients.
- **Collaborative Filtering:** We started with a user-user collaborative filtering algorithm to predict user ratings. This takes the set of k users who have rated a recipe and are most similar to our target user. We compute an average of those k users' ratings, weighted by similarity. We experimented with similarity measures (i.e. cosine similarity, Pearson's correlation) and data transformations.
- **Content-Based Recommendation:** We defined features for recipes and compared recipe profiles to user profiles (users' "ideal" recipes) derived from their highly rated recipes. As with our collaborative filtering algorithm, we experimented with a variety of similarity measures for our system.
- **Final Recommender:** We ultimately used a linear combination of content-based recommendation and collaborative filtering to predict ratings. We ran tests to choose the parameter for our final model.

Results & Analysis

Results

- Computed RMSE for 10 tests with 50 random held-out user ratings each
- Baseline: average rating for each recipe and average rating over all recipes
- Our algorithms performed worse for a random held out rating
- However, this does not necessarily mean our algorithm performs worse for a random user-recipe pair. This is due to the inherent bias in our test set.

Algorithm / Evaluation	Collaborative Filtering	Linear Combination	Average Recipe Rating	Average Overall Rating
Mean RMSE over 10 tests	0.8301877	0.8245097	0.7608214	0.784331

Challenges

- Our model performs poorly compared to baseline due to the sparsity of dataset.
- For content-based recommendation, it is not obvious how to best translate cosine similarity or pearson correlation between a recipe and a user to a meaningful rating for this data set.
- There is inherent bias in our test set. Since people are generally good at picking recipes that they will like, they choose recipes that they end up rating highly.

References

- Leskovec, Rajaraman, Ullman. *Mining of Massive Datasets.* Cambridge University Press, 2016
- United States Department of Agriculture, Economic Research Service, *Economic Information Bulletin Number 121*, February 2015
- Food and Agriculture Organization of the United Nations, 2011
- United States Environmental Protection Agency, Reducing Wasted Food At Home
- United States Environmental Protection Agency, Sustainable Management of Food