

Project Proposal: Can You Cook It?

January 24, 2025

1. Team Members

- Robert Lam
- Natham Hum

2. Overview

The goal of this project is to create an NLP model that can assess the difficulty of a recipe based on its description. By evaluating factors like the number of ingredients, the complexity of the steps, and the required appliances, the system will provide a difficulty rating that can help users determine whether a recipe is suitable for their skill level. This system can benefit both novice and experienced cooks, helping them choose recipes that align with their capabilities, saving time and reducing frustration. In turn, it encourages people to try cooking more often, build their culinary skills, and feel more confident in the kitchen.

The significance of this task is rooted in the common practice of recipe authors advertising their dishes as "easy" or "simple" to attract a broad audience. While these labels are meant to appeal to beginner cooks, they often lead to frustration when the recipe turns out to be more complex than expected. Novice home cooks may attempt these "easy" recipes only to find they require more skill, time, or multitasking than anticipated. This misrepresentation can discourage people from cooking altogether. If there were a straightforward way to assess the true difficulty of a recipe before trying it, beginners and expert home cooks could better identify recipes that match their skill level, or present a challenge, encouraging them to gain confidence and skill in the kitchen.

This NLP project is challenging because it requires assessing a variety of overlapping factors, including ingredient count, step complexity, appliance use, margin of error, and multitasking, all of which contribute differently to a recipe's difficulty. On top of this, recipes are written in diverse ways, with some condensing many steps into one, while others break down even the simplest tasks into multiple steps. These inconsistencies in how recipes are structured can make it difficult to standardize difficulty assessments. Additionally, the varying importance of these factors, depending on the recipe, further complicates the task of accurately determining a recipe's true difficulty.

3. Task Definition

- **Type of Data:** The dataset we will be using contains around 1000 online recipes written by various authors on different foods. We will provide annotators with the URL to the online recipe, allowing them to review the full content.
- **Task Type:** Regression on a continuous scale of 1-5

4. Data Sources and Plan for Data Collection

Data Sources: We will be using the following dataset from kaggle.com:
<https://www.kaggle.com/datasets/thedevastator/better-recipes-for-a-better-life>.

Authors: Roman Negri, Brad Yuan, Ara Mangoyan, Ricky Chu.

This document is specified as CC0: public domain.

The authors of this database request that if the data is used for research the authors are credited.

Plan for Data Collection: Outline how you plan to collect the data and handle any challenges.

5. Dataset Description

Expected Size: Provide the estimated number of data points (e.g., several thousand).

Example Data Points:

1. Example 1: Data Point and Assigned Label
2. Example 2: Data Point and Assigned Label
3. Example 3: Data Point and Assigned Label

6. Team Contract

1. Team Purpose or Mission

Our team's purpose is to collaborate effectively and efficiently to achieve the objectives of our project. We aim to create high-quality deliverables, meet all deadlines, and support each other throughout the project's learning and development process.

2. Duties/Roles and Expectations for Team Members

Team Members' Responsibilities:

- Contribute equally to the project as a whole (delegation is allowed as long as all members agree).
- Complete assigned tasks on time and communicate progress regularly.
- Actively participate in meetings and discussions.
- Provide constructive feedback and support to other team members.

3. Leadership, Facilitation, and Management

- Leadership will be collaborative and based on each member's skill set.
- Decisions will be made collectively, with majority agreement when necessary.
- Project timelines and task allocations will be tracked using Discord to ensure accountability.

4. Conflict Resolution

Conflicts will be addressed promptly and respectfully. The team will:

1. Meet to discuss the issue and allow each member to share their perspective.
2. Identify the root cause of the conflict and work collaboratively to find a resolution.
3. Seek help from a mentor, instructor, or mediator if the conflict cannot be resolved internally.

5. Additional Norms and Ground Rules

- All team members will attend scheduled meetings unless there is a valid reason for absence, or if it is agreed beforehand by the group member.
- Feedback will be given and received constructively and respectfully.
- All team members will uphold academic integrity and ethical standards.
- Each team member will regularly review the project's progress to ensure alignment with goals.

Acknowledgment and Signatures

We, the undersigned, agree to abide by the terms outlined in this contract and to work together to achieve the project's goals.

Name	Signature
Nathan Hum	
Robert Lam	
Omer Karo	