<u>**Research Proposal:**</u>

Classifying AllRecipes.com Recipe Comments as Related to Specific Kinds of Errors and Providing a Metric for the Recipe that Summarizes Whether It Needs Review

<u>Introduction:</u> Home cooks of all skill levels frequent AllRecipes.com for meal inspiration, but the quality of user-submitted recipe content on the site varies widely. I have identified five kinds of errors (related to incorrect temperature, incorrect cook time, missing or misleading steps, incorrect measurements, and missing or superfluous ingredients) that recipes might suffer from. These errors are generally evident in the comments related to a recipe, although different users might highlight different errors or propose alternative solutions to the same error. I believe these comments can be classified according to the kinds of errors they suggest the recipe suffers from, and these classifications could be quantified into a metric that informs stakeholders of which recipes need review in order to improve the quality of the site's content.

<u>Relevant Research:</u>

- In **"Efficient Comment Classification through NLP and Fuzzy Classification,"** the authors compare the performance of supervised and unsupervised machine learning methods on classifying how negative or positive an Amazon product comment is based on its text, which may differ from how the same comment would be perceived based solely on its star rating. They prefer the unsupervised methods, which had higher precision and recall during their evaluation, and which appear more capable of handling language complexities like sarcasm.
  - Authors: Shubham Derhgawen | Himaja Gogineni | Subhasish Chatterjee | Rajesh Tak
  - Published in: *International Journal of Trend in Scientific Research and Development*, Volume 4, Issue 3, April 2020, pp. 1000-1006
  - URL: https://www.ijtsrd.com/papers/ijtsrd30758.pdf
- **"Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis"** explores the complicated reality that words that are typically assigned positive and negative sentiments are actually frequently used in neutral phrases/contexts. The authors find that when these neutral phrases can be identified and excluded, the accuracy of the polarity assessments of the remaining phrases is greatly improved. They also find that assessing words of negation and modification are very important to accurately assessing phrase polarity.
  - Authors: Theresa Wilson | Janyce Wiebe | Paul Hoffmann
  - Published in: *Computational Linguistics*, Volume 35, Issue 3, September 2009, pp. 399–433
  - URL: https://doi.org/10.1162/coli.08-012-R1-06-90
- In **"ADAPT at IJCNLP-2017 Task 4: A Multinomial Naive Bayes Classification Approach for Customer Feedback Analysis,"** the authors classify Microsoft Office customer feedback as either a comment, a request, a bug, a complaint, meaningless, or undetermined using three different methods: two of them based on multinomial Naïve Bayes Classifiers and one based on a machine learning sentiment analysis method. The modified multinomial Naïve Bayes Classifier outperforms the classic version of the same method, as well as the machine learning sentiment analysis method, based on both precision and recall.

Problem Statement: Classify the kinds of errors recipes on AllRecipes.com potentially suffer from based on user comments and score recipes to indicate whether they need review based on the quantity of these error classifications.

Objectives: Highlight a subset of recipes on AllRecipes.com that could benefit from review so that the quality of the site's content could be improved. Further, each recipe highlighted for review would have its potential errors classified so that only a partial review would likely be necessary. This reduces an impossible problem—a full review of every AllRecipes.com recipe—into a much more manageable problem.

Methodology:

- Scrape recipe content and top comments from AllRecipes.com
- Label enough data for training, validating, and testing
- Prepare the comment text data for Natural Language Processing by eliminating stopwords and punctuation, but not numbers, and transforming the remaining unigrams into a logical matrix
- Train multinomial Naïve Bayes Classifier and Support Vector Machines models on the training data and tune them on the validation data
- Select the final model and evaluate its performance on the test data

Evaluation Metrics:

- Precision
- Recall
- F-score

Assumptions: A single comment can suggest a recipe suffers from multiple kinds of errors, so I will have to use any-of classification. In short, this means creating and running a classifier for each kind of error instead of one multinomial classifier for all kinds of errors. Recipes can also be correct without being good, so there will be some low-rated recipes for which the classifier assigns no error labels. There might also be high-rated recipes for which the classifier assigns several error labels, perhaps because users are reading the comments, making their own adjustments, and reviewing the recipe based on how it turned out with the adjustments as opposed to how it would have turned out without them. Lastly, a recipe that has already been corrected might still receive error labels if newer comments haven't yet replaced the ones that resulted in the corrections.